

UNIVERSIDADE ESTADUAL DE CAMPINAS



JÚLIO CÉSAR BORGES

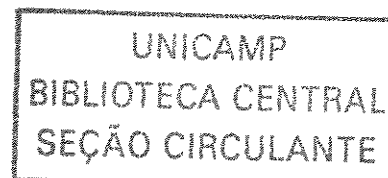
ESTUDOS ESTRUTURAIS DO SISTEMA CHAPERONE MOLECULAR HSP70 HUMANO

Este exemplar corresponde à redação final
da tese defendida pelo(a) candidato (a)
JULIO CESAR BORGES
Carlos Henrique Inácio Ramos
e aprovada pela Comissão Julgadora.

Tese para apresentação ao Instituto de
Biologia da Universidade Estadual de
Campinas, para obtenção do título de
Doutor em Biologia Funcional e Molecular
na área de Bioquímica.

Orientador: Prof. Dr. Carlos Henrique Inácio Ramos

Campinas, 2004



NIDADE	PC
CHAMADA	
UNICAMP	
B644e	
/	EX
COMBO BC/	61434
PROC.	16-86-05
C	☐
D	☒
PREÇO	21,00
DATA	04-1-05
Nº CPD	

Bibid 336796

FICHA CATALOGRÁFICA

B644e

Borges, Júlio César.

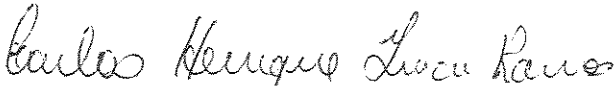
Estudos estruturais do sistema chaperone molecular HSP70 humano / Júlio César Borges. -- Campinas, SP : [s.n.], 2004.

Orientador: Carlos Henrique Inácio Ramos.
Tese (doutorado) - Universidade Estadual de Campinas, Instituto de Biologia.

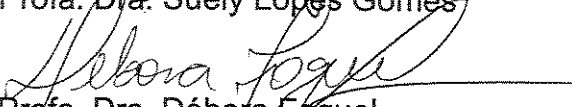
1. Proteínas - Estrutura. 2. Bioquímica. 3. Biofísica.
I. Ramos, Carlos Henrique Inácio. II. Universidade Estadual de Campinas. Instituto de Biologia. III. Título.

Data da Defesa: 04 de Outubro de 2004

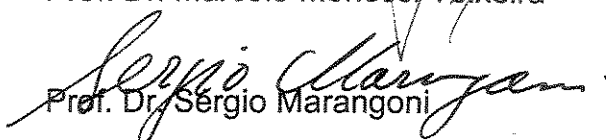
Banca Examinadora


Prof. Dr. Carlos Henrique Inácio Ramos (Orientador)


Profa. Dra. Suely Lopes Gomes


Profa. Dra. Débora Foguel


Prof. Dr. Marcelo Menossi Teixeira


Prof. Dr. Sérgio Marangoni

Profa. Dra. Íris Concepcion Linares de Torriani

Prof. Dr. Jörg Kobarg

AGRADECIMENTOS

Esta foi a parte mais difícil de escrever, pois há tantos a agradecer... Primeiramente, devo agradecer a Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP – pelo apoio financeiro que possibilitou a manutenção da minha vida e pela clareza de pensamento em prol do desenvolvimento científico. Ao Laboratório Nacional de Luz Síncrotron – LNLS – pela ótima infra-estrutura para os diversos experimentos realizados. Ao Conselho Nacional para o Desenvolvimento da Pesquisa – CNPq – vinculado ao Ministério de Ciência e Tecnologia – MCT – pela manutenção do LNLS. Ao Portal Periódicos mantido pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES – pelo acesso aos mais variados títulos de revistas científicas.

Meus agradecimentos ao Prof. Dr. Carlos H.I. Ramos pela oportunidade e orientação, pelas discussões científicas, sugestões nos experimentos realizados, relatórios e artigos científicos e pelo bom convívio nestes anos de trabalho. Tenho muito a agradecer ao Mr. Hannes Fischer e ao Prof. Dr. Aldo Craievich pelo auxílio no tratamento e interpretação dos dados de SAXS das GrpE#1 e Hsp40s. O mesmo deve ser feito em relação ao Mr. Cristiano L.P. Oliveira e à Profa. Dra. Íris Torriani no tratamento e interpretação dos dados de SAXS para as proteínas GrpE#2 e Hsp70.

Um grande agradecimento deve ser feito à Veruska Mazon Soares que foi como uma “fada madrinha” no auxílio nos experimentos e na obtenção dos mais variados materiais necessários para os trabalhos. Aos demais pesquisadores, técnicos e funcionários do LNLS que forneceram as condições adequadas para o desenvolvimento dos trabalhos. Aos vários bons amigos que participaram de maneira direta e indireta desta longa caminhada e àqueles que se não ajudaram, também não atrapalharam.

Devo dedicar um agradecimento especial aos inúmeros professores que influenciaram, direta ou indiretamente, na minha formação acadêmica e científica, que foi em grande parte de cunho público. Aos meus professores no ensino fundamental e médio, meus agradecimentos. Àqueles do curso de Farmácia da Universidade Federal de Minas Gerais – UFMG – sejam eles no ciclo básico ou profissional, que me mostraram as bases fundamentais da ciência, do ensino acadêmico e da profissão farmacêutica. Em especial, devo agradecer ao Prof. Dr. Jorge Luiz Pesquero que me direcionou para a pesquisa científica acadêmica e me mostrou como é satisfatório desenvolvê-la. Meus agradecimentos

aos professores da Universidade Estadual de Campinas – UNICAMP – e aos membros da Pós-Graduação em Biologia Molecular e Funcional do Departamento de Bioquímica pelos ensinamentos, em especial à secretária Andréia A. Vigilato pelo carinho e auxílio nos trâmites burocráticos.

Por fim, agradeço aos meus familiares que sempre me apoiaram nestes longos e árduos anos de estudo, à eles eu dedico este trabalho. Aos meus pais: José Agostinho Borges Filho e Maria da Conceição Borges, que me deram não só a Vida, mas também a educação, as condições de estudo, a persistência necessária para aqui chegar e que sempre foram e são espelhos para uma vida honesta e longa. Aos meus irmãos: Rogério, Robert, Sérgio e Leila, pelo apoio, amizade, compreensão, conversas e exemplos de vida. Aos meus sobrinhos: João Pedro, Priscila, Camila, Gustavo e André Lucas, pela alegria que fazem a minha vida e a de minha família e que me estimulam a ser para eles um exemplo a ser seguido. Aos meus avós, tios, primos e cunhados pelo carinho, atenção e amizade. Enfim, a Deus, por guiar-me por tão tortuosos caminhos.

Obrigado a todos.

Júlio César Borges

“O tempo todo o tempo passa.”

(Arnaldo Antunes)

ÍNDICES

AGRADECIMENTOS.....	v
ÍNDICES	vii
LISTA DE FIGURAS E TABELAS	xiii
LISTA DE ABREVIACÕES	xvi
RESUMO	xx
ABSTRACT	xxi
1 INTRODUÇÃO.....	1
1.1 Enovelamento e agregação de proteínas	1
1.2 Chaperones moleculares e proteínas de choque térmico.....	2
1.2.1 Hsp70	2
1.2.2 Ciclo funcional das Hsp70	6
1.2.3 Co-chaperones da Hsp70.....	6
Hsp40.....	7
GrpE.....	9
1.3 Técnicas biofísicas utilizadas para estudos de estrutura-função.....	11
1.3.1 Dicroísmo circular	11
1.3.2. Fluorescência	12
1.3.3 Calorimetria	13
1.3.4 Ultracentrifugação analítica	15
1.3.5 Espalhamento de raios-X a baixos ângulos	18
2 OBJETIVOS	27
3 MATERIAIS E MÉTODOS.....	29
3.1 Pesquisas em banco de dados e análise de bioinformática	29
3.2 cDNA plasmidial	29
3.3 Reação em cadeia da DNA polimerase.....	31
3.4 Construção dos oligonucleotídeos iniciadores para as PCR	31
3.5 Gel de agarose e determinação da concentração de DNA.....	31
3.6 Vetores para clonagem.....	32
3.6.1 Vetores de transferência	32
3.6.2 Vetor de expressão	32
3.7 Linhagens de <i>Escherichia coli</i>	33

3.8 Meios de cultura	34
3.9 Clonagem	34
3.9.1 Clonagem em vetores de transferência.....	34
Preparação do vetor pUC18.....	34
Preparação do inserto	35
Reação de ligase.....	35
Análise de clones	35
3.9.2 Subclonagem em vetores de expressão	36
3.10 Transformação de bactérias competentes e seleção dos clones transformantes	36
3.11 Extração de DNA plasmidial	36
3.12 Seqüenciamento de DNA	37
3.13 Expressão de proteínas.....	37
3.14 Purificação das proteínas recombinantes.....	38
3.14.1 Lise de bactérias	38
3.14.2 Cromatografia de afinidade	38
3.14.3 Cromatografia de troca iônica.....	38
3.14.3 Cromatografia de exclusão molecular	39
3.14.4 Processo de concentração de proteínas	39
3.15 SDS-PAGE	40
3.16 Determinação da concentração.....	40
3.16.1 Determinação da concentração de proteínas	40
3.16.2 Determinação da concentração de nucleotídeos.....	41
3.17 Espectropolarimetria de dicroísmo circular.....	41
3.17.1 Espectros de dicroísmo circular.....	42
3.17.2 Experimentos de desnaturação térmica monitorados por CD	42
3.17.3 Experimentos de desnaturação química monitorados por CD	43
3.18 Espectroscopia de emissão de fluorescência.....	43
3.19 Experimentos de calorimetria	44
3.19.1 Calorimetria isotérmica de titulação.....	44
3.19.2 Calorimetria de varredura diferencial.....	45
3.20 Ultracentrifugação analítica	45
3.20.1 Velocidade de sedimentação	45
3.20.2 Sedimentação em equilíbrio	47

3.21 Espalhamento dinâmico de luz.....	47
3.22 Espalhamento de raios X a baixos ângulos.....	48
3.22.1 Coleta dos dados.....	48
3.22.2 Tratamento dos dados de SAXS e modelagem <i>ab initio</i>	49
3.23 Avaliação dos dados hidrodinâmicos dos modelos <i>ab initio</i>	51
3.24 Modelagem por homologia	51
4 RESULTADOS	53
4.1 Hsp70 humana	53
4.1.1 Hsp70-1A	53
Pesquisa em banco de dados e obtenção do cDNA plasmidial da Hsp70-1A	53
Oligonucleotídeos para a PCR da Hsp70-1A	54
Clonagem do cDNA da Hsp70-1A.....	54
Expressão e purificação da proteína Hsp70-1A	55
Análise de estrutura secundária da proteína Hsp70-1A por CD.....	56
Desenovelamento térmico da proteína Hsp70-1A monitorado por CD.....	57
Experimentos de fluorescência do triptofano da proteína Hsp70-1A.....	58
Experimentos de ITC para a proteína Hsp70-1A e seus ligantes MgATP e MgADP.....	59
Análise de velocidade de sedimentação da proteína Hsp70-1A por UAL	60
Estimativa do coeficiente de difusão da proteína Hsp70-1A por EDL	60
Análise de sedimentação em equilíbrio da proteína Hsp70-1A.....	61
Experimentos de SAXS para a proteína Hsp70-1A na presença de MgATP e MgADP	61
Modelos <i>ab initio</i> para a proteína Hsp70-1A	63
Análise hidrodinâmica dos modelos <i>ab initio</i> construídos para a proteína Hsp70-1A	64
Superposição das estruturas em alta resolução das Hsp70 com os modelos <i>ab initio</i>	65
4.1.2 Mortalina.....	83
Pesquisa em banco de dados e obtenção do cDNA plasmidial da mortalina.....	83
Oligonucleotídeos para a PCR da mortalina	83
Clonagem do cDNA da mortalina em pET28a.....	84
Expressão e purificação da proteína mortalina	84
Dicroísmo circular para a proteína mortalina.....	85
4.2 Hsp40 humana	90
4.2.1 Hsp40 DjA1	90
Pesquisa em banco de dados e obtenção do cDNA plasmidial da Hsp40 DjA1	90

Oligonucleotídeos para a PCR da Hsp40 DjA1	90
Clonagem do cDNA da Hsp40 DjA1	91
Outros ensaios realizados para a proteína Hsp40 DjA1	92
4.2.2 Deleção carboxi-terminal da proteína Hsp40 DjA1 – DjA1 ₁₋₃₃₂	94
Oligonucleotídeos para a PCR da Hsp40 DjA1 ₁₋₃₃₂	94
Clonagem do cDNA da proteína Hsp40 DjA1 ₁₋₃₃₂	94
Outros ensaios realizados para a proteína Hsp40 DjA1 ₁₋₃₃₂	95
4.2.3 Hsp40 DjB4	97
Pesquisa em banco de dados e obtenção do cDNA plasmidial da Hsp40 DjB4	97
Oligonucleotídeos para a PCR da Hsp40 DjB4	97
Clonagem do cDNA da proteína Hsp40 DjB4	97
Outros ensaios realizados para a proteína Hsp40 DjB4	98
4.3 GrpE	100
4.3.1 GrpE#1	101
Pesquisa em banco de dados e obtenção do cDNA plasmidial da GrpE#1	101
Oligonucleotídeos para a PCR da GrpE#1	101
Clonagem do cDNA da proteína GrpE#1	101
Outros ensaios realizados para a proteína GrpE#1	102
4.3.2 GrpE#2	104
Pesquisa em banco de dados e obtenção do cDNA plasmidial da GrpE#2	104
Oligonucleotídeos para a PCR da GrpE#2	104
Clonagem do cDNA da proteína GrpE#2	105
Expressão e purificação da proteína GrpE#2	106
Espectros de CD das GrpEs humanas	106
Ensaio de estabilidade térmica e química para as proteínas GrpEs humanas	107
Análise hidrodinâmica da proteína GrpE#2	108
Ensaio de SAXS da proteína GrpE#2	108
Modelagem <i>ab initio</i> para a proteína GrpE#2	109
Análise hidrodinâmica dos modelos <i>ab initio</i> gerados para a proteína GrpE#2	111
Superposição dos modelos construídos por modelagem molecular por homologia da proteína GrpE#1 com o modelo <i>ab initio</i> da proteína GrpE#2	111
Comparação da sequência primária das proteínas GrpEs humanas com a EcGrpE	111
5 DISCUSSÃO	125

5.1 Hsp70-1A.....	125
5.1.1 Clonagem do cDNA, expressão e purificação da proteína Hsp70-1A	125
5.1.2 MgADP e MgATP aumentaram a estruturação da proteína Hsp70-1A	128
5.1.3 MgATP e MgADP aumentaram a estabilidade térmica da proteína Hsp70-1A	128
5.1.4 Experimentos de fluorescência do triptofano da proteína Hsp70-1A.....	130
5.1.5 Calorimetria da interação da proteína Hsp70-1A com os nucleotídeos adenosina ...	131
A afinidade da proteína Hsp70-1A por MgADP ou MgATP	131
ΔH e ΔS da interação entre a proteína Hsp70-1A e MgATP ou MgADP.....	133
5.1.6 MgATP e MgADP provocaram mudanças conformacionais na proteína Hsp70-1A..	134
Análise hidrodinâmica da proteína Hsp70-1A	134
Análise por SAXS da proteína Hsp70-1A na ausência e na presença de MgATP ou MgADP	136
Os modelos <i>ab initio</i> da Hsp70-1A nas diferentes condições experimentais	138
Superposição dos modelos <i>ab initio</i> da Hsp70-1A com estruturas em alta resolução dos domínios das Hsp70.....	138
5.1.7 Considerações finais acerca da proteína Hsp70-1A.....	140
5.2 Mortalina.....	144
5.2.1 A proteína mortalina foi obtida como um agregado protéico	145
5.3 Hsp40	146
5.4 GrpE	147
5.4.1 Clonagem do cDNA, expressão e purificação das proteínas GrpEs humanas.....	147
5.4.2 As proteínas GrpEs humanas possuem diferentes quantidades de estrutura secundária.....	148
5.4.3 As proteínas GrpEs humanas possuem diferentes estabilidades.....	149
5.4.4 As proteínas GrpEs humanas possuem estruturas quaternárias alongadas.....	151
Os ensaios de SAXS e os modelos <i>ab initio</i> para as proteínas GrpEs humanas.....	151
As análises hidrodinâmicas das proteínas GrpEs humanas	152
As proteínas GrpEs humanas são proteínas flexíveis.....	153
5.4.5 A similaridade entre as proteínas GrpEs humanas e a proteína EcGrpE.....	154
5.4.6 A relação entre a forma e a termodinâmica das proteínas GrpEs humanas	154
5.4.7 As proteínas GrpEs humanas compartilham similaridades e diferenças com a EcGrpE.....	155
5.5 Discussão geral	157

6 CONCLUSÕES.....	15
6.1 Conclusões gerais	15
6.2 Conclusões específicas sobre a proteína Hsp70-1A	16
6.3 Conclusões específicas sobre a proteína mortalina	16
6.4 Conclusões específicas sobre as proteínas Hsp40	16
6.5 Conclusões específicas sobre as proteínas GrpEs humanas.....	16
7 REFERÊNCIAS BIBLIOGRÁFICAS	16
7.1 Lista de endereços na rede mundial de computadores – Internet.....	17
8 APÊNDICES	17
8.1 Anexo I	17
8.3 Anexo II	19
8.2 Anexo III	22
8.3 Anexo IV	22
9 CURRICULUM VITAE	23

LISTA DE FIGURAS E TABELAS

Figura 1: Modelo para o ciclo de ação das chaperones moleculares do sistema Hsp70 (modificado de Mayer <i>et al.</i> , 2000).	23
Figura 2: Diagrama da estrutura do dímero da GrpE de <i>E. coli</i>	24
Figura 3: Esquema de um experimento de SAXS, relação partícula-imagem e função de distribuição de distâncias – $p(r)$	25
Tabela I: Lista de clones ESTs contendo os cDNAs das proteínas de interesse adquiridos comercialmente para os procedimentos de clonagem.....	30
Tabela II: Temperaturas médias de transição para a Hsp70-1A na ausência e na presença de aditivos.	58
Tabela III: Parâmetros matemáticos e estruturais obtidos pelos experimentos de SAXS da proteína Hsp70-1A.....	63
Tabela IV: Conjunto de parâmetros hidrodinâmicos determinados para a Hsp70-1A na ausência e na presença de MgADP ou MgATP.	64
Figura 4: Alinhamento das seqüências de aminoácidos da Hsp70-1A, da mortalina e da DnaK.....	66
Figura 5: Verificação da clonagem do cDNA da Hsp70-1A no vetor pET28a.....	67
Figura 6: A Hsp70-1A foi altamente expressa e apresentou alto grau de pureza após o processo de purificação.....	68
Figura 7: Cromatografia de exclusão molecular analítica para a proteína Hsp70-1A.....	69
Figura 8: Análise das frações da Hsp70-1A eluídas pela CEM preparativa por SDS-PAGE....	70
Figura 9: Experimentos de CD das frações da proteína Hsp70-1A purificadas pela CEM preparativa.....	71
Figura 10: Experimentos de CD da proteína Hsp70-1A na ausência e na presença dos nucleotídeos adenosina.....	72
Figura 11: Ensaio de desnaturação térmica de 20 a 90 °C da Hsp70-1A na ausência e na presença de nucleotídeos adenosina.	73
Figura 12: Experimentos de fluorescência da Hsp70-1A na ausência e na presença de nucleotídeos adenosina.....	74
Figura 13: Determinação da K_A da Hsp70-1A por seus ligantes: MgATP ou MgADP.	75
Figura 14: Experimentos de velocidade de sedimentação da proteína Hsp70-1A na ausência e na presença de MgADP e MgATP.....	76

Figura 15: Experimentos de sedimentação em equilíbrio da proteína Hsp70-1A.	77
Figura 16: Curvas das $I(q)$ medidas por SAXS para a proteína Hsp70-1A na ausência e na presença de nucleotídeos adenosina.	78
Figura 17: Análise das curvas das $p(r)$ da Hsp70-1A na ausência e na presença de nucleotídeos adenosina.	79
Figura 18: Gráfico de Kratky obtidos das curvas das $I(q)$ da Hsp70-1A medidas por SAXS.	80
Figura 19: Modelos <i>ab initio</i> gerados para a proteína Hsp70-1A a partir dos dados de SAXS nas diferentes condições experimentais.	81
Figura 20: Superposição das estruturas em alta resolução determinadas para os domínios das Hsp70 com os modelos <i>ab initio</i> gerados para a proteína Hsp70-1A.	82
Figura 21: Clonagem do cDNA da mortalina no vetor pET28a.	86
Figura 22: Cromatograma da CEM analítica para a proteína mortalina.	87
Figura 23: Etapas de expressão e purificação da mortalina analisada por SDS-PAGE.	88
Figura 24: Caracterização da estrutura secundária da mortalina por CD.	89
Figura 25: Verificação da clonagem do cDNA da Hsp40 DjA1 no vetor pET28a.	93
Figura 26: Verificação da clonagem do cDNA da deleção carboxi-terminal da Hsp40 DjA1 no vetor pET28a.	96
Figura 27: Verificação da clonagem do cDNA da proteína Hsp40 DjB4 no vetor pET28a.	99
Figura 28: Clonagem do cDNA da proteína GrpE#1 no vetor pET23a.	103
Tabela V: Parâmetros biofísicos das GrpEs humanas e da proteína EcGrpE analisados por CD.	107
Tabela VI: Conjunto de parâmetros matemáticos, estruturais e hidrodinâmicos da proteína GrpE#2 obtidos nos experimentos de SAXS, UAL, EDL e para o modelo <i>ab initio</i>	110
Tabela VII: Homologia na sequência de aminoácidos entre as proteínas GrpE#1, GrpE#2 e EcGrpE.	112
Figura 29: Clonagem do cDNA da proteína GrpE#2 humana no vetor pET28a.	113
Figura 30: Análise do processo de indução e purificação da GrpE#2 analisada por SDS-PAGE 12%.	114
Figura 31: Análise de CD para as proteínas GrpEs humanas.	115
Figura 32: Ensaios de estabilidade térmica para as proteínas GrpE humanas.	116
Figura 33: Desenovelamento das proteínas GrpE humanas por uréia.	117
Figura 34: Experimentos de UAL: velocidade de sedimentação para a proteína GrpE#2.	118
Figura 35: Dados de SAXS obtidos para a proteína GrpE#2.	119

Figura 36: Gráfico de Kratky das proteínas GrpE humanas.	120
Figura 37: Modelos <i>ab initio</i> gerados para a proteína GrpE#2.	121
Figura 38: Comparação do modelo <i>ab initio</i> da GrpE#2 com os modelos da estrutura quaternária da GrpE#1.	122
Figura 39: Alinhamento das seqüências de aminoácidos das proteínas GrpEs humanas e EcGrpE.	123
Figura 40: Posicionamento espacial relativo do triptofano localizado no DLN das proteínas Hsp70-1A e DnaK.....	143

LISTA DE ABREVIações

- ΔC_p : variação capacidade calorífica;
 ΔG : variação da energia livre de Gibbs;
 ΔH : variação da entalpia;
 ΔS : variação da entropia;
 Σ : somatório;
 ε : coeficiente de absorvidade molar;
 η : viscosidade;
 η_{tex} : viscosidade do solvente na temperatura experimental;
 η_{t20} : viscosidade do solvente à 20 °C;
 η_w : viscosidade da água;
 λ : comprimento de onda;
 Θ : elipticidade;
 ρ : densidade;
 ρ_e : densidade eletrônica;
 ρ_f : densidade do fluido;
 $\rho_{20,w}$: densidade do solvente em água e na temperatura de 20 °C;
 ρ_{ex} : densidade do solvente nas condições experimentais;
 σ : desvio padrão;
 ω : velocidade angular;
 $\text{\AA}$: angström (1×10^{-10} metros);
AMP: adenosina mono-fosfato;
ADP: adenosina 5' di-fosfato;
ATP: adenosina 5' tri-fosfato;
BLAST: ferramenta de procura por alinhamento básico local (*Basic Local Alignment Search Tool*);
BSA: albumina de soro bovina (*Bovine Serum Albumin*);
°C: unidade de temperatura – graus Celsius;
C: concentração;
CEM: cromatografia de exclusão molecular;
CD: dicroísmo circular;
cDNA: DNA complementar;

C_m : concentração média de transição;
 C_p : capacidade calorífica;
 D : coeficiente de difusão;
 $D_{20,w}$: coeficiente de difusão em água e na temperatura de 20 °C;
 $D^0_{20,w}$: coeficiente de difusão em água, na temperatura de 20 °C e concentração zero de proteínas;
DAM: modelo de pequenas esferas (*Dummy Atoms Model*);
Deg: grau (*degree*);
DLN: domínio de ligação de nucleotídeos adenina das Hsp70;
DLS: domínio de ligação de substratos das Hsp70;
 D_{max} : distância máxima;
DnaJ_C: domínio carboxi-terminal das Hsp40;
DSC: calorimetria de varredura diferencial (*Differential Scanning Calorimetry*);
DTT: ditioneitol;
EcGrpE: GrpE de *E. coli*;
EDL: espalhamento dinâmico de luz;
EDTA: ácido etilenodiaminotetracético;
EMR: elipticidade molar residual;
EST: etiqueta de sequência expressa (*Expressed Sequence Tag*);
Exp.: Dados experimentais;
 f : coeficiente friccional;
 f/f_0 : razão friccional ou fator de Perrin;
 F_C : Força centrípeta;
 F_{Emp} : Força de empuxo;
 F_f : Força friccional;
Pi: fosfato inorgânico;
FPLC: cromatografia líquida de proteínas por pressão (*Flow Protein Liquide Chromatography*);
 g : aceleração da gravidade;
Gdn-HCl: cloridrato de guanidina;
G/F: região rica em glicinas e fenilalaninas das Hsp40;
GrpE: proteína que auxilia a troca de nucleotídeos das Hsp70 (*GroP-like gene E*);
Hip: proteína de interação com Hsp (*Hsp-interacting protein*);

Hop: proteína organizadora das Hsp70-Hsp90 (*Hsp70-Hsp90 organizing protein*);
 Hsc70: proteína de choque térmico cognata de 70 kDa;
 Hsp: proteína de choque térmico (*Heat shock protein*);
 Hsp170: proteína de choque térmico de 170 kDa;
 Hsp110: proteína de choque térmico de 110 kDa;
 Hsp90: proteína de choque térmico de 90 kDa;
 Hsp70: proteína de choque térmico de 70 kDa;
 Hsp40: proteína de choque térmico de 40 kDa;
 $I_{F\lambda}$: intensidade de fluorescência em um determinado comprimento de onda;
IMAGE Consortium: Integrated Molecular Analysis of Genomes and their Expression Consortium;
 IPTG: Isopropiltiol- β -D-galactosídeo;
 $I(q)$: função da intensidade de espalhamento;
 ITC: calorimetria isotérmica de titulação (*Isothermal Titration Calorimetry*);
 k : constante de Boltzmann ($1,381 \times 10^{-23} \text{ J K}^{-1}$ ou $3,298 \times 10^{-24} \text{ cal K}^{-1}$);
 K: unidade de temperatura – Kelvin;
 kDa: kiloDalton;
 K_A : constante de associação;
 K_D : constante de dissociação;
 LNLS: Laboratório Nacional de Luz Síncrotron;
 l : comprimento do caminho ótico;
 LB: Luria-Bertani;
 m : massa;
 MM : massa molecular;
 n : estequiometria da ligação;
 N_{Av} : número de Avogrado ($6,02214199 \times 10^{23} \text{ mol}^{-1}$);
 NCBI: *National Center of Biotechnology Information*;
 ND: não determinado;
 nm: nanômetro (10^{-9} metros);
 NCS: Número de canais de Shannon;
 pb: par de bases;
 P : Fator de Perrin ou razão friccional;
 PCR: reação em cadeia da DNA polimerase (*DNA Polymerase Chain Reaction*);

PDB: banco de dados de estrutura em alta resolução de proteínas (*Protein Data Bank*);

$p(r)$: função de distribuição de distâncias;

q : ângulo de espalhamento;

Q : quantidade de calor;

r : raio;

r_{esf} : raio da esfera;

R : constante dos gases ($8,314 \text{ J mol}^{-1} \text{ K}^{-1}$ ou $1,987 \text{ cal mol}^{-1} \text{ K}^{-1}$);

R_g : raio de giro;

R_S : raio de Stokes;

s : coeficiente de sedimentação;

s^* : coeficiente de sedimentação aparente;

$s_{20,w}$: coeficiente de sedimentação em água e na temperatura de 20°C ;

$s_{20,w}^0$: coeficiente de sedimentação em água, na temperatura de 20°C e concentração zero de proteínas;

S : Svedberg (10^{-13} segundos);

SAS: *Small angle scattering beam line*;

SAXS: espalhamento de raios-X à baixos ângulos (*Small Angle X-ray Scattering*);

SDS-PAGE: eletroforese em gel de poliacrilamida em presença de duodecil-sulfato de sódio (*sodium duodecyl-sulphate polyacrilamide gel electrophoreses*);

SMC: sítio múltiplo de clonagem;

T : temperatura absoluta;

TKP: Tampão Tris-HCl 25 mmol/L (pH 7,5), contendo NaCl 50 mmol/L , KCl 5 mmol/L , Na_3PO_4 5 mmol/L e 1 mmol/L β -mercaptoetanol;

T_m : temperatura média de transição;

UA: unidade arbitrária;

UAL: ultracentrifugação analítica;

v : velocidade;

V_{bar} : volume parcial específico;

V_{cel} : volume da célula do microcalorímetro;

$[\Theta]$: elipticidade molar residual;

$[L]$: concentração do ligante livre;

$[M]$: concentração do ligante;

$\langle\lambda\rangle$: centro de massa espectral de fluorescência.

RESUMO

A função de uma proteína depende da estrutura nativa obtida a partir de sua estrutura primária de aminoácidos por um processo denominado de enovelamento protéico. Falhas neste processo podem levar à formação de agregados protéicos que têm relação com doenças humanas; como as doenças amilóides. Durante a evolução, as células têm desenvolvido mecanismos de prevenção e controle de qualidade para evitar a agregação de proteínas. Um destes mecanismos é através da síntese das chaperones moleculares. Dentre as várias proteínas desta categoria, destaca-se o sistema Hsp70 que desempenha funções essenciais no metabolismo de proteínas, pois age como um pivô, recebendo e distribuindo proteínas desenoveladas ou substratos entre as demais chaperones moleculares. Assim, o estudo do mecanismo de ação desta maquinaria é importante para compreender o enovelamento protéico no meio intracelular. Como muitas das informações disponíveis para o sistema Hsp70 são do sistema bacteriano, este estudo pretendeu a obtenção de proteínas da família Hsp70: Hsp70-1A e suas co-chaperones Hsp40 (DjA1 e DjB4) e GrpE (GrpE#1 e GrpE#2) de origem humana. Os cDNAs destas proteínas foram clonados e as proteínas foram expressas, purificadas e caracterizadas quanto à relação estrutura e função. Para tanto, foram utilizadas técnicas como: dicroísmo circular, fluorescência, calorimetria, ultracentrifugação analítica e espalhamento de raios X a baixos ângulos. Os resultados principais mostraram que: 1) a Hsp70-1A sofre pequenas mudanças conformacionais induzidas pelos nucleotídeos adenosina tri-fosfato ou adenosina di-fosfato; 2) as duas representantes humanas das subfamílias A e B das Hsp40, a DjA1 e DjB4, respectivamente, possuem atividade chaperone e estrutura quaternária distintas uma da outra; e 3) as GrpEs humanas possuem estrutura quaternária similar, porém diferem na disposição espacial de alguns domínios e em outras propriedades biofísicas, indicando funções específicas ou especializadas no interior da mitocôndria. Estes resultados podem ajudar na compreensão sobre a relação estrutura-função desta importante maquinaria celular.

ABSTRACT

The role of a protein depends on its three-dimensional structure, which is acquired from its amino acid sequence through a process called protein folding. However, mistakes on this process cause protein aggregation, which may be related to some human illness; like amyloids diseases. During the evolution, the cells have developed mechanisms to prevent and to control quality in order to avoid protein aggregation. One of these mechanisms is the molecular chaperones. Among the several proteins from this category, the Hsp70 system has essential functions in the protein metabolism, because it acts as a pivot, receiving and handing out unfolded proteins or substrates among the others molecular chaperones. Thus, the study of the action mechanism of the Hsp70 machinery is important to understand the protein folding process inside the cells. Since, there is a large amount of information available about the Hsp70 system from bacteria, this study intended the characterization of proteins from human Hsp70 system: Hsp70-1A and its co-chaperones Hsp40 (DjA1 and DjB4) and GrpE (GrpE#1 and GrpE#2). The cDNA from those proteins were cloned and the proteins were expressed, purified, and characterized concerning their structure-function relation. In such effort, several techniques were used, like: circular dichroism, fluorescence, calorimetry, analytical ultracentrifugation and small angle X-ray scattering. The main results are: 1) the Hsp70-1A underwent little conformational changes induced by the nucleotides adenosine triphosphate or adenosine di-phosphate; 2) the two representatives of human Hsp40 proteins DjA1 and DjB4, from subfamily A and B, respectively, possess distinct chaperone activities and quaternary structures; and 3) the human GrpEs possess similar quaternary structure, but with discrete differences in the spatial disposition of some domains and with dissimilar biophysical properties, implying in specific or specialized functions inside the mitochondria. These results may help the further understanding of the structure-function relation of that important cell machinery.

1 INTRODUÇÃO

1.1 Enovelamento e agregação de proteínas

A atividade biológica das proteínas está relacionada com a estrutura nativa que a mesma é capaz de obter no meio solúvel. Anfinsen (1973) mostrou que todas as informações necessárias para o enovelamento correto estão contidas na própria sequência primária de aminoácidos que compõem a proteína. O processo pelo qual uma proteína recém sintetizada obtém sua estrutura terciária é denominado de enovelamento protéico. O primeiro estágio deste processo parece envolver o colapso hidrofóbico que pode compensar, termodinamicamente, a entropia desfavorável proveniente da formação de uma estrutura organizada (revisões sobre enovelamento protéico: Privalov (1996); Jaenicke (1999) e suas citações). Para proteínas pequenas, o enovelamento parece ocorrer rapidamente, porém em proteínas maiores este processo pode ser mais lento, passando por etapas que envolvam intermediários parcialmente enovelados. O enovelamento através de intermediários pode ser descrito em três etapas: 1) formação muito rápida de estruturas secundárias; 2) formação rápida do intermediário caracterizado pelo colapso da cadeia peptídica contendo estrutura secundária em um glóbulo compacto sem estrutura terciária rígida; e 3) formação da estrutura protéica nativa (Ptitsyn, 1991). O enovelamento protéico é influenciado por várias propriedades físico-químicas como: energia livre, temperatura, força iônica, acidez, concentração de proteínas, pontes dissulfeto, glicosilação, etc (Ruddon & Bedows, 1997). Algumas proteínas podem interagir umas com as outras em condições não adequadas, como em soluções em alta concentração ou em condições de desnaturação (como o choque térmico), através de regiões hidrofóbicas expostas e formar agregados protéicos que tendem a se precipitar *in vivo* e *in vitro* (Jaenicke *et al.*, 1987; Beissinger & Buchner, 1998).

Agregados protéicos ou proteínas enoveladas incorretamente não são somente problemas para os estudos de proteínas *in vitro* ou um fator de perdas econômicas para as indústrias biotecnológicas e farmacêuticas (Ruddon & Bedows, 1997; Fink, 1998). As falhas no enovelamento protéico, resultando na agregação, também podem estar relacionadas com algumas doenças humanas, tais como: mal de Alzheimer, mal de Parkinson, doença de Creutzfeld-Jacob (prion), fibrose cística, doença de Huntington, deficiência de α_1 -antitripsina, alguns tipos de câncer, entre outras (Ruddon & Bedows, 1997; Beissinger & Buchner, 1998; Chai *et al.*, 1999; Hartl & Hayer-Hartl, 2002). Agregados protéicos podem se formar dentro

das células ou em tecidos causando danos aos órgãos ao ativar a resposta inflamatória, por exemplo. A principal questão é como prevenir tais patologias e quando isto não for possível como reverter a agregação protéica.

As células controlam o problema da agregação protéica e enovelamento incorreto de proteínas através da expressão de uma classe especial de proteínas, em condições normais e em estresse, que são capazes de auxiliar o enovelamento protéico, prevenir a agregação de proteínas e também recuperar proteínas de agregados, entre outras funções. Tais proteínas são chamadas genericamente de chaperones moleculares ou proteínas do choque térmico (Hsp). As chaperones moleculares são temas de várias revisões recentes, como as citadas a seguir: Gething & Sambrook (1992); Hendrick & Hartl (1993), Boston *et al.* (1996); Martin & Hartl (1997); Bukau & Horwich (1998); Beissinger & Buchner (1998); Fink (1999); Miernyk (1999); Hartl & Hayer-Hartl (2002).

1.2 Chaperones moleculares e proteínas de choque térmico

Uma revisão sobre as principais famílias de chaperones moleculares pode ser encontrada na revisão em anexo (Borges, J.C. & Ramos, C.H.I. 2004, Protein folding assisted by chaperones. *Protein Pept. Lett.*, submetido – Anexo I), que relata a importância e mecanismo de ação das principais famílias de chaperones moleculares. Nesta revisão, é enfatizada a importância da família Hsp70 no sistema chaperone celular, no qual elas funcionam como um pivô mediando a troca de substratos e a ação chaperone com as demais famílias de chaperones moleculares. A seguir é apresentada uma descrição mais detalhada das propriedades gerais e características das proteínas que compõem o sistema chaperone Hsp70 e suas principais co-chaperones, as Hsp40 e GrpE, que são objeto deste trabalho.

1.2.1 Hsp70

Uma das mais importantes famílias de chaperones moleculares é a formada pelo sistema Hsp70: Hsp70 (DnaK em *Escherichia coli*) e suas co-chaperones Hsp40 (DnaJ em *E. coli*) e GrpE. Tal sistema está envolvido em uma série de importantes funções no metabolismo protéico intracelular envolvendo o enovelamento de proteínas, nos quais é possível destacar: ligação de polipeptídeos recém sintetizados auxiliando seu enovelamento; prevenção da agregação induzida por diferentes tipos de estresses e recuperação de proteínas de agregados (Mayer *et al.*, 2001). As Hsp70 estão envolvidas em outras funções,

nas quais não envolvem necessariamente o enovelamento de proteínas, como o transporte de proteínas por membranas e para a degradação protéica, na estabilização de fatores de transcrição celular, entre outras (Mayer *et al.*, 2001; Hartl & Hayer-Hartl, 2002). No meio intracelular, as Hsp70 exercem funções de controle de qualidade sobre proteínas enoveladas incorretamente (Glover & Lindquist, 1998; Goloubinoff *et al.*, 1999; Mogk *et al.*, 1999) e em proteínas recém sintetizadas (Hendrick *et al.*, 1993; Frydman *et al.*, 1994; Mayer *et al.*, 2001). Na realidade, o sistema Hsp70 funciona como um “pivô” capaz de receber e distribuir proteínas instáveis e/ou parcialmente enoveladas entre as demais famílias de chaperones moleculares (Anexo I – Borges & Ramos, 2004, submetido).

O mecanismo pelo qual as Hsp70 exercem suas funções ainda não é muito compreendido, porém duas hipóteses são consideradas: 1) as Hsp70 agiriam de maneira passiva, na qual através de múltiplos ciclos de associação e dissociação com o substrato, este se enovelaria espontaneamente (Mayer *et al.*, 2001) e 2) a energia liberada pela hidrólise de adenosina 5' tri-fosfato (ATP) no ciclo das Hsp70 (ver abaixo) provocaria mudanças conformacionais no substrato rompendo algumas barreiras energéticas na via de enovelamento (Pierpaoli *et al.*, 1997; Mayer *et al.*, 2001).

As Hsp70 apresentam massa molecular (*MM*) de aproximadamente 70 kDa e são divididas em dois domínios: o domínio de ligação de nucleotídeos adenosina (DLN – aproximadamente 45 kDa), que possui atividade de hidrólise de ATP, e o domínio de ligação de substratos (DLS – aproximadamente 25 kDa), capaz de ligar seqüências de resíduos de aminoácidos hidrofóbicos. No entanto, existem proteínas classificadas como pertencentes à superfamília das Hsp70 e apresentam seqüência primária correspondente a proteínas de aproximadamente 110 kDa e 170 kDa. Tais proteínas são chamadas de Hsp110 e Hsp170, respectivamente, e apresentam alta similaridade na seqüência de aminoácidos com o DLN das Hsp70. Entretanto, na região correspondente ao DLS, as proteínas Hsp110 e Hsp170 possuem seqüências de aminoácidos não similares inseridos em regiões de alta similaridade com as Hsp70. As funções das Hsp110 e Hsp170 e, mais especificamente, daquelas regiões inseridas, ainda não estão bem descritas (Easton *et al.*, 2000; Mayer *et al.*, 2001).

As Hsp70 são capazes de interagir com seqüências de, no mínimo, 5 resíduos de aminoácidos hidrofóbicos flanqueados por resíduos positivos (Zhu *et al.*, 1996; Rüdiger *et al.*, 1997; Rüdiger *et al.*, 2001) e a afinidade pelos substratos é regulada pela presença do nucleotídeo adenosina no DLN. O DLS apresenta alta afinidade por substratos quando o nucleotídeo adenosina 5' di-fosfato (ADP) está ligado ao DLN, apresentando uma baixa taxa

de ligação de substratos. A baixa afinidade do DLS pelo substrato ocorre quando o ATP está ligado ao DLN, sendo a taxa de ligação de substratos alta nesta condição (Theyssen *et al.*, 1996; Mayer *et al.*, 2000; Mayer *et al.*, 2001). O DLN funciona como uma “concha” (Figura 1), existindo no seu interior o sítio de ligação para os nucleotídeos, que pode ser dividido em dois sub-domínios (DLN1 e DLN2). O DLS também pode ser subdividido em dois sub-domínios: um formado principalmente de folha β -pregueada (DLS1), que funciona como uma “cama” para o substrato e outro composto por uma longa estrutura em α -hélice que pode funcionar como uma “tampa” (DLS2) (Zhu *et al.*, 1996; Mayer *et al.*, 2000). A “tampa” possuiria uma mobilidade que a torna capaz de fechar sobre a “cama” através de uma dobradiça (Zhu *et al.*, 1996) e, por conseguinte, sobre o substrato. Assim, duas possíveis conformações principais da Hsp70, na presença de diferentes nucleotídeos adenosina, poderiam ser descritas: uma “aberta” e outra “fechada” em relação ao substrato. Na ausência de nucleotídeos adenosina, foi descrito a existência de um estado de equilíbrio entre as duas possíveis conformações da tampa (Mayer *et al.*, 2000). A ligação de ADP ao DLN provocaria mudanças conformacionais neste que seriam transmitidas ao DLS, deslocando o equilíbrio para a conformação fechada. A ligação de ATP causaria um efeito oposto, no qual o equilíbrio seria deslocado para a conformação aberta (Mayer *et al.*, 2000). O ciclo de ligação de ATP na Hsp70, a hidrólise de ATP e a troca de ADP por ATP são auxiliados pelas co-chaperones da Hsp70, as Hsp40 e GrpE (Figura 1), e será discutido a seguir.

Embora as proteínas Hsp70 apresentem alta conservação funcional e estrutural, elas podem ser classificadas em 3 subfamílias denominadas como DnaK, Hsc70 e HscA, segundo as suas propriedades de dissociação de nucleotídeos adenosina. As Hsc70 e HscA apresentam afinidade por nucleotídeos adenosina de aproximadamente 20 e 700 vezes, respectivamente, menor do que a afinidade da DnaK por tais ligantes (Brehmer *et al.*, 2001). As subfamílias DnaK, Hsc70 e HscA possuem duas principais diferenças estruturais em duas regiões da estrutura terciária do DLN: 1) uma alça exposta na estrutura localizada na abertura da fenda do sítio de ligação do nucleotídeo adenosina; e 2) a polaridade da interface da fenda de acesso ao sítio de ligação do nucleotídeo (Brehmer *et al.*, 2001). A subfamília DnaK (Hsp70 de procariotos e aquelas encontradas em organelas de eucariotos) possui uma longa alça (A276-R302 na DnaK de *E. coli*), na qual as GrpE seriam capazes de se ligar, e uma região hidrofóbica no topo da fenda e duas pontes salinas responsáveis pela polaridade da interface da fenda. Estes fatores dificultariam o trânsito de ligantes pela fenda (Brehmer *et al.*, 2001). As representantes da subfamília Hsc70 (as Hsp70 de citoplasma de eucariotos)

apresentam a alça menor em 4 resíduos de aminoácidos, não apresentando a região hidrofóbica e uma das pontes salinas na interface da fenda. A subfamília representada pelas HscA (também encontradas em procariotos) apresenta 10 resíduos de aminoácidos a menos naquela alça, e nenhum dos três fatores da interface da fenda (Brehmer *et al.*, 2001). Tais propriedades podem explicar as diferenças nas afinidades das Hsp70 pelos nucleotídeos adenosinas e a necessidade de co-chaperones para auxiliar a troca de nucleotídeos adenosina nas subfamílias DnaK e Hsc70 (Brehmer *et al.*, 2001).

Como descrito acima, a ligação de ATP ou ADP ao DLN é um fator importante para a ação das Hsp70. Portanto, a afinidade das Hsp70 pelos nucleotídeos adenosina também é um fator importante para a ação destas proteínas. As proteínas DnaK de *E. coli* e de *Thermus thermophilus* e também a proteína Hsc70 de cérebro bovino, apresentam constantes de dissociação de ordem submicromolar pelos nucleotídeos MgATP e MgADP (Mayer *et al.*, 2001). Este fator *per se* indica que a associação da Hsp70 com ATP não é o fator limitante do ciclo de ação (ver abaixo), pois em condições fisiológicas, as Hsp70 estão saturadas com ATP (Mayer *et al.*, 2001). A etapa de transição das Hsp70 da forma de baixa afinidade pelo substrato (Hsp70-ATP) para a forma de alta afinidade (Hsp70-ADP) é dependente da hidrólise de ATP, sendo este o fator limitante do ciclo devido à baixa capacidade das Hsp70 em realizar esta hidrólise (Wilbanks *et al.*, 1994; Mayer *et al.*, 2001). A hidrólise de ATP é estimulada pela ligação do substrato ao DLS e também pela interação com as Hsp40 (Mayer *et al.*, 2001 e citações). Entretanto, alguns autores relatam que algumas Hsp70 apresentam constantes de dissociação pelos nucleotídeos adenosina (ATP e ADP) na ordem de micromolar (Mayer *et al.*, 2001). Ha & McKay (1994) discutem que muitos dados da literatura podem conter erros na determinação das constantes de dissociação devido à ausência de etapas de remoção de nucleotídeos adenosinas porventura ligados à proteína. Entretanto, as ações de determinados sais devem ser levadas em consideração (O'Brien & McKay, 1995). Alguns trabalhos mostram a importância de alguns sais para a ligação dos nucleotídeos e para a ação das Hsp70 (Mayer *et al.*, 2001). Por exemplo, íons Ca^{2+} , Mg^{2+} e K^+ foram encontrados ligados a sítios específicos nas estruturas em alta resolução do DLN de algumas Hsp70 (Flaherty *et al.*, 1990; Wilbanks & McKay, 1995). Ademais, a hidrólise de ATP é dependente da concentração de íons K^+ (Ha & McKay, 1994), mas não de íons Na^+ (O'Brien & McKay, 1995). O Pi pode ser importante em reduzir a taxa de liberação de ADP pelo DLN, sendo uma das etapas limitantes da hidrólise de ATP, o que provoca um relativo aumento na constante de afinidade (Ha & McKay, 1994).

1.2.2 Ciclo funcional das Hsp70

O ciclo funcional simplificado das Hsp70 está representado na figura 1 (modificado de Mayer *et al.*, 2000). Como descrito acima, a afinidade das Hsp70 por substratos depende do nucleotídeo adenosina ligado ao DLN – ADP (alta afinidade) e ATP (baixa afinidade) – sendo que a hidrólise de ATP é o fator limitante do ciclo (Mayer *et al.*, 2001). Na presença de ATP, a ligação de substratos ao DLS é caracterizada por uma alta taxa de associação e dissociação, visto que o equilíbrio conformacional do subdomínio DLS2 estaria deslocado para a conformação aberta em relação ao DLS1 e substrato, expondo o sítio de ligação do substrato do DLS. O substrato pode se ligar ao DLS a partir da solução ou ser transportado pelas Hsp40 para as Hsp70. Assim, as Hsp40 poderiam agir como um fator de seleção de substratos para as Hsp70 (Rüdiger *et al.*, 2001). Tanto a presença do substrato ligado ao DLS quanto a interação entre as Hsp40 a Hsp70 são fatores de estimulação, de maneira concomitante, da atividade de hidrólise de ATP (Laufen *et al.*, 1999). A presença de ADP ligado ao DLN resultaria em mudanças conformacionais que deslocariam o equilíbrio do DLS2 para a conformação fechada. Nesta conformação, a Hsp70 apresenta alta afinidade pelo substrato (Montgomery *et al.*, 1999; Mayer *et al.*, 2001). Dependendo da Hsp70 em questão, pode haver a necessidade da ação das co-chaperones, GrpE ou Bag-1, para induzir o desligamento do ADP e a posterior liberação do substrato para o meio ou para as outras chaperones moleculares. Assim, as GrpEs podem reciclar as Hsp70 disponibilizando-as para um outro ciclo de ação (Liberek *et al.*, 1991).

1.2.3 Co-chaperones da Hsp70

As Hsp70 possuem muitas co-chaperones que auxiliam e regulam a sua atividade (Mayer *et al.*, 2001). Muitas proteínas agem no heterocomplexo formado pelas Hsp70-Hsp90, como a Hip e a Hop que interagem com as Hsp70 pela seqüência repetitiva de motivos tetratricopeptídeos (Mayer *et al.*, 2001). Neste trabalho, entretanto, é descrito em detalhes as duas principais co-chaperones da Hsp70: as Hsp40, que estimulam a atividade de hidrólise de ATP das Hsp70; e as GrpE, que são fatores de troca de nucleotídeos no DLN das Hsp70. Detalhes sobre as demais co-chaperones da Hsp70 podem ser encontrados na revisão escrita por Mayer *et al.* (2001).

Hsp40

As Hsp40 são co-chaperones que estimulam a atividade de hidrólise de ATP das Hsp70, selecionam e transportam substratos para as Hsp70, possuindo também atividade chaperone intrínseca (Hendrick *et al.*, 1993; Liberek *et al.*, 1996; Karzai & McMacken, 1996, Laufen *et al.*, 1999). Assim, as Hsp40 auxiliam as Hsp70 no enovelamento de proteínas recém sintetizadas, na prevenção da agregação, no reenovelamento de proteínas de agregados, entre outros processos (Gething & Sambrook, 1992; Hendrick *et al.*, 1993; Liberek *et al.*, 1996; Martin & Hartl, 1997; Fink, 1999; Mayer *et al.*, 2001). A ação das Hsp40 no estímulo à hidrólise de ATP pelas Hsp70 é sub-estequiométrica, na qual elas funcionam de uma maneira sinérgica com o substrato (Pierpaoli *et al.*, 1998; Laufen *et al.*, 1999).

As Hsp40 são proteínas de aproximadamente 40 kDa, encontradas principalmente na forma de dímeros e caracterizadas pela presença de um domínio rico em α -hélice chamado de domínio J, devido a sua identificação na DnaJ, a Hsp40 de *E. coli*. A interação entre o domínio J das Hsp40 e o DLN das Hsp70 ocorre principalmente através de uma sequência de 3 resíduos de aminoácidos – HPD – que são evolutivamente conservados e estão localizados em uma alça exposta ao solvente (Szabo *et al.*, 1994; Greene *et al.*, 1998; Suh *et al.*, 1999). Entretanto, alguns trabalhos mostram a possibilidade de múltiplos sítios de interação entre as Hsp40 e Hsp70 (Qian *et al.*, 2002; Linke *et al.*, 2003) e ensaios de espectroscopia de ressonância de plasma superficial mostram a existência de pelo menos dois sítios de interação entre a DnaK e a DnaJ de *E. coli*, ambos localizados no domínio J da DnaJ (Suh *et al.*, 1999).

As Hsp40 são proteínas ubíquas e apresentam alta diversidade, porém diferem na arquitetura de domínios e na localização celular (Cyr *et al.*, 1994; Cheetham *et al.*, 1998; Mayer *et al.*, 2001). Um exemplo desta diversidade é a presença no genoma humano de 44 possíveis genes que codificam sequências de Hsp40 (Venter *et al.*, 2001). As Hsp40 podem ser divididas em 4 regiões conservadas (Figura 1a do Anexo II): o domínio J, que é capaz de interagir com as Hsp70 estimulando a atividade de hidrólise de ATP, é geralmente localizado na região amino-terminal da proteína e é composto principalmente por estrutura secundária do tipo α -hélice (Szabo *et al.*, 1994; Greene *et al.*, 1998; Suh *et al.*, 1999). Após o domínio J, pode ser identificada uma região de 30 a 70 resíduos de aminoácidos onde muitos são resíduos de glicina e fenilalanina (G/F), que pode funcionar como um espaçador entre o domínio J e o restante da estrutura da Hsp40 e ser responsável por uma flexibilidade entre

tais regiões (Szyperski *et al.*, 1994; Karzai & McMacken, 1996). A região G/F é essencial para a função das Hsp40 testadas (Karzai & McMacken, 1996; Yan & Craig, 1999), pois o domínio J isolado é incapaz de interagir com as Hsp70 (Szabo *et al.*, 1996; Karzai & McMacken, 1996; Suh *et al.*, 1999). A região central de algumas Hsp40 apresenta um domínio rico em resíduos de cisteínas formado por 4 repetições do motivo CXXCXGXC (onde X é um aminoácido qualquer). Esta região central se enovela de maneira dependente de zinco, onde dois conjuntos de motivos ligam 1 íon de zinco (Szabo *et al.*, 1996; Banecki *et al.*, 1996; Martinez-Yamout *et al.*, 2000). O domínio localizado na região carboxi-terminal, chamado de domínio DnaJ_C, é formado principalmente de estrutura em folha β -pregueada e se enovela como uma ferradura na estrutura dimérica (Sha *et al.*, 2000). A função da região rica em cisteínas e do domínio DnaJ_C das Hsp40 é a interação com substratos, seqüestrando-os do meio, prevenindo a agregação e transportando-os para as Hsp70 (Szabo *et al.*, 1996; Banecki *et al.*, 1996; Sha *et al.*, 2000; Rüdiger *et al.*, 2001; Lee *et al.*, 2002). A recente determinação da estrutura em alta resolução da região central rica em cisteínas juntamente com o domínio DnaJ_C com um substrato peptídico ligado, mostra a participação de ambos domínios na ligação de substratos (Li *et al.*, 2003).

Como discutido acima, as Hsp40 também podem funcionar como chaperones moleculares (Langer *et al.*, 1992; Szabo *et al.*, 1994; Rüdiger *et al.*, 2001). O sítio de ligação de substratos da DnaJ reconhece seqüências de 8 resíduos de aminoácidos que apresentam similaridade com aqueles reconhecidos pela DnaK (Rüdiger *et al.*, 2001). Deste modo, a DnaJ pode agir como um fator selecionador de substratos para a DnaK (Rüdiger *et al.*, 2001). A ação chaperone das Hsp40 ocorre por interação hidrofóbica com os substratos (Rüdiger *et al.*, 2001; Lee *et al.*, 2002; Li *et al.*, 2003) e dois possíveis mecanismos de ação são propostos para o transporte de substratos das Hsp40 para as Hsp70. Em uma das possibilidades, as Hsp40 se ligariam ao mesmo sítio hidrofóbico que as Hsp70, não ocorrendo formação de complexos Hsp40-Hsp70 via substrato. Na outra possibilidade, as Hsp40 e Hsp70 interagiriam através de diferentes sítios hidrofóbicos nas proteínas não enoveladas, permanecendo ligadas pelo substrato por algum tempo formando um complexo transiente Hsp40-Hsp70-substrato (Rüdiger *et al.*, 2001). Laufen *et al.*, (1999) propõem que o ciclo funcional das Hsp40 com as Hsp70 ocorre em múltiplas etapas: 1) interação entre o domínio J com o DLN das Hsp70 no estado ligado ao ATP; 2) transferência do substrato para o DLS das Hsp70; 3) comunicação da ligação do substrato transmitido pelo DLS ao DLN, o que resulta na estimulação completa da hidrólise de ATP seguida pelo fechamento do DLS2

sobre o substrato e 4) dissociação da Hsp40 do complexo formado entre Hsp70-ADP-substrato. Entretanto, devido à grande diversidade das Hsp40 e a existência de múltiplos contatos entre as Hsp40-Hsp70 (Suh *et al.*, 1999; Qian *et al.*, 2002; Linke *et al.*, 2003), tal ciclo funcional pode ser constituído de mais etapas, dependendo das proteínas em questão.

A figura 1a, apresentada no Anexo II, sumariza a arquitetura de domínios das Hsp40 e as classifica em 3 subfamílias, a saber: A) proteínas que apresentam todos os domínios descritos acima; B) proteínas que não possuem o domínio rico em cisteínas e C) aquelas que apresentam somente o domínio J, porém não necessariamente localizado na região amino-terminal da proteína (Cheetham *et al.*, 1998; Mayer *et al.*, 2001). As Hsp40 das subfamílias A e B são encontradas na forma oligomérica de dímeros, cujo sítio de dimerização ocorre pela interação de aminoácidos localizados na extremidade carboxi-terminal (Sha *et al.*, 2000; Li *et al.*, 2003).

GrpE

As GrpE são co-chaperones específicas para a subfamília DnaK das Hsp70 (Brehmer *et al.*, 2001) e são essenciais para a viabilidade de bactérias em todas as temperaturas (Langer *et al.*, 1992). A GrpE age como um fator de troca do nucleotídeo ADP ligado à DnaK (Szabo *et al.*, 1994), sendo importante para a reciclagem desta última (Liberek *et al.*, 1991). A estrutura em alta resolução da GrpE de *E. coli* (EcGrpE – entre os resíduos 34-197), em complexo protéico com o DLN da DnaK de *E. coli*, mostra a EcGrpE como um dímero rico em estrutura secundária do tipo α -hélice e com um formato alongado (Harrison *et al.*, 1997). A figura 2 representa, de maneira esquemática, a estrutura de GrpE de *E. coli* e a caracteriza em 4 principais regiões. A região carboxi-terminal (EcGrpE₁₃₈₋₁₉₇) é formada principalmente por estrutura secundária do tipo folha β -pregueada e é capaz de interagir com o DLN da DnaK induzindo um movimento de abertura do DLN (Harrison *et al.*, 1997). Desta forma, a interação da GrpE com a DnaK auxilia na liberação da molécula de ADP ligada ao DLN (Harrison *et al.*, 1997; Brehmer *et al.*, 2001). A região central da proteína (EcGrpE₈₉₋₁₃₇) é formada por um feixe de 4 hélices que possivelmente age como um centro para a dimerização da proteína (Gelinas *et al.*, 2002; Mehl *et al.*, 2003). A região amino-terminal (EcGrpE₄₀₋₈₈) é formada por duas longas α -hélices que parecem estar conectadas ao domínio carboxi-terminal e funcionam como um sensor térmico, como discutido abaixo (Grimshaw *et al.*, 2001; Gelinas *et al.*, 2002; Gelinas *et al.*, 2003; Grimshaw *et al.*, 2003).

Para a GrpE de *T. thermophilus* o domínio carboxi-terminal é o responsável pela função de termo-sensor (Groemping & Reinstein, 2001). Os primeiros 34 resíduos da GrpE estão ausentes na estrutura em alta resolução, sugerindo que esta região poderia não possuir estrutura secundária bem definida, pois o cristal só foi obtido após a proteólise limitada daquela região (Harrison *et al.*, 1997). Entretanto, esta região é importante para a interação com o DLS da DnaK reduzindo a afinidade desta pelos substratos. Esta propriedade pode explicar a necessidade do formato alongado apresentada pela GrpE (Harrison *et al.*, 1997; Mally & Witt, 2001; Chesnokova *et al.*, 2003). A outra possibilidade recentemente levantada é que as GrpEs podem evitar a interação de substratos com o DLS das Hsp70 (Brehmer *et al.*, 2004).

A GrpE é capaz de agir como um sensor térmico para o sistema Hsp70, de forma que as Hsp70 possam manter os substratos em solução nas condições de estresse térmico (Grimshaw *et al.*, 2001; Groemping & Reinstein, 2001; Grimshaw *et al.*, 2003). Um aumento de temperatura no meio provoca uma perda de estrutura nas longas hélices, desestabilizando o domínio carboxi-terminal e reduzindo a sua capacidade de interação com o DLN da DnaK (Gelinas *et al.*, 2003). Assim, a DnaK, em células submetidas ao estresse térmico, estaria permanentemente ligada ao ADP e também ao substrato. Deste modo, a DnaK evita a agregação do substrato e não o libera para o meio até que ocorra a redução da temperatura e a reativação da GrpE (Grimshaw *et al.*, 2001). Em baixas temperaturas, a GrpE recupera a estrutura nas longas hélices, recuperando a estabilização do domínio carboxi-terminal e a sua capacidade de interagir com o DLN da DnaK. Assim, após a liberação do ADP do DLN da DnaK e a ligação do ATP a este domínio, o substrato é liberado do DLS para o meio, podendo se enovelar (Grimshaw *et al.*, 2001).

A maior parte das informações experimentais disponíveis sobre as GrpE se referem ao exemplar de *E. coli* e em menor grau àquele de *T. thermophilus*, organismos que possuem apenas uma isoforma da proteína. Entretanto, organismos eucarióticos possuem mais de uma isoforma de GrpE, ambas localizadas na mitocôndria (Schlicher & Soll, 1997; Naylor *et al.*, 1998; Choglay *et al.*, 2001). As proteínas de mamíferos têm a habilidade de interagir e estimular a atividade de hidrólise de ATP pela Hsp70 mitocondrial (Naylor *et al.*, 1998) que pertence à subfamília DnaK (Brehmer *et al.*, 2001). Embora a necessidade de duas isoformas de GrpE na mitocôndria não esteja ainda solucionada, Naylor *et al.* (1998) mostra que elas têm diferentes padrões de expressão e sugere que elas podem ser ativas em diferentes situações e/ou reguladas de maneira pós-transcricional e/ou pós-traducional.

Não foi identificada nenhuma proteína homóloga a GrpE localizada no citoplasma de células eucarióticas, porém existe uma proteína capaz de influenciar a troca de nucleotídeos adenosina das Hsp70 citoplasmáticas. Tal proteína recebe o nome de Bag-1 e pertence a uma família de proteína antiapoptóticas que se ligam a um grande número de proteínas celulares (Fink *et al.*, 1999). A Bag-1 existe na forma monomérica (Stuart *et al.*, 1998) e é formada por um feixe de 3 α -hélices. Ela é capaz de deslocar o ADP ligado à Hsp70 citoplasmática e também recrutar as Hsp70 para alvos específicos (Briknarova *et al.*, 2001; Takayama & Reed, 2001).

1.3 Técnicas biofísicas utilizadas para estudos de estrutura-função

As proteínas apresentam várias propriedades biofísicas que podem ser analisadas através da interação com ondas eletromagnéticas, liberação ou absorção de calor e também pela interação da proteína com o meio solúvel no qual que ela se encontram. Para a realização deste trabalho, foram avaliadas algumas propriedades biofísicas das proteínas em solução, para a caracterização da relação estrutura-função das chaperones moleculares humanas Hsp70, Hsp40 e GrpE. Abaixo é apresentada uma breve introdução teórica sobre as técnicas utilizadas.

1.3.1 Dicroísmo circular

Uma molécula (cromóforo) cuja estrutura é quiral, ou seja, que possui um centro de assimetria, é capaz de interagir com a luz circularmente polarizada a direita e a esquerda, assim é a técnica de espectropolarimetria de dicroísmo circular (CD) (para revisão: Woody, 1995; Kelly & Price, 1997). As proteínas são moléculas que satisfazem o requisito acima devido à presença, em sua estrutura, de dois cromóforos principais: as ligações amídicas e as estruturas aromáticas. As primeiras são responsáveis pelo sinal característico no comprimento de onda (λ) do ultravioleta distante (abaixo de 250 nm) sendo utilizadas como sondas para estruturas secundárias de proteínas. As estruturas aromáticas das proteínas são responsáveis pelo sinal no ultravioleta próximo (entre 250-300 nm) e são utilizadas para avaliar a estrutura terciária de proteínas. Considerando estas características, a técnica de CD é utilizada para estimar a quantidade de estrutura secundária de proteínas e avaliar as mudanças conformacionais ocasionadas por ligantes, agentes desnaturantes (acidez, uréia, cloridrato de guanidina – Gdn-HCl, etc) e temperatura (Woody, 1995; Kelly & Price, 1997).

Através da análise dos sinais característicos de diferentes estruturas secundárias nos espectros de CD de proteínas foi mostrada a possibilidade de classificá-las em 5 classes (Levitt & Chothia, 1976): 1) toda α (principalmente α -hélices), 2) toda β (principalmente folhas β -pregueadas), 3) $\alpha + \beta$ (regiões α e β separadas), 4) α/β (regiões com α e β misturadas), e 5) randômica (predominantemente desordenada). Esta avaliação também pode ser feita com o auxílio de programas computacionais que avaliam o espectro de CD de uma proteína. Estes programas comparam o espectro a ser analisado com um banco de dados de espectros de CD de proteínas que já possuem suas estruturas em alta resolução determinadas. Este procedimento aumenta a precisão na estimativa da quantidade de cada tipo de estrutura (para revisão: Greenfield, 1996). A avaliação da estabilidade de uma proteína, por CD, pode fornecer informações sobre a presença de domínios com diferentes estabilidades e sobre processos dinâmicos, como interações com ligantes (Woody, 1995).

1.3.2. Fluorescência

A fluorescência é uma técnica espectrofotométrica baseada em um fenômeno quântico no qual um elétron de um determinado cromóforo, após absorver uma determinada quantidade de energia, atinge um estado excitado e, ao retornar ao seu estado basal, emite parte da energia absorvida na forma de luz. O restante da energia absorvida pelo cromóforo pode ser perdida no retorno dos elétrons excitados em outros processos que a emissão de radiação eletromagnética como: calor ou processos químicos e vibracionais. Desta forma, a quantidade de energia emitida na forma de ondas eletromagnéticas é relativamente inferior à energia absorvida (energia de excitação) e o λ da luz emitida é maior do que aquele da luz absorvida. Naturalmente, existem 3 aminoácidos susceptíveis ao processo de fluorescência: principalmente triptofano, tirosina e em menor escala fenilalanina. Porém, é possível utilizar sondas fluorescentes ligadas a determinados aminoácidos, aminoácidos modificados ou compostos que emitem fluorescência quando interagem com a proteína (Lakowicz, 1983).

O aminoácido triptofano é sensível ao ambiente no qual ele se encontra, apresentando diferentes características de fluorescência que são dependentes do grau de exposição ao meio hidrofílico ou hidrofóbico. Assim, aliado ao seu alto rendimento quântico, sua baixa frequência em proteínas e possibilidade de excitação específica (ver abaixo), a técnica de fluorescência do triptofano se tornou uma das mais utilizadas em biofísica molecular, principalmente em estudos relacionados a mudanças conformacionais de proteínas. Quando

o resíduo de triptofano se encontra bem exposto ao solvente, ele gasta uma quantidade maior de energia no estado excitado com a reorganização das moléculas de água que o envolvem e emite fluorescência em λ menos energéticos (próximos a 355 nm). Já quando o resíduo de triptofano se encontra em uma região no interior da proteína, geralmente hidrofóbica e com baixa acessibilidade a água, a emissão de fluorescência ocorrerá em comprimentos de onda mais energéticos (310 nm a 335 nm), pois não haverá grande perda de energia com a reorganização de moléculas de água. A mudança de exposição do triptofano ao ambiente pode ser induzida por ligantes ou pela desnaturação e fornecem informações sobre as mudanças conformacionais e sobre a estabilidade da proteína (Lakowicz, 1983).

Os principais λ de excitação para o triptofano, tirosina e fenilalanina estão no intervalo entre 260 nm e 295 nm. Porém, é possível separar bem as contribuições nos espectros de emissão entre o triptofano e tirosina. A contribuição da fenilalanina é muito baixa quando comparada com o triptofano e tirosina. Utilizando λ de excitação próximos a 260 nm é possível excitar principalmente resíduos de tirosina e em 295 nm os resíduos de triptofano serão preferencialmente excitados. Se uma determinada proteína possui apenas tirosinas, o máximo de emissão de fluorescência se encontra em torno de 305-310 nm no estado nativo e 335 nm no estado desnaturado (dependente da exposição das tirosinas ao solvente). Porém o rendimento quântico da fluorescência do triptofano é maior que para as tirosinas. Assim, dependendo da proporção destes fluoróforos e da proximidade entre eles, é possível observar comportamentos diferenciados do espectro de emissão, tais como: supressão de fluorescência, deslocamento do máximo de emissão ou mudanças no centro de massa espectral de fluorescência. A acessibilidade do solvente ao triptofano e em menor escala à tirosina, também podem ser determinadas utilizando agentes supressores de fluorescência tais como os iodetos, acrilamida, entre outros (Lakowicz, 1983).

1.3.3 Calorimetria

Calorimetria é a área da termodinâmica que estuda a troca de calor envolvido em uma reação química, seja ele oriundo do processo de interação entre dois agentes químicos ou nas mudanças estruturais ocorridas durante a reação. Para sistemas biológicos, existem duas técnicas principais que são utilizadas para obter informações termodinâmicas a partir de diferentes aparelhos. Porém, o princípio funcional destes aparelhos é baseado na

manutenção do equilíbrio térmico entre duas células revestidas por jaquetas adiabáticas: a célula da amostra e a da referência. Se ocorrer um processo exo ou endotérmico na célula da amostra, o aparelho detecta a diferença entre as células e compensa a diferença na célula de amostra. A diferença de compensação, em watts, constitui o sinal observado e que é analisado pelo equipamento. Várias informações termodinâmicas podem ser obtidas pela calorimetria, que aliada a informações estruturais, podem fornecer informações sobre estabilidade, função e identificação de regiões importantes para interações proteína-ligante e/ou proteína-proteína (Pierce *et al.*, 1999).

A calorimetria de titulação isotérmica (ITC) é utilizada para medir diretamente o calor liberado ou absorvido por uma interação química. A reação é iniciada pela adição de um componente químico que interage com o receptor. Quando estes componentes interagem, a energia envolvida, na forma de calor, pode ser liberada ou absorvida. O calor envolvido na interação, por mol de ligante injetado, é analisado em função da razão molar entre ligante/receptor e é proporcional à quantidade de ligante injetado. Nas primeiras adições, todo o ligante adicionado interage com os receptores livres e como ocorrem adições subseqüentes do ligante – titulação – os receptores ficam saturados e o calor envolvido na interação diminui até alcançar o calor de diluição do ligante no tampão. Este processo é dependente da concentração do ligante e do receptor. A variação da entalpia (ΔH), estequiometria da ligação (n) e a constante de associação (K_A) podem ser obtidos simultaneamente em um mesmo experimento. A variação da energia livre de Gibbs (ΔG) e a variação da entropia (ΔS) da interação podem ser calculadas a partir dos valores experimentais seguindo a relação abaixo (para revisão: Pierce *et al.*, 1999):

$$\Delta G = -RT \ln K_A = \Delta H - T\Delta S \quad [\text{Equação 1}]$$

onde R é a constante dos gases e T é a temperatura absoluta. As medidas da ΔH e ΔS em diferentes temperaturas podem fornecer informações importantes em relação a variação da capacidade calorífica (ΔC_p) pela seguinte relação:

$$\Delta C_p = \frac{d(\Delta H)}{dT} = T \frac{d(\Delta S)}{dT} \quad [\text{Equação 2}]$$

A ΔC_p é um bom indicador de mudanças hidrofóbicas provocadas pela interação proteína-ligante e fornece informações sobre a hidratação do sistema ou mudanças neste coeficiente (para revisão: Pierce *et al.*, 1999; Jelesarov & Bosshard, 1999).

A técnica de calorimetria de varredura diferencial (DSC) é usada para medir a ΔC_p de um sistema submetido a um aumento contínuo de temperatura. O aparelho mantém a temperatura das células sempre iguais e neste processo de compensação de energia, o calorímetro detecta a energia liberada pelo sistema ocasionada por um processo de desenovelamento das proteínas. Se o processo de desenovelamento térmico da proteína for reversível, é possível, através da integração da curva da C_p em função da temperatura, obter a ΔH e a ΔS envolvidas no processo através das seguintes equações, respectivamente:

$$\Delta H = \int_{T_1}^{T_2} C_p dT \quad [\text{Equação 3}]$$

e

$$\Delta S = \int_{T_1}^{T_2} \frac{C_p}{T} dT \quad [\text{Equação 4}]$$

Com os dados de ΔH e ΔS do sistema e com o auxílio da equação 1, é possível determinar a ΔG do desenovelamento da proteína, um importante parâmetro para quantificar e comparar a estabilidade de proteínas (Santoro & Bolen, 1992). Outro parâmetro extraído pelo DSC é a temperatura média de transição ($T_m = \Delta H/\Delta S$, quando $\Delta G = 0$), sendo o pico máximo da fase de transição no desenovelamento térmico. No T_m , obedecida as condições de equilíbrio e em um sistema de dois estados, 50% das proteínas estão enoveladas e outros 50% se encontram desenoveladas. Em uma análise de transição de dois estados existe o equilíbrio somente entre os estados nativo e desenovelado (Shnyrov *et al.*, 1997).

1.3.4 Ultracentrifugação analítica

A técnica de ultracentrifugação analítica (UAL) é poderosa para entender o comportamento de macromoléculas em solução. A UAL é uma técnica que pode fornecer informações importantes para a caracterização de proteínas em solução. Através desta técnica é possível determinar a *MM* de uma macromolécula e obter informações sobre a heterogeneidade do sistema em relação à sua massa, ou seja, de quantas subunidades o

sistema é formado e qual a proporção entre elas. Esta técnica também ajuda a determinar se a macromolécula forma agregados e se estes são reversíveis ou irreversíveis. O coeficiente de sedimentação (s), é um parâmetro específico da partícula, sendo particularmente útil para monitorar mudanças conformacionais de proteínas (Laue, 2001). Informações sobre a termodinâmica de um sistema associativo também podem ser adquiridas, pois é possível obter a estequiometria e a constante de equilíbrio envolvida na interação. Tais informações podem ser obtidas em função de alterações provocadas no meio, por exemplo, modificando a temperatura ou adicionando ligantes, modificando o pH, etc. Várias revisões de qualidade estão disponíveis tais quais as escritas por Laue & Stafford (1999), Stafford (1997), Cole & Hansen (1999), Schuck (2000), Laue (2001), e Lebowitz *et al.* (2002).

A técnica de UAL é absoluta, isto é, não necessita de padrões externos e se baseia em teorias bem fundamentadas da física de movimentos. Quando um objeto está sob centrifugação, é possível caracterizar algumas forças às quais o objeto está submetido. Em condições de velocidade constante, a partícula é forçada a migrar para o fundo do tubo pela força centrífuga (F_C) que depende da massa da partícula e aceleração. Durante a sua migração, a partícula sofre a ação de forças que retardam o seu progresso: força de empuxo (F_{Emp}) e da força friccional (F_f). A F_{Emp} depende unicamente da densidade do fluido (ρ_f) e do volume do objeto em função da massa e não depende da composição ou da forma do objeto, sendo igual ao peso do fluido deslocado. A F_f é caracterizada pelo atrito que a partícula provoca ao sistema, sendo proporcional à velocidade v do objeto e depende do formato da partícula. Em uma situação de velocidade constante, estas forças se anulam e a seguinte equação pode ser feita:

$$F_C + F_{Emp} + F_f = \omega^2 r m + (-\omega^2 r m V_{bar} \rho_f) + (-fv) = 0 \quad [\text{Equação 5}]$$

onde ω é a velocidade angular; r é o raio; m é igual a massa do objeto; V_{bar} é o volume parcial específico da partícula e f é o coeficiente friccional. O sinal negativo da F_{Emp} e F_f indica que estas forças são contrárias à F_C . Portanto, a velocidade de uma partícula nestas condições pode ser definida pelo rearranjo da equação 5 em função da velocidade:

$$v = \frac{[\omega^2 r m (1 - V_{bar} \rho_f)]}{f} \quad [\text{Equação 6}]$$

Considerando a massa $m = MM/N_{Av}$, onde MM é a massa molecular e N_{Av} é o número de Avogrado e rearranjando a equação 6, a seguinte equação pode ser obtida, sendo ela conhecida como equação de Svedberg:

$$\frac{MM(1 - V_{bar}\rho_f)}{N_{Av}f} = \frac{v}{\omega^2 r} \equiv s \quad [\text{Equação 7}]$$

O termo $v/\omega^2 r = s$, é a velocidade da partícula por unidade de aceleração gravitacional, chamado de coeficiente de sedimentação s e é dado em 10^{-13} segundos ou Svedberg (S), em homenagem ao “pai” da UAL, Theodore Svedberg. O s aumenta em função da massa da partícula e é inversamente proporcional ao f . Em outras palavras: uma alta MM aumenta o s , enquanto a forma assimétrica o reduz (veja a seguir).

Se for considerado que o f é o fator responsável pelo atrito provocado ao sistema e é dependente do tamanho e da forma do objeto, partículas alongadas ou volumosas apresentarão maior atrito ao sistema quando comparado a partículas compactas e/ou esféricas de mesma MM . Assim, informações sobre a forma da partícula podem ser obtidas diretamente em experimentos de UAL, pela monitoração direta do s . Uma esfera é a forma que sofre a menor fricção e seu f , também conhecido como f_0 , depende de seu raio r e pode ser dado pela lei de Stokes (R_S é o raio hidrodinâmico ou raio de Stokes):

$$f_0 = 6\pi\eta R_S \quad [\text{Equação 8}]$$

onde o η é a viscosidade da solução. Outras formas (elipsóides, oblatas, prolatas, cilíndricas, etc) podem ser determinadas em relação a uma esfera, de mesma MM , pela razão friccional f/f_0 , que, para uma relação do tipo comprimento/altura, será igual ou maior que 1.

O f está, por sua vez, relacionado ao coeficiente de difusão (D) de uma partícula pela seguinte equação, conhecida como equação de Einstein:

$$D = \frac{RT}{N_{Av}f} \quad [\text{Equação 9}]$$

onde R é a constante dos gases e T é a temperatura absoluta. O D pode ser medido experimentalmente a partir da técnica de espalhamento dinâmico de luz (EDL) (Schönfeld & Behlke, 1998). Rearranjando as equações de Svedberg (equação 7) e de Einstein (equação 9) em função do f é possível definir a seguinte equação:

$$MM = \frac{sRT}{D(1 - V_{bar}\rho_f)} \quad [\text{Equação 10}]$$

Assim, a MM de uma partícula pode ser diretamente estimada a partir das medidas experimentais do s e do D .

Pela utilização dos princípios físicos descritos acima, dois tipos principais de técnicas podem ser aplicados utilizando a UAL. A velocidade de sedimentação, uma técnica baseada em princípios hidrodinâmicos, que utiliza altas velocidades para sedimentar uma partícula e o fluxo de sedimentação da partícula é acompanhado. Utilizando esta técnica é possível obter informações quanto à MM e forma da partícula, diretamente através do valor de s .

A técnica de sedimentação em equilíbrio é baseada em princípios termodinâmicos, e através da utilização de velocidades menores é possível produzir uma pequena sedimentação que contrapõem a força de sedimentação e a força de difusão da partícula. No estado de equilíbrio provocado pela lei da ação das massas, é possível obter informações sobre a MM (de maneira independente da forma da partícula), sobre as constantes de afinidade e também da estequiometria do sistema associativo.

1.3.5 Espalhamento de raios-X a baixos ângulos

A técnica de espalhamento de raios-X a baixos ângulos (SAXS) fornece informações estruturais a partir das flutuações ou heterogeneidades da densidade eletrônica de partículas com dimensões características da ordem de 10 a 1.000 angströms (Å). Vários tipos de partículas apresentam tais flutuações nesta escala de dimensões, entre elas as proteínas. Os dados obtidos a partir das curvas de SAXS contêm informações sobre o tamanho e a forma das partículas espalhadoras. Para soluções de proteínas em condições não-ideais nas quais as proteínas compõem sistemas de interação interparticulares, dados sobre a correlação espacial da interação também podem ser obtidas (para revisão: Koch *et al.*, 2003; Svergun & Koch, 2003).

As ondas eletromagnéticas associadas aos raios X incidentes provocam oscilações dos elétrons dos átomos componentes da amostra irradiada. Quando a energia dos fótons incidentes é bem diferente da energia de ligação dos elétrons, estes se comportam como se fossem livres. Estas partículas livres carregadas oscilantes produzem ondas espalhadas coerentes em relação à onda incidente em todas as direções com intensidade definida pela lei de Thompson, resultando em um efeito elástico. Nestas condições o feixe espalhado não sofre mudança no λ incidente, ou seja, não perde energia. Também existe o espalhamento inelástico ou Compton, com mudança no λ do feixe espalhado em relação ao incidente, além de outros tipos de espalhamento. Para ângulos pequenos o espalhamento inelástico é pequeno e pode ser desprezado. Geralmente para experiências de espalhamento em solução, o ângulo de espalhamento (q), chamado de 2Θ é ≤ 10 graus (Figura 3a) para um λ de aproximadamente 1,5 Å (Koch *et al.*, 2003; Svergun & Koch, 2003).

As proteínas são partículas compostas por átomos leves (H, C, N, O e P) e apresentam densidade eletrônica constante (ρ_e) de aproximadamente 0,40 elétrons/Å³. As soluções aquosas utilizadas para acondicionar proteínas, contendo pequenos solutos, apresentam ρ_e média de 0,33 elétrons/Å³. Ou seja, entre as proteínas e o meio aquoso existe um contraste médio na ρ_e de aproximadamente 0,07 elétrons/Å³. Em soluções contendo 1 mol/L de NaCl, por exemplo, a ρ_e da solução passa de 0,33 para 0,35 elétrons/Å³, diminuindo ainda mais o contraste eletrônico entre a proteína e o solvente, exigindo mais tempo de coleta ou soluções protéicas mais concentradas (nem sempre é possível concentrar muito, pois senão as proteínas interagem ou agregam mudando o padrão de espalhamento). Portanto, as soluções protéicas produzem um espalhamento muito fraco. Por exemplo, uma solução de 1 mg/mL de uma proteína de 80 kDa espalha 4 fótons em cada 100.000 incidentes. Isto explica imediatamente a grande vantagem de se usar feixes de raios X produzidos por fontes síncrotron que possuem um fluxo de fótons muito maior do que aparelhos de raios X convencionais.

No caso de espalhamento de partículas em solução, as amostras são isotrópicas, ou seja, as partículas estão em movimento browniano e podem assumir todas as orientações possíveis em relação ao feixe incidente. Através da promediação em todas as orientações possíveis em relação ao feixe incidente, a expressão da intensidade de espalhamento $I(q)$ é dada por:

$$I(q) = 4\pi \int_0^{\infty} p(r) \frac{\sin rq}{rq} dr \quad [\text{Equação 11}]$$

onde $q = |\vec{q}| = 4\pi \sin \theta / \lambda$ é o módulo do vetor de espalhamento, sendo que para ângulos pequenos é possível aproximar $q \approx 2\pi(2\theta) / \lambda$. Através da aplicação de uma transformada de Fourier indireta na equação acima é possível obter a seguinte relação:

$$p(r) = r^2 V \gamma(r) = 2 \int_0^{\infty} r q I(q) \sin rq dq \quad [\text{Equação 12}]$$

onde $p(r)$ é a função de distribuição de distâncias ou a função de distribuição de pares. Para partículas homogêneas ($p(r) = \text{constante}$), a curva da $p(r)$ representa o histograma de distâncias entre todos os pares de pontos (elementos de volume) da amostra (Figura 3c). Assim, é possível identificar a distância máxima ou diâmetro máximo (D_{max}) da partícula como sendo o ponto em que a curva atinge o eixo das distâncias (eixo da coordenada). A função $y(r)$ é chamada de função de auto-correlação ou função característica da partícula e representa a probabilidade de se encontrar um ponto (elemento de ρ_e) a uma distância r de um dado ponto (Figura 3b-c). Através da integração desta função probabilística sobre a superfície de uma esfera de raio r e sobre o volume V é obtida a curva da $p(r)$ como definida na equação 12, que pode fornecer uma idéia do tamanho e a forma da proteína. A curva da $p(r)$ contém as mesmas informações que a curva da $I(q)$, entretanto a primeira possui uma representação espacial mais intuitiva e fornece informações sobre o tamanho e forma da partícula pela análise direta do perfil da curva (Svergun & Koch, 2003). Por exemplo, partículas globulares apresentam curvas da $p(r)$ com um formato de sino; partículas alongadas resultam em curvas da $p(r)$ cujo máximo está deslocado para pequenas distâncias; enquanto formas planas também mostram o máximo da curva da $p(r)$ deslocado para pequenas distâncias, porém com o formato da curva assume uma forma de máximo largo. Curvas da $p(r)$ com múltiplos máximos podem representar subunidades ou domínios de diferentes tamanhos relativamente separados, onde um máximo mostra as distâncias intra-subunidades e as demais as relações entre as subunidades (Figura 3c; Svergun & Koch, 2003).

Um importante parâmetro estrutural que pode ser obtido pela técnica de SAXS e pode ser utilizado na análise de mudanças conformacionais de macromoléculas em solução, é o raio de giro (R_g). O R_g de uma macromolécula pode ser definido de forma semelhante ao “momento de inércia” da mecânica clássica, ou seja, como sendo a raiz quadrada da distância quadrática média entre os elétrons da partícula e o centro de massa eletrônico do objeto espalhador:

$$R_g = \left[\frac{1}{\Delta m} \int \Delta \rho(r) |r|^2 dV_r \right]^{1/2} \quad [\text{Equação 13}]$$

onde $\Delta m = (m_1 - \rho_e V)$, sendo m_1 o número total de elétrons e V o volume da macromolécula. O valor de R_g pode ser obtido a partir da porção da curva da $I(q)$ a muito baixos ângulos para a qual é válida a aproximação de Guinier (Guinier & Fournet, 1955):

$$I(q) = N(\Delta m)^2 e^{-\frac{1}{3} q^2 R_g^2} \quad [\text{Equação 14}]$$

A curva da $p(r)$ fornece informações mais confiáveis sobre o R_g . Porém, em soluções diluídas de proteínas ($\leq 1\text{mg/mL}$), os erros estatísticos nas medidas da $I(q)$ são altos e não é possível coletar informações com qualidade suficiente numa ampla faixa de q para poder determinar uma $p(r)$ confiável. Neste caso, a curva experimental fica estatisticamente confiável somente em ângulos muito baixos ($q \rightarrow 0$) e o R_g deve ser obtido a partir da lei de Guinier (equação 14).

As partículas em solução podem apresentar diferentes graus de compactidade e os dados de SAXS, quando analisados pelo perfil do gráfico de Kratky, podem fornecer tais informações. O gráfico de Kratky é obtido pela relação $I(q) \times q^2$ versus q e os perfis das curvas resultantes são característicos para proteínas compactas, com segmentos flexíveis ou proteínas completamente desenoveladas. Proteínas compactas apresentam o gráfico de Kratky com um perfil que lembra um sino, cujos máximos e mínimos são bem definidos. Partículas parcialmente enoveladas não apresentam um máximo e mínimo bem definidos, o que sugere a presença de segmentos com flexibilidade na proteína espalhadora. Partículas desenoveladas não apresentam máximos e mínimos e a curva tende a obedecer a uma função de crescimento linear (Doniach *et al.*, 1995).

Os dados de SAXS podem ser utilizados para construir modelos *ab initio* das proteínas analisadas. Tais modelos são de baixa resolução e detalhes atômicos das proteínas não podem ser obtidos. Todavia, podem ser obtidas informações sobre o tamanho, formato e sobre as posições relativas dos domínios da partícula (Svergun, 1999; Svergun & Koch, 2003). As informações contidas nas curvas de SAXS também são úteis para interpretar e confrontar, através de programas computacionais, as possíveis diferenças entre a estrutura de proteínas em solução e aquela apresentada em alta resolução (Svergun & Koch, 2003).

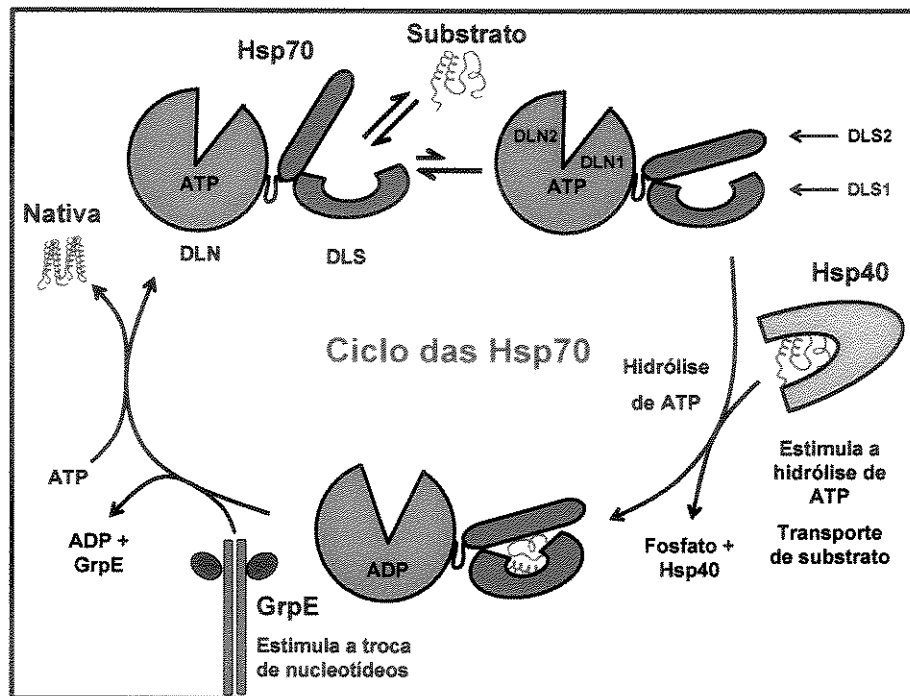


Figura 1: Modelo para o ciclo de ação das chaperones moleculares do sistema Hsp70 (modificado de Mayer *et al.*, 2000). As Hsp70 podem ser divididas em dois domínios principais, o domínio de ligação de nucleotídeos (DLN) que possui uma forma de “concha” e o domínio de ligação de substratos (DLS). Na presença de ATP ligado no interior da fenda da “concha” do DLN, a Hsp70 sofre pequenas mudanças conformacionais, as quais transmitiriam informações ao DLS, deslocando o equilíbrio do subdomínio DLS2 para uma conformação aberta. Nesta conformação, a Hsp70 apresenta um alto fluxo de associação e dissociação pelo substrato (Theyssen *et al.*, 1996; Mayer *et al.*, 2000). O substrato livre pode se ligar à Hsp70 a partir do meio ou pode ser transportado pela Hsp40 para a Hsp70. Através de um estímulo concomitante do substrato e da Hsp40, a Hsp70 hidrolisa o ATP em ADP + Pi. Na presença de ADP ligado na fenda do DLN, o equilíbrio conformacional do subdomínio DLS2 seria deslocado para uma conformação fechada sobre o DLS1, retendo o substrato ligado à Hsp70 e prevenindo a agregação. Algumas Hsp70 necessitariam da ação de co-chaperones GrpE (ou Bag-1) para auxiliar na dissociação da molécula de ADP ligada no DLN, possibilitando a ligação de ATP ao sítio, que deslocaria o equilíbrio conformacional do subdomínio DLS2 para a conformação aberta. Assim, o substrato pode ser liberado para o meio para se enovelar, ou repetir o ciclo com a Hsp70 ou com outra chaperone molecular (Anexo I – Borges & Ramos, 2004).

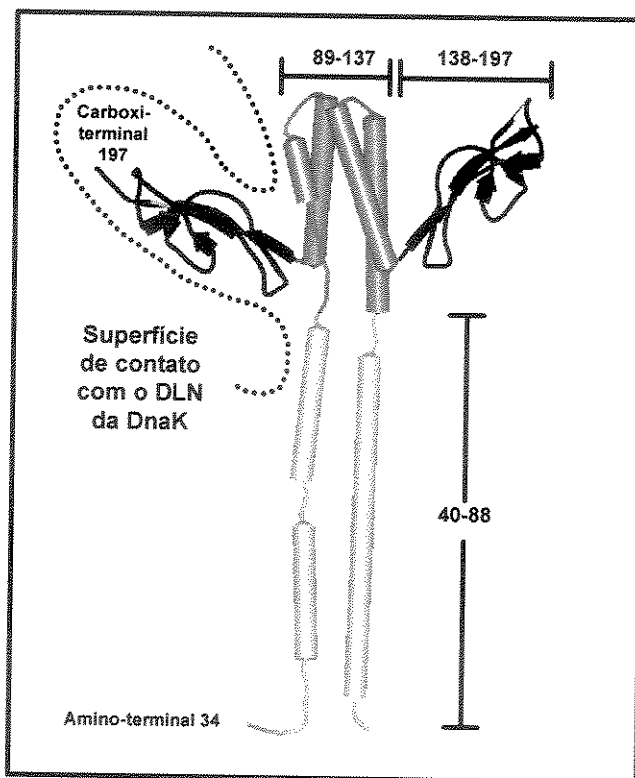


Figura 2: Diagrama da estrutura do dímero da GrpE de *E. coli*. A figura foi produzida usando o programa WebLab ViewerLite a partir do arquivo depositado no bando de dados *Protein Data Bank* (PDB) código nº 1DKG referente à estrutura em alta resolução da EcGrpE (EcGrpE₃₄₋₁₉₇) em complexo protéico com o DLN da DnaK (Harrison *et al.*, 1997). As regiões da EcGrpE são designadas como se segue: cinza claro, EcGrpE₄₀₋₈₅ formada pelas longas α -hélices; cinza escuro, feixe de 4 α -hélices (EcGrpE₈₉₋₁₃₇); em preto, o domínio carboxi-terminal estruturado em folha β -pregueada (EcGrpE₁₃₈₋₁₉₄). O DLN da DnaK não está representado na figura, mas a linha pontilhada representa a região a qual ele está ligado: o monômero proximal da GrpE, o outro é denominado como monômero distal. Note a diferença de angulação do domínio carboxi-terminal entre os monômeros proximal e distal. A região amino-terminal não foi obtida no cristal devido a uma etapa de proteólise da EcGrpE e foi sugerido que esta região não apresenta estrutura organizada (Harrison *et al.*, 1997).

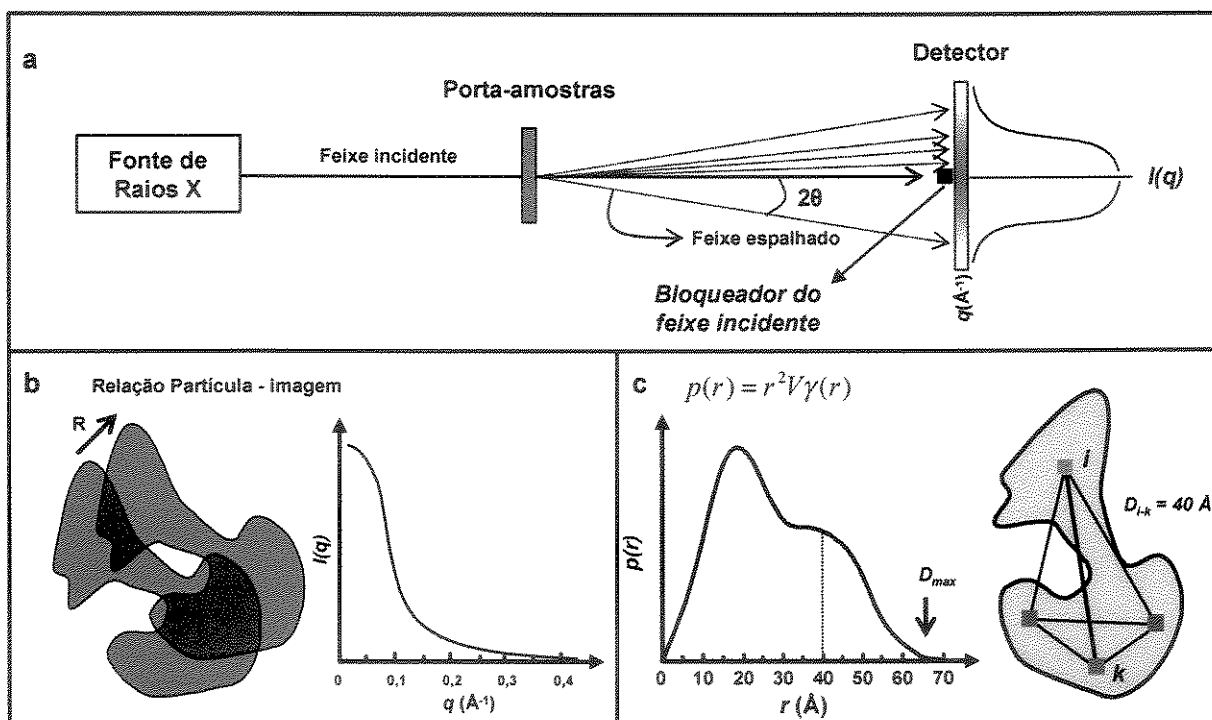


Figura 3: Esquema de um experimento de SAXS, relação partícula-imagem e função de distribuição de distâncias – $p(r)$. a) Desenho esquemático de um experimento de SAXS. Repare o ângulo formado entre o feixe incidente e o feixe espalhado sendo definido como 2θ . A função do “bloqueador do feixe incidente” é evitar que este, devido a sua alta intensidade, atinja o detector e danifique-o. b) Relação partícula-imagem e a curva da $I(q)$ (direita) das curvas de espalhamento de raios X a baixos ângulos; c) função de distribuição de distâncias da partícula resultante da função de autocorrelação e o modelo *ab initio* que interpreta tal curva (direita). Note a presença de dois máximos devidos à presença de dois domínios relativamente separados na partícula e a caracterização do D_{max} como sendo o ponto onde a curva atinge o eixo das coordenadas. À direita, a partícula com elementos de volume i - k com distância de $40 \text{ \AA}$. A curva da $p(r)$ representa a frequência da identificação da distância entre estes elementos de volume.

2 OBJETIVOS

O objetivo geral deste trabalho foi obter proteínas do sistema Hsp70 (Hsp70, Hsp40 e GrpE) oriundas de humano e estudar as propriedades estruturais das proteínas para compreender mais sobre a relação entre a estrutura das proteínas e suas funções. Para isso, foram utilizadas técnicas de biofísica molecular, tais como: dicroísmo circular, fluorescência, calorimetria, ultracentrifugação analítica e espalhamento de raios X a baixos ângulos.

Os objetivos específicos deste trabalho foram:

- 1) Produzir duas proteínas da família Hsp70: uma representante da subfamília Hsc70 (Hsp70-1A) que são localizadas no citoplasma; e uma pertencente à subfamília DnaK (mortalina). Um outro objetivo para se obter a mortalina foi para desenvolver estudos de interação com as GrpEs humanas.
- 2) Produção de duas proteínas da família Hsp40. Uma pertencente à subfamília A (DjA1), que apresenta as principais regiões conservadas das Hsp40 (domínio J, região rica em glicinas e fenilalaninas, domínio rico em cisteínas e o domínio DnaJ_C). A outra proteína pertence à subfamília B (DjB4), que apresenta as regiões conservadas descritos acima, exceto aquela rica em cisteínas.
- 3) Obter um mutante de deleção da região carboxi-terminal da DjA1 (DjA1₁₋₃₃₂) para analisar as propriedades oligoméricas da DjA1, pela eliminação do sítio de dimerização daquelas proteínas.
- 4) Produzir e caracterizar as duas GrpEs, GrpE#1 e GrpE#2, que são codificadas pelo genoma humano.

3 MATERIAIS E MÉTODOS

Neste tópico estão apresentados os Materiais e Métodos gerais utilizados no desenvolvimento dos procedimentos experimentais. As técnicas de biologia molecular foram desenvolvidas, com algumas adaptações, conforme descrito por Sambrook *et al.* (1990). Portanto, foram realizadas algumas adaptações nas metodologias e os detalhes serão apresentados no tópico Resultados e nas legendas das respectivas figuras.

3.1 Pesquisas em banco de dados e análise de bioinformática

Os cDNAs moldes para as proteínas de interesse foram inicialmente identificados através de pesquisa no banco de dados de EST (*Expressed Sequence Tag*). Para isto, foram utilizados as seqüências de aminoácidos e de nucleotídeos das proteínas previamente identificadas no banco de dados *GenBank* do NCBI (*National Center of Biotechnology Information* – www.ncbi.nlm.nih.gov). A ferramenta de bioinformática BLAST (ferramenta computacional de procura por alinhamento básico local – www.ncbi.nlm.nih.gov/BLAST/) foi utilizada para a comparação entre duas seqüências de nucleotídeos (BLASTn). Para alguns casos também foi utilizada a ferramenta BLAST na sua versão para comparação de seqüências de aminoácidos com seqüências de nucleotídeos (tBLASTn). As avaliações de bioinformática foram feitas utilizando as ferramentas disponíveis na rede mundial de computadores. A ferramenta de alinhamento múltiplo ClustalW (clustalw.genome.ad.jp/) e LALIGN (www.ch.embnet.org/software/LALIGN_form.html) foram utilizadas para avaliar a identidade entre diferentes seqüências de aminoácidos. A predição da localização intracelular das proteínas foi realizada com o auxílio do programa MITOPROT II 1.0a4 (www.mips.biochem.mpg.de/cgi-bin/proj/medgen/mitofilter). Alguns parâmetros físico-químicos e o coeficiente de absorvidade molar das proteínas recombinantes produzidas e purificadas foram calculados a partir da seqüência de aminoácidos pelo programa ProtParam Tool (ca.expasy.org/tools/protparam.html)

3.2 cDNA plasmidial

Os cDNAs de interesse foram obtidos pela aquisição comercial de clones EST obtidos e armazenados pelo sistema *IMAGE Consortium* (*Integrated Molecular Analysis of Genomes and their Expression Consortium*) pelas empresas Incyte Genomics, Life Technologies ou Invitrogen. A tabela I mostra a relação de clones adquiridos.

Tabela I: Lista de clones ESTs contendo os cDNAs das proteínas de interesse adquiridos comercialmente para os procedimentos de clonagem.

Proteína alvo	Código GenBank	Nome do clone depositado no GenBank
Hsp70	BE891351	601431889F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE: 3917086 5', mRNA sequence
	BE893670	601436469F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE: 3921521 5', mRNA sequence
	BE901434	601674773F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE: 3957403 5', mRNA sequence gi 10390612 gb BE901434.1 BE901434[10390612]
	BE791812	601583071F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE: 393772 5', mRNA sequence gi 10213010 gb BE791812.1 BE791812[10213010]
Hsp40	AW247277	2820635.5prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE: 2820635 5', mRNA sequence
	AI209061	qg28f10.x1 NCI_CGAP_Kid3 Homo sapiens cDNA clone IMAGE:17624 3' similar to gb:L08069 DNAJ PROTEIN HOMOLOG 2 (HUMAN); mRNA sequence
	AA081471	zn06a01.r1 Stratagene hNT neuron (#937233) Homo sapiens cDNA clone IMAGE:546600 5' similar to TR:G1127833 G1127833 HEAT SHOCK PROTEIN HSP40 HOMOLOG.; mRNA sequence
	AA464833	zx43h03.r1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone IMAGE: 789269 5' similar to TR:G575891 G575891 HSP40.; mRNA sequence
GrpE	BE614754	601281320F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE: 3623092 5' mRNA sequence
	BE615384	601280650F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE: 36225696 5' mRNA sequence
	AI431866	ti26b04.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE: 2131567 3' similar to TR:P97576 P97576 MT-GRPE PRECURSOR.; mRNA sequence

3.3 Reação em cadeia da DNA polimerase

A reação em cadeia da DNA polimerase (PCR) foi realizada com o auxílio dos seguintes kits comerciais: *PCR SuperMix*[®] (Life Technologies) ou *ReadyMix Taq PCR Reaction Mix*[®] (Sigma) de acordo com o protocolo dos fabricantes. A reação foi realizada em um termociclador (*Mastercycler Personal*[®] – Eppendorf). O programa modelo utilizado foi constituído por 7 etapas: 1) desenovelamento inicial do cDNA molde em uma temperatura de 95 °C por 5 minutos, seguido por um outro ciclo 2) de desenovelamento do cDNA molde na mesma temperatura, porém por 1 minuto; 3) anelamento dos oligonucleotídeos (ver abaixo) em temperaturas previamente determinadas considerando o oligonucleotídeo de menor temperatura de anelamento, como calculado pela fórmula $[2 \times (A+T) + 4 \times (G+C)]$, que leva em consideração o conteúdo de cada tipo de base nitrogenada existente no oligonucleotídeo, com o auxílio do programa Oligo Analysis & Plotting Tool (oligos.qiagen.com/oligos/toolkit.php). Entretanto, a temperatura de anelamento foi ainda reduzida em 5 °C para compensar a inserção dos sítios de restrição. 4) a etapa de polimerização, em temperatura de 72 °C, foi realizada por um tempo dependente do tamanho do produto de PCR a ser obtido (1 minuto/1.000 pb). A quinta etapa do programa foi a progressão geométrica ou amplificação do cDNA realizado pela repetição dos passos 2 – 4 por 36 vezes. A sexta etapa foi uma polimerização final em temperatura de 72 °C por 5 minutos. A última etapa foi o armazenamento da reação a 4 °C por um tempo indefinido.

3.4 Construção dos oligonucleotídeos iniciadores para as PCR

Os oligonucleotídeos iniciadores foram construídos a partir das seqüências do cDNA de interesse. Para facilitar a clonagem em vetores de expressão (veja a seguir), algumas bases foram substituídas para criar sítios para as enzimas de restrição *Nde* I ou *Nhe* I no oligonucleotídeo iniciador 5' na região do códon de iniciação do cDNA alvo, e *Bam*H I ou *Xho* I no oligonucleotídeo iniciador 3' na região do códon de terminação do cDNA alvo.

3.5 Gel de agarose e determinação da concentração de DNA

O DNA foi analisado através de eletroforese (sistema *Horizon 58*[®] – Life Technologies) em gel de agarose contendo 10 mg/mL de agarose em tampão Tris-Borato 45 mmol/L contendo EDTA (ácido etilenodiaminotetracético) 1 mmol/L (pH 8,0). Após a aplicação do DNA no gel, este foi submetido a uma corrente elétrica de 75 Volts, por aproximadamente

uma hora e trinta minutos. Em seguida o gel foi incubado em tampão descrito acima, contendo 0,5 µg/mL de brometo de etídeo, por 30 minutos, e visualizado no transluminador de UV T1202® (Sigma). Foi utilizada uma câmera digital DC120® (Kodak) acoplada a um computador contendo o programa EDAS120® (*Electrophoresis Documentation and Analysis System 120* – Kodak) para o registro das imagens dos géis. A determinação da concentração de DNA foi feita através da comparação entre a intensidade da banda de DNA marcada com brometo de etídeo e a intensidade de bandas padrão de DNA (λ DNA/HindIII, GIBCO BRL).

3.6 Vetores para clonagem

3.6.1 Vetores de transferência

O objetivo das clonagens dos produtos de PCR dos cDNAs de interesse, obtidos a partir dos oligonucleotídeos desenhados com sítios de restrição, em vetores de transferência foi facilitar as digestões com as enzimas de restrição adequadas para a posterior subclonagem nos vetores de expressão (veja a seguir). Foram utilizados os vetores pUC18 (Pharmacia) e o pGEMT (kit *pGEMTeasy*® – Promega) como sistemas de transferência.

3.6.2 Vetor de expressão

Para a expressão da proteína de interesse, os cDNAs-alvo foram subclonados nos vetores de expressão pET23a® e/ou pET28a® (Novagen). Estes vetores de expressão estão sob o controle do sistema T7 no qual o cDNA alvo é clonado sob o controle do promotor do gene 10 do bacteriófago T7 (Studier *et al.*, 1990). A expressão do cDNA alvo ocorre somente na presença da RNA polimerase do vírus T7, devido à especificidade do promotor. As cepas de bactérias *E. coli* BL21(DE3) (veja a seguir) possuem, integrado ao seu genoma, o prófago do bacteriófago T7, que é lisogênico. Neste prófago, a expressão da T7 RNA polimerase está sob o controle do promotor lacUV5 do operon lac, que é mantido reprimido na presença de glicose e na ausência de seu indutor natural, a 1,6 aldolactose, um derivado da galactose (Studier *et al.*, 1990). O produto do gene *lac I* representa o repressor que, quando ligado ao operador O (presente no prófago do T7), impede a expressão da T7 RNA polimerase, não ocorrendo a transcrição do cDNA alvo. Porém, na presença de IPTG (isopropil-β-D tiogalactopiranosídeo), que é um análogo do indutor natural (a 1,6-aldolactose), o repressor se desliga do operador liberando a expressão da T7 RNA polimerase do bacteriófago T7.

Esta, por sua vez, atua na transcrição do cDNA alvo clonado no vetor do sistema pET. Como a T7 RNA polimerase é cinco vezes mais eficiente que a RNA polimerase da *E. coli* e é produzida em grandes quantidades, em pouco tempo, praticamente, o único produto protéico sintetizado pela bactéria será aquele representado pelo cDNA alvo. Assim, na presença da RNA polimerase do bacteriófago T7, o cDNA alvo é altamente transcrito, resultando em grande quantidade do produto protéico. Outra vantagem desse promotor é a sua especificidade, ou seja, a RNA polimerase da *E. coli* não transcreve o cDNA alvo, o que permite controlar o momento exato em que ele será expresso (Studier *et al.*, 1990).

O pET28a possui uma sequência de DNA localizada em posição anterior ao sítio múltiplo de clonagem (SMC) do plasmídeo que codifica um peptídeo de 20 resíduos de aminoácidos (MGSSHHHHHSSGLVPRGSH), contendo uma sequência de 6 resíduos de histidinas seguidas (Novagen). Esta sequência recebe a denominação de cauda de poli-histidina e a sua função é facilitar o processo de purificação do produto protéico, codificado pelo cDNA adequadamente clonado em sua extremidade 3', através de cromatografias de afinidade (ver item 3.15.2 descrito abaixo). Para isto, é utilizada uma resina contendo um metal acoplado (geralmente níquel ou cobalto) aos quais os anéis imidazólico de histidinas vicinais são capazes de se ligarem e formarem complexos. A sequência de aminoácidos descrita acima também possui um sítio de hidrólise para a enzima proteolítica trombina que é capaz de clivar a ligação peptídica entre os aminoácidos Arg17 e Gly18 (Novagen). Assim, as proteínas codificadas pelo cDNA, que foi clonado no SMC do pET28a em fase de leitura com o cDNA da cauda de poli-histidina, serão expressas em fusão a uma sequência de 20 resíduos de aminoácidos acoplados na extremidade amino-terminal. Se necessário, a maior parte desta sequência de aminoácidos pode ser eliminada da proteína pela enzima proteolítica trombina, a fim de se obter uma proteína sem a cauda de poli-histidina.

3.7 Linhagens de *Escherichia coli*

A linhagem de bactérias *E. coli* DH5 α foi utilizada para multiplicação e manutenção de vetores plasmidiais.

A cepa de bactérias *E. coli* BL21(DE3)[®] (Stratagene) foi utilizada para expressar as proteínas de interesse. Esta cepa possui baixos níveis de produção de proteases, que poderiam degradar a proteína alvo, e possui ainda o gene DE3 responsável pela síntese da enzima T7 RNA polimerase controlado pelo promotor lacUV5 e ativado pela adição de IPTG.

A linhagem de bactérias *E. coli* BL21(DE3)pLysS[®] (Stratagene), além das características acima descritas, apresenta expressão aumentada de lisozima que inibe a T7 RNA polimerase, regulando a expressão da proteína de interesse clonada no vetor do sistema pET. Esta linhagem pode ser utilizada quando a proteína alvo se apresenta tóxica para a bactéria.

A cepa de bactérias *E. coli* BL21(DE3) Códon Plus RP[®] (Stratagene), também apresenta as características descritas acima para a cepa *E. coli* BL21(DE3), porém possui suplementação de RNA transportadores para códons raros de arginina e prolina. Estes códons raros são aqueles pouco freqüentes nas bactérias, sendo que estas não possuem RNA transportadores que os reconhecem.

3.8 Meios de cultura

O meio de cultura utilizado para o crescimento de bactérias, expressão de proteínas e para a repicagem das cepas foi o meio LB (Luria-Bertani), cuja composição é: triptona 10 g/L, extrato de levedura 5 g/L, NaCl 10 g/L, e pH 7,0 (Sambrook *et al.*, 1990). O meio foi suplementado de um antibiótico (carbenecilina 50 µg/mL, canamicina 30 µg/mL ou cloranfenicol 50 µg/mL) a fim de selecionar as bactérias que continham DNA plasmidial de interesse.

3.9 Clonagem

3.9.1 Clonagem em vetores de transferência

O protocolo para a clonagem dos cDNAs de interesse no vetor pUC18 está descrito abaixo e as instruções do fabricante (Promega) foram seguidas para as clonagens utilizando o kit *pGEMT easy*.

Preparação do vetor pUC18

O pUC18 (Pharmacia) foi submetido à digestão exaustiva com a enzima de restrição *Sma* I (New England BioLabs) conforme protocolo do fabricante. Esta enzima foi escolhida, pois sua atividade é em corte cego (*'Blunting'*) e existir um sítio de restrição para ela na região do SMC do vetor pUC18. Após a digestão, o plasmídeo foi incubado com a enzima fosfatase alcalina ácida intestinal (New England BioLabs) conforme o protocolo do

manufaturador. Este processo foi realizado para evitar a ligação do pUC18 sem a presença do inserto de interesse. A eficiência da ação de ambas as enzimas foi monitorada pela transformação de cepas de bactérias *E. coli DH5α* tanto com ou sem reação de ligação (ver abaixo).

Preparação do inserto

O produto de PCR foi purificado do gel de agarose 1% após eletroforese (Sambrook *et al.*, 1990) e foi submetido à reação de reparo e fosforilação das extremidades utilizando as enzimas *Klenow* e *T4 Polinucleotídeo Kinase* (New England Biolabs), respectivamente, conforme instruções do manufaturador. A reação de reparo foi necessária para se obter o fragmento de DNA sem a presença de extremidades coesivas e aptas a se ligarem no vetor preparado como descrito no tópico anterior. A fosforilação das extremidades do inserto foi necessária para catalisar a reação de ligação, visto que o vetor foi desfosforilado para evitar sua própria ligação (ver tópico anterior).

Reação de ligase

Após inativação das enzimas *Klenow* e *T4 PNK*, o produto de PCR foi ligado a 50 ng de pUC18 previamente tratado com as enzimas *Sma* I e fosfatase alcalina. A reação de ligação foi realizada no mesmo tampão da reação anterior utilizando 3 unidades de *T4 DNA Ligase* (GIBCO BRL) de acordo com o protocolo do manufaturador por 16 horas à 16 °C. Os produtos de ligação foram utilizados para transformar células como descrito a seguir no tópico 3.10 de Material e métodos.

Análise de clones

Como os cDNAs de interesse foram clonados no SMC do vetor pUC18 e este se localiza dentro do cDNA que codifica da enzima β -galactosidase que degrada um açúcar sintético 5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside (X-gal – Sigma) levando à formação de um produto azul (Sambrook *et al.*, 1990). Os clones que apresentaram ausência de coloração azul na presença de X-gal e IPTG (Sigma) foram submetidos à análise por restrição com as enzimas utilizadas para confecção dos oligonucleotídeos.

3.9.2 Subclonagem em vetores de expressão

Tanto o vetor de expressão (pET28a ou pET23a) quanto os vetores de transferência (pUC18 ou pGEMT) que apresentaram o cDNA alvo clonado foram submetidos à digestão exaustiva com as respectivas enzimas conforme o protocolo do fabricante. Após eletroforese em agarose 1% (Sambrook *et al.*, 1990), as bandas correspondentes ao fragmento de interesse e o vetor de expressão, digeridos com as enzimas adequadas, foram extraídas do gel e purificadas usando o *kit QIAEX II* (Qiagen) conforme o protocolo do fabricante. Aproximadamente 200 ng do inserto de interesse foram incubados com aproximadamente 50 ng de vetor de expressão digerido juntamente com 3 unidades da enzima *T4 DNA Ligase* (GIBCO BRL), de acordo com o protocolo do fabricante, por 16 horas à 16 °C.

3.10 Transformação de bactérias competentes e seleção dos clones transformantes

A transformação de bactérias pelos vetores de interesse foi realizada através do método de eletroporação, utilizando o eletroporador *MicroPulser™ Electroporation Apparatus* (Bio-Rad) conforme protocolo do fabricante. As células eletrocompetentes foram obtidas como descrito no protocolo do fabricante. Quando necessário, o DNA plasmidial foi previamente precipitado com etanol absoluto e acetato de sódio 3 mol/L (pH 5,2) para evitar a interferência do sal e aumentar a eficiência da eletro-transformação.

O produto de transformação foi espalhado em meio LB sólido contendo o antibiótico adequado (carbenecilina 50 µg/mL ou canamicina 30 µg/mL) e as placas submetidas a crescimento em temperatura constante de 37 °C por 16 horas. A seleção dos clones transformantes foi feita através da análise de restrição com as enzimas utilizadas para a clonagem no vetor de expressão.

3.11 Extração de DNA plasmidial

A obtenção do DNA plasmidial de interesse foi realizada através do método de lise alcalina em pequena escala como descrito em Sambrook *et al.* (1989). Para a obtenção do DNA plasmidial com um grau de pureza maior, principalmente para as PCR e para o seqüenciamento de DNA, foi utilizado o mesmo método de lise alcalina, porém com auxílio de kit de extração de DNA plasmidial, *QIAprep Spin Miniprep Kit* (Qiagen), de acordo com o protocolo do fabricante.

3.12 Seqüenciamento de DNA

O seqüenciamento de DNA foi feito em seqüenciador automático 377 *ABI Prism* (Perkin Elmer). A PCR, utilizando o método de dideoxi nucleotídeos, foi realizada com o kit *Big Dye Terminator* (Perkin Elmer) de acordo com o protocolo do fabricante e foi conduzida em termociclador *Mastercycler Personal* (Eppendorf). Os oligonucleotídeos universais que interagem especificamente com os respectivos vetores e que foram utilizados para o seqüenciamento dos cDNA de interesse clonados nos SMC foram: *T7 promoter* (5' TAATACGACTCACTATAGGG 3'), *T7 terminator* (5' TATGCTAGTTATTGCTCAG 3'), *M13 Forward* (5'CACGACGTTGTAAAACGAC 3'), *M13 Reverse* (5' GGATAACAATTCACACAGG 3') e *SP6* (5' GATTTAGGTGACACTATAG 3').

3.13 Expressão de proteínas

O sistema de expressão utilizado para a produção das proteínas recombinantes foi o sistema pET descrito por Studier *et al.* (1990). Tal sistema está descrito em detalhes no item 3.6.2 de Materiais e métodos. Bactérias *E. coli* das linhagens *BL21(DE3)*[®], *BL21(DE3) pLysS*[®] ou *BL21(DE3) codon plus RP*[®] (Stratagene), previamente eletro-transformadas pelo DNA plasmidial de interesse como descrito acima (item 3.10 de Material e métodos), foram crescidas em meio de cultura LB, contendo 30 µg/mL de canamicina ou 50 µg/mL de carbenecilina e 50 µg/mL de cloranfenicol (para as cepas *BL21(DE3) pLysS*[®] ou *BL21(DE3) Codon plus RP*[®]), sob agitação constante de 200 ciclos/min à 37 °C. Este crescimento foi monitorado até que a densidade ótica a 600 nm atingisse 0,8 unidades de absorbância. Nesta fase de crescimento, as culturas celulares foram induzidas à expressão pela adição de IPTG na concentração final de 0,4 mmol/L. Em alguns casos, a temperatura foi reduzida para obter a proteína na fração solúvel do lisado (ver abaixo). O período de indução ótimo foi determinado para cada proteína recombinante em ensaios de indução em pequena escala, onde 50 mL de meio contendo as células foram induzidas nas mesmas condições e alíquotas retiradas de hora em hora. O nível de indução foi avaliado por eletroforese em gel de poliacrilamida em presença de duodecil-sulfato de sódio (SDS-PAGE – item 3.15 de Material e métodos). As culturas de células, após a indução, foram centrifugadas à 15.000 x *g* por 5 min e à temperatura de 4 °C, e o sedimento bacteriano armazenado à – 80°C para posterior lise e purificação das proteínas de interesse.

3.14 Purificação das proteínas recombinantes

3.14.1 Lise de bactérias

O sedimento bacteriano contendo as proteínas recombinantes foram lisadas com o auxílio do aparelho sonicador *Sonifier 450* (Branson) em tampão de lise contendo Tris-HCl 50 mmol/L (pH 8,0), KCl 100 mmol/L ou 500 mmol/L e EDTA 1 mmol/L (15 mL/L de cultura) após 30 minutos de incubação em gelo com 5 unidades de *DNase* (GIBCO BRL) e 30 µg/mL de lisozima (Sigma) e centrifugadas a 24.100 x g, por 30 minutos, em temperatura de 4 °C. O sobrenadante foi filtrado em filtro *Millipore* 0,2 µm e utilizado para a purificação da proteína de interesse.

3.14.2 Cromatografia de afinidade

Para a purificação das proteínas recombinantes fusionadas à cauda de poli-histidina fornecida pelo vetor de expressão pET28a (para detalhes veja item 3.6.2 de Material e métodos), foi utilizada a coluna de afinidade *HiTrap Chelating Column 1 x 5 mL* (Amershan Pharmacia Biotech) previamente carregada com níquel de acordo com o protocolo do fabricante. O processo de purificação foi feito com a coluna de afinidade acoplada ao sistema de cromatografia líquida de proteínas por pressão (FPLC) *ÄKTA FPLC* (Amershan Pharmacia Biotech). O protocolo básico foi realizado com uma etapa de equilíbrio da coluna, que foi lavada com 15 mL de tampão fosfato de sódio 20 mmol/L (pH 7,4), NaCl 500 mmol/L e imidazol 20 mmol/L, seguida pela injeção da amostra e eluição da fração não ligante com 20 mL do mesmo tampão. Após estes passos, a amostra foi eluída com 15 mL de tampão fosfato de sódio 20 mmol/L (pH 7,4), NaCl 500 mmol/L, imidazol 75 mmol/L para eluição de contaminantes porventura ligados à resina. Então, as proteínas de interesse foram eluídas por 500 mmol/L de imidazol no tampão descrito acima. A última etapa foi a lavagem da coluna com 10 mL de tampão fosfato de sódio 20 mmol/L (pH 7,4), NaCl 500 mmol/L e imidazol 20 mmol/L. Todo o processo foi realizado com monitoramento constante da absorbância a 280 nm, da temperatura, da pressão do sistema FPLC e da condutância.

3.14.3 Cromatografia de troca iônica

Técnicas de cromatografia de troca iônica foram utilizadas para a primeira etapa da purificação da GrpE#1 clonada em vetor de expressão na ausência da cauda de poli-histidina

(pET23a). O ponto isoelétrico teórico para a GrpE#1 foi estimado pelo programa ProtParam Toll (ca.expasy.org/tools/protparam.html) como sendo 6,0 e foi utilizado para escolher a resina de troca aniônica *Q Sepharose* (Amershan Pharmacia Biotech) e as condições de pH utilizadas. Foram realizados ensaios piloto em colunas *HiTrap Q Sepharose FF 1 mL* (Amershan Pharmacia Biotech) acopladas ao sistema *ÅKTA FPLC* (Amershan Pharmacia Biotech) utilizando como solução tampão Tris-HCl 25 mmol/L (pH 8,0) sem NaCl e contendo 1 mol/L NaCl. Após a estimativa, por técnicas de gradiente salino, da concentração de NaCl suficiente para a eluição da proteína, o processo foi refinado para a purificação em passos de eluição isocrática por NaCl. Uma coluna *HiLoad 16/20* (Amershan Pharmacia Biotech) foi empacotada com aproximadamente 15 mL resina *Q Sepharose* (Amershan Pharmacia Biotech) e foi utilizada para a purificação em alta escala.

3.14.3 Cromatografia de exclusão molecular

A etapa final da purificação das proteínas recombinantes foi feita pela técnica de cromatografia de exclusão molecular (CEM) utilizando a coluna *HiLoad Superdex 200 pg 26/60* acoplada ao sistema *ÅKTA FPLC* (Amershan Pharmacia Biotech). Esta metodologia também foi utilizada para realizar alguns experimentos analíticos com o auxílio das colunas *Superdex 200 HR 16/30* e *Superose 12 HR 10/30* (Amershan Pharmacia Biotech) acopladas ao sistema descrito acima. O protocolo experimental foi aquele sugerido pelo fabricante para cada uma das colunas. De maneira geral, a uma coluna, previamente lavada e equilibrada com o tampão de interesse, foi aplicado um volume de amostra correspondente a até 5% de seu volume de coluna e realizada a corrida com um volume de tampão correspondente a 1,5 vezes o volume da coluna. O tampão utilizado nas CEM foi Tris-HCl 25 mmol/L (pH 7,5) contendo 150 ou 500 mmol/L de NaCl (conforme especificado no tópico Resultados e na legenda das figuras) e agente redutor (DTT ou β -mercaptoetanol). Todo o processo foi realizado com monitoramento constante da absorbância a 280 nm, da temperatura, da pressão do sistema FPLC e da condutância.

3.14.4 Processo de concentração de proteínas

Para obter as proteínas em alta concentração, foram utilizados o sistema *Amicon Ultra 10.000 MWCo* ou a técnica de osmose reversa, utilizando poli-etileno glicol 8.000 (Sigma), seguidas por diálise exaustiva contra o tampão de interesse.

3.15 SDS-PAGE

As amostras coletadas em cada etapa dos processos de expressão e de purificação foram diluídas em tampão de amostra para eletroforese (Tris-HCl 50 mmol/L, pH 6,8, DTT 100 mmol/L, SDS 2%, azul de bromofenol 0,1% e glicerol 10%), fervidas durante 5 min, no caso de proteínas em culturas de células, e centrifugadas por 2 min em uma microcentrífuga *Eppendorff* em temperatura ambiente. Estas amostras foram submetidas, juntamente com um padrão de *MM*, à eletroforese em SDS-PAGE (Laemmli, 1970). A eletroforese foi conduzida com a utilização de um aparato *Mini-Protean II Dual Slab Cell* (Bio-Rad), com um gel de empacotamento de 5%, gel de separação de 10% ou 12% de concentração peso/volume de acrilamida/bis-acrilamida, para proteínas de *MM* entre 40 a 70 kDa e 20 a 50 kDa, respectivamente. A corrida foi realizada em voltagem constante de 175 Volts por aproximadamente 45 min. O gel foi corado por 20 min em etanol:ácido acético:água 5:1:15 (v:v:v) e *Comassie Brilliant Blue R* (Bio-Rad) 0,25% e descorado em ácido acético:etanol:água 3:2:35 (v:v:v).

3.16 Determinação da concentração

3.16.1 Determinação da concentração de proteínas

A concentração de proteínas foi determinada por técnicas espectroscópicas, através do método descrito por Edelhock (1967), utilizando os seguintes comprimentos de onda: 282, 280, 279, 278 e 276 nm. Para isto, os coeficientes de absorvidade molar (ϵ), nestes comprimentos de onda, foram calculados utilizando a seguinte equação:

$$\epsilon(\lambda) = n_{Trp} \times \epsilon_{Trp} + n_{Tyr} \times \epsilon_{Tyr} + n_{Cys} \times \epsilon_{Cys} \quad [\text{Equação 15}]$$

onde, $\epsilon(\lambda)$ é o coeficiente de absorvidade molar ($\text{mol/L}^{-1} \cdot \text{cm}^{-1}$) da proteína de interesse em um determinado λ ; n_{Trp} - número de triptofanos da proteína; n_{Tyr} - número de tirosinas e n_{Cys} é o número de cisteínas. ϵ_{Trp} , ϵ_{Tyr} e ϵ_{Cys} são os coeficientes de absorvidade molar do triptofano, da tirosina e da cisteína, respectivamente. Os valores dos ϵ destes aminoácidos variam de acordo com o λ e foram calculados pelo programa ProtParam Tool (ca.expasy.org/tools/protparam.html) de acordo com Gill & von Hippel (1989). Os parâmetros calculados e utilizados para quantificação das proteínas são apresentados conjuntamente no Anexo III.

As amostras das proteínas purificadas foram diluídas em tampão contendo fosfato de sódio (pH 6,5) na concentração final de 20 mmol/L e Gdn-HCl 6 mol/L, o espectro de absorvância coletado entre os λ de 250-300 nm após 15 minutos de incubação na bancada. Após a subtração do respectivo branco, onde o tampão no qual a proteína se encontrava foi diluído no tampão desnaturante descrito acima, as leituras nos λ descritos acima foram utilizadas para estimar a concentração de proteínas pela equação de Beer-Lambert:

$$A = \epsilon \times l \times C \quad \text{[Equação 16]}$$

onde A é a leitura de absorvância; l , comprimento do caminho ótico em centímetros e C , concentração em mol/L. A concentração molar determinada correspondeu à média das concentrações encontradas para cada λ entre 276 e 282 nm e multiplicado pelo fator de diluição. As leituras de absorvância, para cada λ relacionado, foram realizadas em triplicada.

3.16.2 Determinação da concentração de nucleotídeos

A concentração de nucleotídeos adenosina foi também determinada por técnicas espectrométricas utilizando a equação de Beer-Lambert descrita acima (equação 16). Os nucleotídeos ADP e ATP foram obtidos da Sigma como sais de sódio. Os nucleotídeos foram diluídos para uma leitura ótima entre 0,5 e 1,0 em λ de 259 nm. Foi utilizado o ϵ de 15.400 mol/L⁻¹.cm⁻¹ para os nucleotídeos adenosina, obedecendo a correlação absorvância 280/260 nm de 0,16 (Sambrook *et al.*, 1990), após subtração do respectivo branco. A equação de Beer-Lambert foi resolvida e corrigida pelo fator de diluição.

3.17 Espectropolarimetria de dicroísmo circular

Os experimentos de CD foram realizados em um espectropolarímetro *J-810* (JASCO), acoplado a um sistema interno de controle de temperatura (*Peltier Type Control System PFD 425S* – JASCO) e controlador externo de temperatura (*NESLAB RTE Series* – NESLAB). As cubetas de quartzo usadas possuíam caminho ótico de 1 mm ou 0,1 mm e as proteínas foram diluídas suficientemente para fornecer um sinal em 222 nm entre – 20 e – 40 mDeg. Os programas utilizados para registro e tratamento dos dados foram o *Spectra Manager*® (JASCO) e o *Origin*® (Microcal), respectivamente. Os espectros de CD foram normalizados em elipticidade molar residual (*EMR* ou $[\Theta]$) pela aplicação da seguinte equação:

$$EMR = \frac{\Theta \times MM}{10 \times C \times l \times n} \quad [\text{Equação 17}]$$

onde Θ é a elipticidade (em mDeg), MM é a massa molecular (em g/mol), C é a concentração da proteína (mg/mL), l é o comprimento do caminho ótico (cm) e n o número de resíduos de aminoácidos da proteína. Os dados de EMR foram representados em função do λ (nm).

3.17.1 Espectros de dicroísmo circular

Os espectros de CD foram coletados em temperatura constante de 20 °C. Os espectros de CD em ultravioleta distante iniciavam em 260 nm e cobriam até a banda na qual a voltagem do detector atingisse o limite de 600 Volts. Bandas com voltagens superiores, porventura coletadas, ocasionadas pela presença de sal ou excesso de proteínas foram descartados dos espectros analisados. Espectros cumulativos de 32 coletas (velocidade de coleta = 100 nm/min e tempo de resposta do detector de 2 segundos) foram analisados como descrito acima, após subtração do espectro do respectivo tampão (acumulados por 16 vezes). A porcentagem de estrutura secundária do tipo α -hélice, folha β -pregueada e estrutura randômica foi calculado com o auxílio do programa CDNN Deconvolution (Versão 2 – Bioinformatik.biochemtech.uni-halle.dee/cdnn). Este programa compara o espectro de CD normalizado em EMR com um banco de dados de 13, 23 ou 33 espectros em EMR de proteínas cuja estrutura em alta resolução é determinada (Böhm *et al.*, 1992).

3.17.2 Experimentos de desnaturação térmica monitorados por CD

Os ensaios de desnaturação térmica foram realizados em cubetas de quartzo de 1 mm com incremento da temperatura controlado pelo sistema interno de controle de temperatura descrito acima. As velocidades de incremento de temperatura e as concentrações das proteínas utilizadas estão descritas na legenda das respectivas figuras. As curvas de desnaturação térmica ($n = 3$) foram obtidas nas mesmas condições e a média das curvas normalizadas em função da EMR como descrito acima. A T_m foi estimada através da análise do máximo do ajuste de curvas gaussianas da primeira derivativa das curvas de desnaturação após “smoothing” Savitzky-Golay realizado pelo programa Origin® (Microcal).

3.17.3 Experimentos de desnaturação química monitorados por CD

Ensaio de desnaturação química foram realizados usando como desnaturantes uréia e Gdn-HCl, sendo ambos preparados no mesmo dia do ensaio, no respectivo tampão da proteína em questão e com a concentração do agente desnaturante determinada utilizando refratometria (método de face). A solução de proteína, na presença do agente desnaturante, foi preparada por diluição de uma mesma quantidade de proteína em um mesmo volume de solução do agente químico desnaturante em diferentes concentrações. A temperatura da cela foi mantida constante à 20 °C e os espectros (cumulativos de 4 coletas) foram adquiridos na região de 260-205 nm após 45 minutos de incubação na bancada. As leituras de CD das proteínas a 222 nm, após subtração do respectivo tampão, foram utilizadas para a construção dos gráficos de *EMR* em função da concentração de agente desnaturante.

3.18 Espectroscopia de emissão de fluorescência

Os experimentos de fluorescência foram realizados utilizando um fluorímetro *Aminco Bowman Series 2[®]* (SLM-AMINCO), cubetas de quartzo de 1 cm x 1 cm de caminho ótico e proteínas em concentrações de 2 µmol/L. Os programas utilizados para registro e tratamento dos dados foram o SLMAB2[®] (SLM-AMINCO) e o Origin[®] (Microcal), respectivamente. A temperatura da câmara de leitura foi mantida constante à 20 °C através de um controlador de temperatura digital *Polyscience*. Os espectros de emissão de fluorescência foram obtidos, após excitação dos resíduos de triptofano a λ 295 nm, com a abertura para passagem de luz de 2 nm e de emissão máxima a 330 nm (abertura 4 nm). O equipamento foi inicialmente configurado para uma mesma concentração de proteína. O espectro de emissão de fluorescência do tampão (3 leituras cumulativas) foi subtraído do espectro de emissão de fluorescência da proteína (3 leituras cumulativas). Os pontos para a composição do espectro de emissão de fluorescência foram coletados a cada 1 nm dentro do intervalo espectral de 305 a 420 nm. Os dados foram analisados pelo perfil do espectro de emissão obtido em função da intensidade de fluorescência pela aplicação da integral da área espectral na faixa de 320 a 380 nm (calculado pelo programa Origin[®]) e pelo cálculo do centro de massa espectral ($\langle \lambda \rangle$) considerando a seguinte equação: $\langle \lambda \rangle = \Sigma(\lambda \times IF_{\lambda}) / \Sigma IF_{\lambda}$. O espectro de emissão de fluorescência da Hsp70-1A também foi analisado pela interpretação por duas curvas Gaussianas (Origin[®]) para estimar o λ máximo de emissão de fluorescência para cada resíduo de triptofano.

3.19 Experimentos de calorimetria

3.19.1 Calorimetria isotérmica de titulação

Os experimentos de ITC para a determinação da afinidade da Hsp70-1A pelos nucleotídeos adenosina foram realizados a 20 °C utilizando um aparelho *VP-ITC*[®] (Microcal). A solução de Hsp70-1A foi exaustivamente dialisada contra tampão TKP (Tris-HCl 25 mmol/L, pH 7,5, contendo NaCl 50 mmol/L, KCl 5 mmol/L, Na₃PO₄ 5 mmol/L e 1 mmol/L β-mercaptoetanol) e aproximadamente 15 μmol/L de proteína (concentração de proteína quantificada na própria amostra aplicada como descrito no item 3.16.1 de Materiais e métodos) foi adicionada à célula de 1,4 mL do microcalorímetro. As soluções de MgADP ou MgATP utilizadas foram preparadas no tampão TKP e apresentavam concentrações entre 300 a 350 μmol/L (a concentração de nucleotídeos foi determinada a partir de uma alíquota da própria amostra utilizada na titulação como descrito no item 3.16.2 de Materiais e métodos). Alíquotas de 5 μL dos nucleotídeos foram injetadas em intervalos regulares de 6 minutos após a primeira injeção de 1 μL, cujo objetivo foi eliminar gases na ponta da seringa. O calor da reação (Q) por injeção (μcal/seg) foi determinado pela integração das áreas dos picos usando o programa Origin[®] (Microcal) e é proporcional ao volume da célula do microcalorímetro (V_{cel}), à variação da concentração do ligante ($[L]$) e à ΔH . O calor de diluição dos nucleotídeos foi determinado em experimentos no quais os nucleotídeos foram titulados no tampão TKP para verificar a existência de alguma dependência no perfil de diluição dos nucleotídeos. O calor de diluição da proteína foi determinado da linha de base do final da titulação e depois comparada com o calor de diluição do nucleotídeo para a subtração do calor de interação observado. O calor envolvido na interação por mol de nucleotídeo injetado foi analisado contra a razão molar entre [nucleotídeo]/[proteína]. A ΔH , a estequiometria da ligação (n) e a constante de associação (K_A) foram estimados pelo ajuste da curva com o auxílio do programa descrito acima, que interpreta a seguinte equação quadrática:

$$Q = \frac{(n[M]\Delta HV_{cel})}{2} \left\{ 1 + \frac{[L]}{(n[M])} + \frac{1}{(nK_A[M])} - \sqrt{\left[\left(1 + \frac{[L]}{(n[M])} + \frac{1}{(nK_A[M])} \right)^2 - \frac{4[L]}{(n[M])} \right]} \right\}$$

[Equação 18]

O processo de ajuste experimental dos dados envolveu: 1) suposição inicial do n , K_A (em mol/L⁻¹) e ΔH (em cal/mol), que foi feita automaticamente pelo programa Origin® (Microcal); 2) cálculo do $\Delta Q(i)$ para cada injeção e comparação deste valores com o calor medido para a injeção experimental correspondente; 3) melhora dos valores iniciais de n , K_A e ΔH ; e 4) interação dos procedimentos acima até que melhoras significantes não ocorressem. Assim, os parâmetros n , K_A e ΔH foram otimizados para ajustar a curva de titulação pelo programa Origin® (Microcal) utilizando um processo de regressão não-linear (revisado por: Pierce *et al.*, 1999). O valor de ΔS (em cal/mol/°C) foi calculado a partir da dos dados de K_A e ΔH , pela aplicação da equação 1 (item 1.3.3 de Introdução).

3.19.2 Calorimetria de varredura diferencial

Esta metodologia está descrita no manuscrito apresentado no Anexo IV (Borges, J.C., Fischer, H., Craievich, A.F., Hansen, L.D., & Ramos, C.H.I. (2003) Free human mitochondrial GrpE is a symmetric dimer in solution. *J. Biol. Chem.*, 278, 35337-35344) e as informações adicionais para os experimentos estão descritas na legenda da respectiva figura.

3.20 Ultracentrifugação analítica

As análises utilizando a técnica de UAL, velocidade de sedimentação e sedimentação em equilíbrio, foram realizadas em um equipamento *Beckman Optima XL-A analytical ultracentrifuge* (Beckman Coulter), em temperatura constante de 20 °C, utilizando o rotor AN-60Ti (Beckman Coulter) e a varredura da absorbância em função do raio foi realizada em comprimentos de onda de 230 a 238 nm dependendo da concentração de proteína.

3.20.1 Velocidade de sedimentação

Os experimentos de velocidade de sedimentação foram realizados em velocidades apropriadas para a sedimentação de cada proteína, sendo descritas no item Resultados e nas legendas das respectivas figuras. As análises de dados envolvendo a interpretação da curva de absorbância *versus* o raio da célula foram feitas utilizando um processo de regressão não-linear através do pacote do programa Origin® (Microcal) fornecido pelo fabricante do equipamento. Os métodos de análise de dados *van Holde-Weisheit* (Van Holde & Weisheit, 1978) (gráfico do coeficiente de sedimentação), *Second Moment* (Goldberg, 1953) e *Sedimentation Time Derivative* (distribuição da integral de $g(s^*)$) (Stafford,

1994) foram utilizadas para analisar os dados de velocidade de sedimentação. Os métodos utilizados permitem o cálculo do coeficiente de sedimentação aparente (s^*), o coeficiente de difusão (D) e a MM a partir da equação 10.

O cálculo do valor do s é importante para caracterizar mudanças no tamanho e na forma das macromoléculas causadas por mudanças nas condições do meio. Porém, o valor obtido pelos programas, o s^* , não leva em consideração as contribuições do solvente e da temperatura nos cálculos executados pelos programas descritos acima. Assim, o valor de s^* calculado deve ser sempre extrapolado das condições observadas para as condições padrões (água e temperatura de 20 °C), ou $s_{20,w}$ (Laue, 2001). O valor de $s_{20,w}$ pode ser calculado pela seguinte relação:

$$s_{20,w} = s^* \left[\frac{(1 - V_{bar} \rho_{20,w})}{(1 - V_{bar} \rho_{ex})} \right] \left(\frac{\eta_{tex}}{\eta_{t20}} \right) \left(\frac{\eta_s}{\eta_w} \right) \quad [\text{Equação 19}]$$

onde, $\rho_{20,w}$ é a densidade da água à 20 °C, ρ_{ex} é a densidade do solvente nas condições experimentais, η_{tex} é a viscosidade do solvente na temperatura experimental, η_{t20} é a viscosidade do solvente à 20 °C e η_w é a viscosidade da água. Tal extrapolação elimina contribuições do solvente e temperatura no valor de s^* que não foram considerados pelos cálculos iniciais determinados nos métodos de análise, como discutido acima. O programa Sednterp (www.jphilo.mailway.com/download.htm) foi utilizado para realizar a operação descrita pela equação 19 e também para estimar os valores de: 1) V_{bar} (em mL/g) a 20 °C, a partir da seqüência primária de aminoácidos da proteína em questão; 2) densidade ρ (em g/mL) e 3) viscosidade (η) do tampão na temperatura de 20 °C.

O valor de $s_{20,w}$ pode ser medido em várias concentrações e a extrapolação da curva para o valor zero permite a obtenção de $s_{20,w}^0$ que representa uma propriedade única da macromolécula sem os efeitos dependentes da concentração da proteína na solução (Laue, 2001). Assim, mudanças observadas no valor de $s_{20,w}^0$ ocasionadas por diferentes condições experimentais (temperatura, pH, adição de ligantes, etc), podem ocorrer somente em decorrência de variações na massa ou de mudanças conformacionais da macromolécula (Laue, 2001).

3.20.2 Sedimentação em equilíbrio

Os experimentos de sedimentação em equilíbrio também foram realizados em pelo menos 3 velocidades diferentes. As amostras foram aceleradas à uma velocidade inicial constante (veja no item Resultados e nas legendas das respectivas figuras) por aproximadamente 16 horas, e através de leituras sucessivas, foi verificado se o estado de equilíbrio foi alcançado, para então realizar a coleta dos dados com 25 cumulativas por amostra. O método de *Self-Association* foi utilizado para analisar os dados de sedimentação em equilíbrio utilizando diferentes modelos de associação para cada proteína. A distribuição de proteína pela célula foi interpretada pela seguinte equação (Johnson, 1981):

$$C = C_0 e^{\left[\frac{MM (1 - V_{bar} \rho) \omega^2 (r^2 - r_0)}{2RT} \right]} \text{ [Equação 20]}$$

onde C é a concentração de proteína em uma determinada posição radial r ; C_0 é a concentração de proteína na posição radial r_0 ; ω é a velocidade angular do rotor e e representa uma função exponencial. As análises foram feitas com um conjunto mínimo de 9 curvas oriundas de 3 velocidades diferentes em uma mesma concentração de proteína e o resultado final dado pela média dos resultados nas 3 concentrações utilizadas. O programa Sednterp foi utilizado para estimar os valores de V_{bar} (em mL/g) a 20 °C a partir da seqüência primária de aminoácidos da proteína em questão e também a densidade ρ (em g/mL) para a realização dos cálculos.

3.21 Espalhamento dinâmico de luz

Os dados de D foram estimados por experimentos de espalhamento dinâmico de luz (EDL) que foram realizados em um equipamento *DynaPro-MS/X* (Protein Solutions) em temperatura constante de 20 °C. As amostras foram previamente centrifugadas a 14.000 x g por 20 minutos a 20 °C. A função de autocorrelação foram medidas em um ângulo de espalhamento de 90 graus, com 600 acumulações de 10 segundos cada e analisados pelo programa fornecido pelo fabricante. O valor de D está relacionado com o raio de Stokes (R_s) pela equação de Stokes-Einstein:

$$D = \frac{kT}{(6\pi\eta R_s)} \quad [\text{Equação 21}]$$

onde k é a constante de Boltzmann, T a temperatura absoluta e η a viscosidade do solvente.

O coeficiente de difusão, assim como o coeficiente de sedimentação, sofre efeitos dependentes da viscosidade e concentração (Ralston, 1993; Wilson, 2003). Assim, o valor de D observado (D_{obs}) foi normalizado para condições padrões de água e temperatura de 20 °C, com o auxílio da equação descrita abaixo (Ralston, 1993):

$$D_{20,w} = D_{obs} \left(\frac{293,2}{T} \right) \left(\frac{\eta_{T,w}}{\eta_{20,w}} \right) \left(\frac{\eta_s}{\eta_w} \right) \quad [\text{Equação 22}]$$

A partir do valor $D_{20,w}$ determinado para diferentes concentrações de proteína foi realizada a extrapolação para concentração de proteína igual a zero e assim obter o $D_{20,w}^0$ (Ralston, 1993) de forma similar ao realizado para o $s_{20,w}^0$ (ver tópico 3.20.1 de Material e métodos). A técnica de EDL também foi utilizada como forma de controle de qualidade para avaliar a monodispersidade das amostras para os demais ensaios de caracterização hidrodinâmica, espectroscópicas e calorimétricas.

3.22 Espalhamento de raios X a baixos ângulos

3.22.1 Coleta dos dados

As medidas de SAXS foram realizadas na linha de luz síncrotron SAS no Laboratório Nacional de Luz Síncrotron (LNLS), Campinas, São Paulo, Brasil. A aquisição de dados foi feita utilizando um detector multicanal unidimensional sensível à posição. Assim, foi possível adquirir uma curva de intensidade de espalhamento em função do número do canal do detector. O λ do raio incidente utilizado foi de 1,488 Å e a distância entre o porta-amostra e o detector foi escolhida para adquirir o espalhamento entre $0,01 < q < 0,4 \text{ Å}^{-1}$. Assim, dependendo da proteína em questão, foram utilizadas distâncias amostra-detector entre 60 a 140 cm (detalhes apresentados no tópico Resultados). Foi utilizada uma cela porta-amostras na forma de capilar ou em janelas, ambas constituídas de mica, para reduzir ao mínimo possível o espalhamento do sistema e favorecer o espalhamento da amostra. Para o tratamento de dados, todas as curvas foram corrigidas pelos efeitos associados à não

homogeneidade da resposta do detector e pelo decaimento da fonte de luz síncrotron. Os dados foram tratados pelo programa Origin® (Microcal) ou pelo programa TRAT1D (Oliveira, 2003). As características experimentais utilizadas para cada uma das proteínas nas medidas de SAXS estão descritas no capítulo de Resultados e na legenda da respectiva figura de espalhamento.

Para obter informações sobre a *MM* da partícula espalhadora foi utilizada a proteína soroalbumina bovina (BSA; *MM* = 69 kDa) como padrão externo de *MM*, sendo esta diluída no respectivo tampão da proteína de interesse. A partir da razão entre o valor extrapolado de $I(q) \rightarrow 0$ para a proteína de interesse e da BSA, foi obtido uma estimativa da *MM* da partícula espalhadora.

A quantidade de informações obtidas através de uma curva de SAXS pode ser representada através do cálculo do teorema da amostragem de Shannon. A quantidade de informação, ou número de canais de Shannon (*NCS*), pode ser representada pela equação $NCS = D_{max}(q_{max} - q_{min})/\pi$ e com este valor foi possível estimar o número de funções e qual o tipo de modelagem a serem utilizadas para ajustar os dados.

3.22.2 Tratamento dos dados de SAXS e modelagem *ab initio*

A função de distribuição de distâncias $p(r)$ e o raio de giro (R_g) das proteínas estudadas por SAXS foram calculados, a partir das curvas da $I(q)$ corrigidas e normalizadas, pela transformada de Fourier indireta com o auxílio do programa GNOM (Semenyuk & Svergun, 1991). Os modelos em baixa resolução foram construídos a partir dos dados experimentais de SAXS através de um método *ab initio* implementado pelos programas DAMMIM (Svergun, 1999) e GASBOR (Svergun *et al.*, 2001).

O método DAMMIM consiste em construir uma esfera cujo diâmetro corresponde ao D_{max} previamente determinado pela análise da curva da $p(r)$. A esfera é preenchida por uma grade regular de pontos correspondentes a um empacotamento hexagonal de pequenas esferas onde o raio da esfera é muito menor que o D_{max} ($r_{est} \ll D_{max}$). Uma constante (obtida automaticamente pelo uso da interpretação linear $q^4 I(q)$ versus q^4 pelo programa de determinação da forma DAMMIM) foi subtraída dos dados experimentais para garantir que a intensidade a altos ângulos decai como q^{-4} , seguindo a lei de Porod para um modelo de densidade de 2 níveis (Porod, 1982). Este processo reduz a contribuição do espalhamento devido a pequenas flutuações da estrutura interna da proteína e produz uma aproximação da

“curva da forma espalhadora” (i.e. a intensidade espalhadora produzida pelo volume excluído da partícula com uma densidade eletrônica espacial constante).

A estrutura do modelo de pequenas esferas (DAM) é definida pela configuração X , assumindo o valor de cada esfera correspondente ao solvente como {0} ou partícula do soluto {1}. A curva de intensidade de espalhamento do DAM é computada e comparada com a curva da $I(q)$ determinada experimentalmente, e o modelo é progressivamente modificado por sucessivas tentativas de minimização da função $f(X)$ e ajuste da curva experimental. A $f(X)$ pode ser definida como:

$$f(X) = \chi^2 + \alpha P(X) \quad [\text{Equação 23}]$$

onde α é uma constante positiva, $P(X)$ é um função de penalidade que evita, nos modelos, pontos “perdidos” ou não conectados à estrutura principal. χ é a discrepância entre a curva da $I(q)$ experimental e do modelo (Svergun & Stuhmann, 1991) e é dado por:

$$\chi = \sqrt{\frac{1}{N-1} \sum_{j=1}^N \left[\frac{I(q_j) - I_{\text{exp}}(q_j)}{\sigma(q_j)} \right]^2} \quad [\text{Equação 24}]$$

onde N é o número de pontos experimentais obtido do teorema de Shannon, $I_{\text{exp}}(q_j)$ é a intensidade experimental, e $\sigma(q_j)$ é o desvio padrão no ponto j^{th} .

O programa GASBOR (Svergun *et al.*, 2001) utiliza uma representação espacial definida por um conjunto de resíduos de aminoácidos moldes (DR), sendo que o número de DR utilizado é igual ao número de resíduos de aminoácidos da proteína em questão. O programa busca a configuração espacial destes DR que melhor descreve os dados experimentais. Assim, é possível, de maneira similar ao DAMMIM, obter um modelo *ab initio* compacto da proteína utilizando um sistema de pequenas esferas interconectadas, na qual cada DR estará a uma distância regular de um respectivo DR (Svergun & Koch, 2003).

Em suma, os programas DAMMIM e GASBOR procuram por um modelo cuja configuração espacial seja compacta e interconectada, minimizando a função $f(X)$ definida pela equação 23. Eles iniciam a partir de uma configuração espacial correspondente a uma esfera de raio $r = D_{\text{max}}/2$, cujo D_{max} foi previamente determinado pela $p(r)$ através do programa GNOM, preenchendo o modelo *ab initio* com pequenas esferas cujos $r_{\text{esf}} \ll D_{\text{max}}/2$

(Svergun, 1999). DAMMIM e GASBOR constroem modelos *ab initio* que apresentam um bom ajuste entre sua curva da $I(q)$ e aquela obtida experimentalmente, modificando a fase dos átomos do soluto {1} e solvente {0}. Como os dados obtidos não são em alta resolução, o modelo final representa o cálculo da média da superposição de 10 a 20 cálculos independentes de modelos através do programa DAMAVER (Volkov & Svergun, 2003).

3.23 Avaliação dos dados hidrodinâmicos dos modelos *ab initio*

O programa HydroPro (García de la Torre *et al.*, 2000) foi utilizado para estimar o coeficiente de difusão D , o raio de giro R_g , o coeficiente de sedimentação s e a distância máxima D_{max} dos modelos *ab initio* gerados a partir dos dados de SAXS. O programa HydroPro é capaz de criar uma monocamada de mini-esferas na superfície do modelo *ab initio*; o local de exposição da partícula às forças hidrodinâmicas (Carrasco & García de la Torre, 1999). Os cálculos das propriedades hidrodinâmicas foram realizados como descrito por Carrasco & García de la Torre (1999) e foram feitos para mini-esferas de diferentes tamanhos e extrapolados para zero. Os cálculos das propriedades convergem para o valor correto com a redução do tamanho das mini-esferas e com o aumento do número de mini-esferas necessárias para formar a monocamada. Este processo reduz o erro inerente à superfície não coberta pelo arranjo hexagonal das mini-esferas (García de la Torre *et al.*, 2000).

O programa HydroPro foi configurado com os raios dos elementos atômico oriundos dos modelos *ab initio*. O fator sigma de 5 a 8 (como indicado pelo fornecedor) e o raio das mini-esferas entre 6 a 2 Å. O programa Sednterp foi utilizado para fornecer o V_{bar} (em mL/g) a 20 °C. A densidade ρ (em g/mL) e viscosidade η (em poise) utilizadas foram aquelas observadas em água, para se obter dados em condições padrão. A razão friccional f/f_0 , ou fator de Perrin (P), que representa o quanto uma partícula é assimétrica em relação a uma esfera de mesma MM , foi estimado pelo programa SolPro (García de la Torre *et al.*, 1997).

3.24 Modelagem por homologia

Esta metodologia está descrita no manuscrito apresentado no Anexo IV (Borges, J.C., Fischer, H., Craievich, A.F., Hansen, L.D., & Ramos, C.H.I. (2003) Free human mitochondrial GrpE is a symmetric dimer in solution. *J. Biol. Chem.*, 278, 35337-35344) e as informações adicionais para os experimentos estão contidas na legenda da respectiva figura.

4 RESULTADOS

Os resultados estão apresentados divididos por proteínas, sendo que, para as proteínas Hsp70-1A, mortalina e GrpE#2, todos os resultados são apresentados abaixo. Para as proteínas Hsp40 e GrpE#1, os resultados de pesquisa em banco de dados até os resultados de clonagem estão apresentados abaixo. Os demais resultados estão descritos nos manuscritos em anexo (Anexo II e Anexo IV) que são apresentados como parte deste documento.

4.1 Hsp70 humana

Com o intuito de compreender a relação estrutura-função das proteínas Hsp70, foram obtidos os cDNAs de uma representante humana da subfamília Hsc70 (Hsp70-1A, presente no citoplasma) e de uma representante da subfamília DnaK (mortalina, presente na mitocôndria). A Hsp70-1A apresenta 60% de similaridade na sequência de aminoácidos com a proteína DnaK de *E. coli*. A mortalina apresenta 75% e 65% de similaridade com a DnaK e com a Hsp70-1A, respectivamente (Figura 4).

Neste tópico são apresentados os resultados de clonagem, expressão, purificação, ensaios espectrofotométricos, estruturais e hidrodinâmicos da proteína Hsp70-1A humana e clonagem, expressão, purificação e ensaios espectrofotométricos da mortalina.

4.1.1 Hsp70-1A

Pesquisa em banco de dados e obtenção do cDNA plasmidial da Hsp70-1A

A sequência do cDNA da proteína Hsp70 humana (*Homo sapiens heat shock 70kDa protein 1A* – código GenBank número BC002453 – gene HSPA1A) foi utilizada para buscar e identificar ESTs similares no banco de dados GenBank divisão Human EST. Foram encontrados dois clones EST (BE891351 e BE893670; Tabela I), que apresentavam similaridade com a Hsp70-1A e estavam disponíveis para a aquisição comercial junto ao sistema do *IMAGE Consortium*. O cDNA do clone EST número BE891351 foi submetido ao seqüenciamento parcial das extremidades internas do SMC do vetor pCMV-SPORT6 e a análise das seqüências obtidas confirmou a presença do cDNA da proteína Hsp70-1A humana.

Oligonucleotídeos para a PCR da Hsp70-1A

O oligonucleotídeo para a região 5' do cDNA da Hsp70-1A foi desenhado para substituir algumas bases com o objetivo de criar um sítio de restrição para a enzima de restrição *Nde* I (5' GGAACCCATATGGCCAAAGC 3') em fase de leitura e anterior ao códon de iniciação ATG (metionina) da proteína. Assim, o cDNA pôde ser clonado em fase de leitura com o cDNA da cauda de poli-histidina do pET28a. Para a região 3', o oligonucleotídeo foi desenhado para criar, pela substituição de bases, um sítio de restrição para a enzima *Bam*H I (5' GAAAGGATCCTAATCTACC 3') após 4 bases do códon de terminação natural da fase de leitura do cDNA da Hsp70-1A. Outros 3 oligonucleotídeos foram desenhados (5' CCTCTCGCAGGCGGTGC 3', 5' AGGCGTACCTGGGCTACC 3' e 5' GCTCAGGGGATCCTAGTCGTT 3') para o seqüenciamento da região interna do cDNA da Hsp70-1A.

Clonagem do cDNA da Hsp70-1A

As células do clone EST número BE891351 (Tabela I) foram crescidas em meio líquido e o sedimento celular foi utilizado para extrair o DNA plasmidial, com o auxílio do kit *QIAprep MiniPrep Kit* (Qiagen), para ser utilizado como molde para a PCR. Esta reação forneceu um produto de aproximadamente 1.900 pb correspondente ao tamanho esperado para o cDNA completo da proteína Hsp70-1A (código GenBank número BC002453). Este produto de PCR foi purificado do gel de agarose 1% e utilizado para a clonagem em pUC18 como descrito no item 3.9 de Material e Métodos. Desta clonagem, foi selecionada uma colônia contendo vetores pUC18 que liberaram um fragmento de aproximadamente 1.900 pb após incubação com as enzimas *Nde* I e *Bam*H I. A análise da seqüência parcial das extremidades do SMC, obtidas pelo seqüenciamento da região interna do SMC do pUC18 previamente isolado, com os oligonucleotídeos *M13 Forward* e *M13 Reverse*, confirmou a clonagem do cDNA da proteína Hsp70-1A em pUC18: o vetor pUC18aHsp70-1A.

O fragmento liberado pelo vetor pUC18aHsp70-1A, através da digestão pelas enzimas *Nde* I e *Bam*H I, foi utilizado para ser subclonado em pET28a como descrito no item 3.9 de Material e Métodos. A figura 5 representa o perfil eletroforético da digestão enzimática de vetores pET28aHsp70-1A, obtidos de uma colônia previamente selecionada, que liberaram fragmentos de aproximadamente 1.900 pb a partir da incubação com as enzimas *Nde* I e *Bam*H I. A análise das seqüências obtidas pelo seqüenciamento das extremidades internas

do SMC do vetor pET28a selecionado, utilizando os oligonucleotídeos *T7 promoter* e *T7 terminator* (GIBCO BRL), confirmou a clonagem do cDNA da proteína Hsp70-1A no vetor pET28a entre os sítios de restrição para as enzimas *Nde* I e *Bam*H I. A análise das seqüências, obtidas pelo seqüenciamento com os oligonucleotídeos desenhados para se anelarem no interior do cDNA da Hsp70-1A, mostrou que a seqüência clonada correspondia à seqüência depositada para o cDNA da Hsp70-1A (código GenBank número BC002453) e em fase de leitura com a seqüência de DNA da cauda de poli-histidina na extremidade amino-terminal. Assim foi obtido o vetor pET28aHsp70-1A, responsável pela expressão de uma proteína, em fusão à cauda de poli-histidina na região amino-terminal, constituída por 661 resíduos de aminoácidos e *MM* igual a 72.215 g/mol.

Expressão e purificação da proteína Hsp70-1A

A figura 6 mostra o resultado do processo de expressão e purificação, por cromatografia de afinidade, da proteína Hsp70-1A como descrito nos itens 3.13 e 3.14 de Materiais e métodos. A proteína foi expressa pela indução por IPTG, sendo identificada no sobrenadante do lisado (Figura 6). A purificação da Hsp70-1A foi realizada em dois passos cromatográficos: cromatografia de afinidade em resina contendo níquel e CEM. A fração separada pela cromatografia de afinidade foi submetida a um tratamento com a enzima fosfatase alcalina ácida, para a eliminação de qualquer vestígio de nucleotídeos adenosina (Theyssen *et al.*, 1996). A fosfatase alcalina ácida desfosforila o grupamento fosfato ligado à ribose dos nucleotídeos levando à formação de Pi e a um nucleosídeo. Após este passo, a amostra da Hsp70-1A foi submetida a uma diálise contra o tampão TKP e, em seguida, submetida a uma CEM como descrito no item 3.14.3 de Material e Métodos utilizando o tampão TKP.

Na CEM, a proteína Hsp70-1A apresentou um perfil de eluição de 3 espécies com *MM* aparente de 600 kDa, 200-150 kDa e 70 kDa (Figura 7), sendo que todas as frações migraram com *MM* aparente correspondente a 70 kDa na SDS-PAGE, mesmo na presença de DTT (Figura 8). A fração de 600 kDa não se apresentou em equilíbrio com as demais frações observadas nas condições testadas, pois sua reaplicação à CEM não resultou no aparecimento das demais frações (Figura 7). A fração 200-150 kDa apresentou ser, provavelmente, uma espécie em equilíbrio dímero-monômero, pois quando ela foi reaplicada à CEM, houve a eluição da proteína nas frações correspondente a 200-150 kDa e 70 kDa,

sendo esta última a espécie principal (Figura 7). A fração de 70 kDa quando reaplicada logo após a primeira cromatografia apresentou um perfil de eluição de uma espécie monomérica (Figura 7). Após estocagem, a fração de 70 kDa passou a apresentar a espécie de 200-150 kDa, porém com predominância da espécie de 70 kDa, quando analisada por CEM analítica (dados não mostrados). Para evitar o comportamento de polidispersidade, todos os experimentos foram realizados com a fração de 70 kDa recém purificada e com a sua monodispersidade monitorada por EDL. Os resultados indicam que a Hsp70-1A foi expressa em quantidade, e foi purificada, a partir do sobrenadante do lisado, com pureza superior a 95% (Figura 8).

Análise de estrutura secundária da proteína Hsp70-1A por CD

As 3 frações obtidas pelo processo de purificação (como descrito acima) foram analisadas por CD. A fração de 70 kDa apresentou um espectro característico de proteína constituída de estrutura secundária do tipo α -hélice (Figura 9). A fração de 600 kDa apresentou menor quantidade de estrutura secundária do que a fração de 70 kDa (Figura 9). Entretanto, a proteína eluída na fração de 600 kDa apresentou estruturação em α -hélices e folha β -pregueada, como pode ser avaliado pelo perfil do espectro de CD (Figura 9). Provavelmente a fração de 600 kDa constituía um agregado protéico da Hsp70-1A, formado pela alta indução heteróloga, não estando em equilíbrio com a forma nativa, como relatado no tópico anterior, e foi descartada das outras análises. A fração de 200-150 kDa apresentou o mesmo perfil de espectro de CD que a fração de 70 kDa (dados não mostrados). Entretanto devido a polidispersividade da amostra 200-150 kDa e o objetivo de estudar a influência da ligação de nucleotídeos adenosina na estrutura da proteína Hsp70-1A por técnicas como SAXS e UAL, tal fração também foi descartada para as demais análises.

A análise do perfil do espectro de CD apresentado pela Hsp70-1A sugere que esta proteína possui estrutura secundária constituída de α -hélices. A análise da quantidade de estrutura secundária para o espectro de CD da Hsp70-1A feita pelo programa CDNN Deconvolution mostrou que a Hsp70-1A é constituída por aproximadamente 30% de α -hélices, 18% de folha β -pregueada, 17% de alças e 35% de estrutura randômica (erro < 10%). A adição de MgAMP não provocou mudanças significativas no espectro de CD da Hsp70-1A entre os comprimentos de onda de 200 a 260 nm. Entretanto, uma pequena redução na intensidade do sinal foi observada na região de 190 a 200 nm (Figura 10). A

adição dos nucleotídeos MgADP ou MgATP nas amostras da Hsp70-1A induziram pequenas alterações no perfil do espectro de CD na região de 208 nm em comparação com os espectros da proteína na ausência de nucleotídeos e na presença de MgAMP. Na região de 190 a 200 nm, os nucleotídeos MgADP e MgATP induziram um aumento considerável na quantidade de luz absorvida à esquerda, sendo o aumento similar para os dois nucleotídeos (Figura 10). Nas amostras de Hsp70-1A na presença de MgATP ou MgADP, a quantidade de estrutura estimada pelo programa CDNN Deconvolution foi de: 36% de α -hélices, 15% de folha β -pregueada, 16% de alças e 31% de estrutura randômica (erro < 10%). Em comparação à quantidade de estrutura estimada para a proteína Hsp70-1A sem aditivos, a adição dos nucleotídeos MgADP ou MgATP provocaram um ganho de estrutura em α -hélices, o que sugere uma estruturação da proteína induzida pela ligação dos nucleotídeos.

Desenovelamento térmico da proteína Hsp70-1A monitorado por CD

Os ensaios de desnaturação térmica (20 a 90 °C) da Hsp70-1A, no tampão TKP, foram acompanhados por CD (Figura 11). A Hsp70-1A sem aditivos apresentou Tms nas seguintes temperaturas: 46, 68 e 82 °C. A desnaturação térmica da Hsp70-1A nestas condições foi irreversível, mesmo que a amostra fosse aquecida somente até 50 °C. A proteína sofreu agregação seguida de precipitação quando foi aquecida até 90 °C (um precipitado esbranquiçado foi observado na cubeta). A tabela II sumariza as Tms, nas diversas condições testadas, determinadas pela análise da primeira derivativa da curva de desenovelamento térmico com a interpretação da área da transição por curvas gaussianas, nas quais o máximo do ajuste foi considerado o valor da Tm. A Tm foi dada pela média de 3 curvas de desenovelamento (mais o desvio padrão da média). Considerando que os dados foram coletados em intervalos de 0,5 °C, este deve ser considerado o erro mínimo na Tm.

As Tms apresentadas pela curva de desnaturação da Hsp70-1A na presença de AMP não sofreram modificações em relação à proteína sem aditivos (Tabela II). Porém, na presença de ADP ou ATP, a Tm₁ e a Tm₃ sofreram um aumento de aproximadamente 4 °C (Tabela II). A presença de MgCl₂ não provocou mudanças nos Tms da curva de desenovelamento da Hsp70-1A. Mas na presença de MgADP ou MgATP, a curva de desnaturação da Hsp70-1A apresentou um aumento, no Tm₁, de 9 e 7 °C, respectivamente, e com o mesmo efeito com as amostras com apenas ADP ou ATP no Tm₃.

Tabela II: Temperaturas médias de transição para a Hsp70-1A na ausência e na presença de aditivos. O desenovelamento térmico da Hsp70-1A (6 $\mu\text{mol/L}$), realizado no tampão TKP na ausência e na presença dos diferentes aditivos (200 $\mu\text{mol/L}$ MgCl_2 , AMP, MgAMP, ADP, MgADP, ATP ou MgATP), foi acompanhado por CD a 222 nm.

Hsp70-1A	T_{m1} ($^{\circ}\text{C}$)	T_{m2} ($^{\circ}\text{C}$)	T_{m3} ($^{\circ}\text{C}$)
Sem aditivos	$46,2 \pm 0,5$	$68,0 \pm 0,5$	$82,5 \pm 0,5$
MgCl_2	$46,5 \pm 0,5$	$67,0 \pm 0,9$	$81,2 \pm 0,5$
AMP	$46,6 \pm 0,5$	$66,5 \pm 0,5$	$82,4 \pm 0,2$
ADP	$50,6 \pm 0,5$	$67,6 \pm 0,5$	$86,7 \pm 0,5$
ATP	$50,2 \pm 0,5$	$67,5 \pm 0,8$	$85,8 \pm 0,5$
MgAMP	$46,8 \pm 0,5$	$65,7 \pm 0,5$	$81,1 \pm 0,7$
MgADP	$55,3 \pm 0,5$	$67,6 \pm 0,5$	$85,8 \pm 0,5$
MgATP	$53,1 \pm 0,5$	$67,8 \pm 0,6$	$86,0 \pm 0,5$

Experimentos de fluorescência do triptofano da proteína Hsp70-1A

A figura 12 apresenta os espectros de emissão de fluorescência de triptofano (excitação em 295 nm) da Hsp70-1A no tampão TKP, na ausência de nucleotídeos, na presença de 1 mmol/L de MgAMP, MgADP e MgATP e em condições desnaturantes (Gnd-HCl 6 mol/L). A estrutura primária da Hsp70-1A possui dois resíduos de triptofano (Figura 4 – W90 localizado no DLN e W580 localizado na “tampa” do DLS) e a análise dos espectros de emissão de fluorescência da figura 12 sugere que um deles pode estar protegido do solvente e o outro pode estar exposto ao solvente. A Hsp70-1A apresentou em seu espectro de emissão de fluorescência, um λ máximo de intensidade de fluorescência em aproximadamente 336 nm e um “ombro” em aproximadamente 355 nm. A interpretação do espectro citado acima por duas curvas Gaussianas mostrou a presença de dois máximos em aproximadamente 330 e 355 nm. Na presença de 6 mol/L de Gnd-HCl, o λ máximo de intensidade de emissão de fluorescência foi observado em aproximadamente 356 nm, o mesmo λ máximo observado para o triptofano em água (dados não mostrados). A presença de MgAMP ou de MgADP não induziram mudanças perceptíveis no espectro de emissão de fluorescência da Hsp70-1A. Entretanto, a presença de MgATP reduziu a intensidade de fluorescência em aproximadamente 7% da área da integral da curva, porém sem modificar o

λ máximo de intensidade do espectro de emissão de fluorescência. Tanto que não foi observada mudança no $\langle\lambda\rangle$ dos espectros de emissão de fluorescência da Hsp70-1A devido à presença de nucleotídeos adenosina. Exceto para a Hsp70-1A na presença de Gnd-HCl, o $\langle\lambda\rangle$ da Hsp70-1A, na ausência e na presença de nucleotídeos adenosina, calculado foi de aproximadamente 345 nm. Os resultados representam a média de 3 preparações independentes coletadas com o equipamento configurado para a amostra de Hsp70-1A sem aditivos.

Experimentos de ITC para a proteína Hsp70-1A e seus ligantes MgATP e MgADP

Experimentos de ITC foram realizados para determinar a K_A da Hsp70-1A pelos nucleotídeos MgADP e MgATP à temperatura constante de 20 °C. A Hsp70-1A, recém purificada, foi preparada no tampão TKP em uma concentração de proteína de aproximadamente 15 $\mu\text{mol/L}$. A concentração de proteínas foi determinada para uma alíquota da mesma amostra utilizada no experimento de ITC, como descrito no item 3.16.1 de Material e métodos. As soluções utilizadas de MgATP ou MgADP foram preparadas no tampão TKP em uma concentração de aproximadamente 300-350 $\mu\text{mol/L}$. A concentração de nucleotídeos foi determinada espectrofotometricamente a partir de uma alíquota da mesma solução utilizada no experimento, conforme descrito no item 3.16.2 de Material e métodos. Os dados do calor médio liberado pela titulação do nucleotídeo na solução protéica saturada pelo mesmo, que foi observado nos 10 últimos pontos de titulação (calor de diluição do titulante), foram subtraídos, após normalização pela concentração, da curva experimental. A curva do calor de diluição do MgATP no tampão TKP apresentou um maior calor liberado nas primeiras titulações que reduziu nas demais titulações, porém o mesmo não foi observado para o MgADP (dados não mostrados). Esta tendência da titulação do MgATP, no tampão TKP, também foi subtraída da curva experimental, após o procedimento de normalização pela concentração e subtração do calor de diluição do titulante da curva de titulação do MgATP no tampão TKP, como definido acima.

A figura 13 representa os experimentos de ITC após o tratamento dos dados utilizando o programa Origin[®] (Microcal). Os ajustes dos dados para os nucleotídeos MgADP e MgATP foram feitos utilizando um sítio de ligação por molécula de Hsp70-1A e os resultados representam a média da análise de 3 experimentos individuais. Os dados sugerem uma maior afinidade da Hsp70-1A por MgADP do que por MgATP ($3,8 \pm 0,2 \times 10^6 \text{ mol/L}^{-1}$ versus

$1,4 \pm 0,1 \times 10^6 \text{ mol/L}^{-1}$, respectivamente). Os dados de ΔH sugerem valores significativamente diferentes entre ADP e ATP, sendo mais negativo para ADP do que para ATP ($-5,8 \pm 0,1$ versus $-4,6 \pm 0,1 \text{ kcal/mol}$, respectivamente). O ΔS foi relativamente maior para a titulação com ATP em comparação ao ADP ($12,4 \pm 0,1$ versus $10,2 \pm 0,6 \text{ cal/mol}^\circ\text{C}$, respectivamente). Os resultados mostraram que a Hsp70-1A possui constantes de dissociação ($K_D = 1/K_A$) pelos nucleotídeos MgATP e MgADP na faixa submicromolar (aproximadamente $0,71$ versus $0,26 \text{ } \mu\text{mol/L}$, respectivamente) e nortearam as condições de saturação, das amostras de Hsp70-1A pelos nucleotídeos adenosina, para os experimentos de UAL (ver abaixo).

Análise de velocidade de sedimentação da proteína Hsp70-1A por UAL

Os experimentos de UAL para a proteína Hsp70-1A, utilizando a técnica de velocidade de sedimentação, foram realizados na presença e na ausência dos nucleotídeos adenosina. Em todos os experimentos realizados, a proteína se comportou como uma solução monodispersa, não apresentando sinais evidentes de agregação ou oligomerização. A concentração de Hsp70-1A utilizada foi de 150 a 1.000 $\mu\text{g/mL}$ (2 a 13,8 $\mu\text{mol/L}$) em tampão TKP e também em soluções contendo 200 $\mu\text{mol/L}$ de MgATP ou MgADP que foram utilizados para avaliar as mudanças nas propriedades hidrodinâmicas da molécula de Hsp70-1A.

A figura 14a mostra a análise de um conjunto de dados de velocidade de sedimentação pelo método *Sedimentation Time Derivative* (distribuição da integral $g(s^*)$ – item 3.20.1 de Material e Métodos) da Hsp70-1A (600 $\mu\text{g/mL}$), na ausência e na presença de MgADP ou MgATP. Cada conjunto de dados foi analisado com o mesmo número de curvas obtidas. A figura 14b mostra o cálculo do $s_{20,w}^0$ da Hsp70-1A na concentração zero de proteína, na ausência e na presença de nucleotídeos: $s_{20,w}^0 = 4,09 \pm 0,01 \text{ S}$; $s_{20,w}^0 \text{ ATP} = 4,44 \pm 0,03 \text{ S}$ e $s_{20,w}^0 \text{ ADP} = 4,21 \pm 0,01 \text{ S}$.

Estimativa do coeficiente de difusão da proteína Hsp70-1A por EDL

Os dados de D da Hsp70-1A foram estimados pela técnica de EDL em 4 diferentes concentrações de proteína (0,5, 1,0, 1,75 e 2,5 mg/mL), normalizados para as condições padrão em água e na temperatura de 20 $^\circ\text{C}$ ($D_{20,w}$) pela aplicação da equação 22 e extrapolados para a concentração de zero (mg/mL) de proteína ($D_{20,w}^0$). Para isto, foi realizado um procedimento similar ao realizado para a determinação do $s_{20,w}^0$ (Figura 14b).

Assim, os erros devido à dependência da concentração foram reduzidos. Os $D_{20,w}^0$ obtidos para a Hsp70-1A nas diferentes condições foram: $D_{20,w}^0 = 5,4 \pm 0,1 \cdot 10^{-7} \text{cm}^2 \text{s}^{-1}$; $D_{20,w}^0 \text{ ATP} = 5,7 \pm 0,1 \cdot 10^{-7} \text{cm}^2 \text{s}^{-1}$ e $D_{20,w}^0 \text{ ADP} = 5,5 \pm 0,1 \cdot 10^{-7} \text{cm}^2 \text{s}^{-1}$. A aplicação da equação 10, utilizando os dados de $s_{20,w}^0$ e de $D_{20,w}^0$, mostraram que a Hsp70-1A possui MM de aproximadamente 70 kDa ($MM_{\text{sem aditivos}} = 69 \pm 1 \text{ kDa}$; $MM_{\text{ATP}} = 72 \pm 1 \text{ kDa}$; $MM_{\text{ADP}} = 71 \pm 1 \text{ kDa}$).

Análise de sedimentação em equilíbrio da proteína Hsp70-1A

Os experimentos de sedimentação em equilíbrio da Hsp70-1A, que foi previamente preparada no tampão TKP, foram realizados em 3 diferentes velocidades (6.000, 8.000 e 10.000 rpm) e em 3 concentrações diferentes (100, 250 e 500 $\mu\text{g/mL}$). Em todas as condições, as curvas foram bem ajustadas como uma única espécie exponencial, resultando em uma partícula de $MM = 70 \pm 2 \text{ kDa}$. A figura 15 mostra o resultado das análises dos dados obtidos com a proteína na concentração de 250 $\mu\text{g/mL}$ em 3 velocidades diferentes. A análise dos residuais mostra a boa qualidade do ajuste das curvas para uma partícula de aproximadamente 70 kDa. Os dados também foram analisados como um sistema de auto-associação e eles foram bem ajustados como um sistema mono-disperso como descrito acima. Os dados sugerem uma única espécie protéica em solução e que o processo de associação, observado nos experimentos de CEM, não ocorreu em virtude das condições utilizadas: baixa concentração protéica, o tampão TKP e o tempo de utilização da amostra.

Experimentos de SAXS para a proteína Hsp70-1A na presença de MgATP e MgADP

Os experimentos de SAXS foram realizados na linha SAS do LNLS (Kellermann *et al.*, 1997) utilizando um porta-amostras de mica na forma de capilar, em duas distâncias diferentes entre o porta-amostra e o detector (750 e 860 mm), e λ de 1,488 Å. As medidas foram realizadas em várias concentrações da proteína Hsp70-1A no tampão TKP (3,50, 6,70, 10,8 e 15,6 mg/mL) na ausência e na presença de 1 mmol/L de MgADP ou MgATP. Os dados foram tratados utilizando o programa TRAT1D (Oliveira, 2003) e foram normalizadas pela concentração para se obter as curvas das $I(q)$. As partes das curvas das $I(q)$, em alto e baixo “ q ”, oriundas a partir de curvas de diferentes concentrações de proteína e coletadas em diferentes distâncias, foram unidas para se obter dados de boa qualidade estatística em um grande intervalo de “ q ”. As curvas das $p(r)$ e os R_{gs} da Hsp70-1A nas diferentes condições

experimentais foram calculados, a partir das curvas das $I(q)$ descritas acima, pela transformada de Fourier indireta do programa GNOM (Semenyuk & Svergun, 1991). A figura 16 representa as curvas das $I(q)$ de SAXS para a proteína Hsp70-1A, em diferentes condições e com o ajuste realizado pelo programa GNOM para o cálculo das curvas das $p(r)$ (Figura 17).

A estimativa de MM para a Hsp70-1A a partir dos dados de SAXS, nas diferentes condições experimentais, foi feita utilizando a proteína BSA como padrão de MM secundário. Os dados obtidos são apresentados na tabela III e sugerem que as amostras sem aditivos e com MgADP (77 kDa) estão em boa concordância com o esperado para a Hsp70-1A na forma de monômeros (72.215 g/mol). Isto forneceu indícios de que não houve a formação de agregados nas amostras utilizadas nas medidas. Entretanto, a Hsp70-1A na presença de MgATP apresentou uma MM de aproximadamente 100 kDa, indicando a possibilidade da presença de agregação nesta condição experimental, o que pode prejudicar a análise destes dados.

As curvas das $p(r)$ obtidas para a Hsp70-1A nas diferentes condições (Figura 17), sugerem que a presença dos nucleotídeos adenosina (MgATP e MgADP) provocou pequenas mudanças conformacionais na proteína em relação à amostra sem aditivos. Porém, em todas as condições, a proteína apresentou uma conformação prolata alongada, com um D_{max} de aproximadamente 125-130 Å (Tabela IV). A análise do R_g da Hsp70-1A, obtidos a partir das curvas das $p(r)$ na diferentes condições experimentais, também sugeriu que os nucleotídeos adenosina provocaram pequenas mudanças conformacionais na partícula espalhadora (Tabela IV).

A figura 18 representa o gráfico de Kratky dos dados da Hsp70-1A nas condições testadas. O perfil da curva da $I(q) \times q^2$ versus q sugere que a proteína Hsp70-1A apresentou grau de compactação similar, tanto na ausência quanto na presença de MgATP ou MgADP. Ou seja, a Hsp70-1A não apresentou segmentos evidentes com alta flexibilidade. Os perfis das curvas de Kratky mostram a presença de um ombro nas 3 curvas, sugerindo que a proteína apresenta uma certa assimetria estrutural e que a presença de nucleotídeos adenosina induziram pequenas mudanças conformacionais na Hsp70-1A.

Modelos *ab initio* para a proteína Hsp70-1A

Os modelos *ab initio* obtidos para a Hsp70-1A, a partir dos dados de SAXS na ausência e na presença dos nucleotídeos adenosina, foram construídos pelo programa GASBOR. Como entrada de dados ao programa, foi fornecido o número de resíduos de aminoácidos da proteína (661) e o programa calculou a configuração destes DR que melhor ajustou os dados experimentais da curva da $I(q)$. Como os dados experimentais de SAXS não são em alta resolução (Tabela III), os modelos finais construídos (Figura 19) são representados pela média do preenchimento dos DR no espaço pela superposição de 10 modelos *ab initio* gerados independentemente. O cálculo da média da superposição foi realizado pelo programa DAMAVER. Os modelos *ab initio* são apresentados na figura 19 e o ajuste da curva da $I(q)$ dos modelos nos dados experimentais, apresentadas no painel inferior da figura 19, mostra a qualidade dos dados. A análise dos modelos *ab initio* construídos sugere que a proteína Hsp70-1A possui a mesma forma geral (prolata alongada). Todavia, o modelo *ab initio* obtido dos dados experimentais na presença de MgATP, é mais diferenciado dos outros dois (sem aditivos e na presença de MgADP). A tabela III apresenta os parâmetros matemáticos e estruturais obtidos dos dados experimentais de SAXS e dos modelos *ab initio* gerados pelo programa GASBOR.

Tabela III: Parâmetros matemáticos e estruturais obtidos pelos experimentos de SAXS da proteína Hsp70-1A. Os dados de SAXS foram obtidos na ausência e na presença de nucleotídeos adenosina e são comparados com aqueles gerados para os modelos *ab initio*.

Parâmetros matemáticos e estruturais	Hsp70-1A					
	Sem Aditivos		MgADP		MgATP	
	SAXS	<i>ab initio</i>	SAXS	<i>ab initio</i>	SAXS	<i>ab initio</i>
NCS [#]	-	10,01	-	9,935	-	10,32
Discrepância χ	0,187	0,468	0,122	0,427	0,099	0,413
Resolução (Å)	25	25	25	25	25	25
MM (kDa)	76 ± 10	-	77 ± 11	-	100 ± 15*	-

[#] $NCS = D_{max} (q_{max} - q_{min})/\pi$ representa uma medida da quantidade de informação contida na curva de SAXS; *esta amostra provavelmente sofreu efeitos de agregação de proteína.

As análises das propriedades hidrodinâmicas dos modelos *ab initio* foram realizadas com o auxílio do programa HydroPro (García de la Torre *et al.*, 2000). Os parâmetros hidrodinâmicos para cada condição foram obtidos para cada um dos 10 modelos *ab initio* obtidos de maneira independente para as 3 diferentes condições experimentais, e são representados pela média dos dados mais os desvios-padrão da média. Para isto, foram utilizados a *MM* correspondente ao monômero, η e ρ em condições padrão (água e 20 °C) e V_{bar} de 0,734 mL/mg que foi estimado a partir da sequência primária da proteína Hsp70-1A mais a cauda de poli-histidina. Estas propriedades foram estimadas pelo programa Sednterp. A tabela IV apresenta os parâmetros hidrodinâmicos obtidos experimentalmente e os compara com os parâmetros estimados pelo programa HydroPro, a partir dos modelos *ab initio* gerados a partir dos dados de SAXS.

Tabela IV: Conjunto de parâmetros hidrodinâmicos determinados para a Hsp70-1A na ausência e na presença de MgADP ou MgATP. Os dados experimentais (Exp.) hidrodinâmicos são comparados com os dados obtidos para os modelos *ab initio* gerados a partir dos dados de SAXS utilizando o programa HydroPro (em água e 20 °C). $s_{20,w}^0$ (experimentos de velocidade de sedimentação), $D_{20,w}^0$ (experimentos de EDL), *MM* (calculada a partir dos dados de $s_{20,w}^0$ e de $D_{20,w}^0$, pela equação 10), D_{max} e R_g (dados das curvas da $p(r)$ – Figura 17).

Parâmetros hidrodinâmicos	Hsp70-1A					
	Sem Aditivos		MgADP		MgATP	
	Exp.	HydroPro	Exp.	HydroPro	Exp.	HydroPro
$s_{20,w}^0$ (S)	4,09 ± 0,01	4,06 ± 0,02	4,21 ± 0,01	4,15 ± 0,01	4,44 ± 0,03	3,98 ± 0,0
$D_{20,w}^0$ (10 ⁻⁷ cm ² s ⁻¹)	5,4 ± 0,1	5,20 ± 0,03	5,5 ± 0,1	5,30 ± 0,04	5,7 ± 0,1	5,09 ± 0,0
<i>MM</i> (kDa)	69 ± 1	-	71 ± 1	-	72 ± 1	-
D_{max} (Å)	~ 126	134 ± 1	~ 125	133 ± 1	~ 130	139 ± 1
R_g (Å)	38,1 ± 0,4	37,8 ± 0,4	37,4 ± 0,4	37,2 ± 0,3	39,1 ± 0,5	38,0 ± 0,

Superposição das estruturas em alta resolução das Hsp70 com os modelos ab initio

A figura 20 representa a superposição dos modelos *ab initio* construídos para a Hsp70-1A, a partir dos dados de SAXS nas três diferentes condições experimentais, com as estruturas em alta resolução disponíveis no banco de dados PDB. A superposição foi realizada manualmente com o auxílio do programa WebLab ViewerLite e são apresentadas em duas posições através da rotação de 90 graus no eixo axial. A análise da figura mostra que, de modo geral, os modelos *ab initio* construídos para a Hsp70-1A acomodaram bem os modelos em alta resolução dos domínios de proteínas Hsp70. Porém, dependendo da condição experimental, a posição relativa destes domínios foi modificada para melhorar a acomodação no modelo *ab initio*. Isto sugere que os nucleotídeos adenosina provocam mudanças conformacionais gerais na Hsp70-1A sem resultar em mudanças no formato geral da proteína.

Hsp70-1A	1	<u>MAKAAAIGIDLGTTSCVGVFQHGKVEI IANDQGNRTTSPSYVAFD-DTERLIGDAAKNQVALNPONTVFD</u>
Mortalina	1	<u>MIKGAVVGIDLGTTNSCVAVMEGKQAKVLENAEGARTTSPSVVAFDADGERLVGMPAKRQAVTNPNNTFYA</u>
DnaK	1	<u>--MGKIIGIDLGTTNSCVAIMDGTTPRVLENAEGDRTPSIIAYTQDGETLVGQPAKQAVTNPNQNTLFA</u>
		***** : *
Hsp70-1A	70	<u>AKRLIGRKFGDPVVQSDMKHWPQVINDGDKPKVQVSYKGETKAFYPPEISSMVLTKMKEIAEAYLGYPV</u>
Mortalina	71	<u>TKRLIGRRYDDPEVQKDIKNVPFKIVRASNG---DAWVEAHGKLYSPSQIGAFVLMKMKETAENYLGHTA</u>
DnaK	69	<u>IKRLIGRRFQDEEVQDVSIMPFKIIAADNG---DAWVEVKGQKMAPQISAEVLKMKKTAEDYLGEFV</u>
		***** : *
Hsp70-1A	140	<u>TNAVITVPAYFNDSQRQATKDAGVIAGLNVLRIINEPTAAAIAYGLDRTGKGERNVLI FDLGGGTFDVS I</u>
Mortalina	138	<u>KNAVITVPAYFNDSQRQATKDAGQISGLNVLRVINEPTAAALAYGLDK--SEDKVIAVYDLGGGTFD I S I</u>
DnaK	136	<u>TEAVITVPAYFNDAQRQATKDAGRIAGLEVKRIINEPTAAALAYGLDKG-TGNRTIAVYDLGGGTFD I S I</u>
		***** : *
Hsp70-1A	210	<u>LTIDDG---IFEVKATAGDTHLGGEDFDNRLNVHFVEEFKRKHKKDISQNKRAVRLRTACERAKRTLS</u>
Mortalina	206	<u>LEIQKG---VFEVKSTNGDTHLGGEDFDQALLRHIVKEFKRETGVDLTKDNMALQVRVREAAEKAKCELS</u>
DnaK	205	<u>IEIDEVDGEKTFEVLATNGDTHLGGEDFDRLINYLVEEFKKDQGDILRNDPLAMQRLKEAAEKAKCELS</u>
		***** : *
Hsp70-1A	276	<u>SSTQASLEIDSLF---EGIDFYTSITRARFEELCSDLFRSTLEPVEKALRDAKLDKAIHDLVLVGGG</u>
Mortalina	272	<u>SSVQTDINLPYLTMDSSG-PKHLNMKLTRAQFEGIVTDLIRRTIAPCQKAMQDAEVSKSDIGEVILVGGM</u>
DnaK	275	<u>SAQQTVDNLPYITADATG-PKHMNIKVTRAKLESLEVDLVNRSIEPLKVALQDAGLSVSDIDDVILVGGQ</u>
		***** : *
Hsp70-1A	341	<u>TRIPKVQKLLQDFFNGRDLNKSINPDEAVAYGAAVQAAILMGDKSENVQDLLLLDVAPLSLGLTAGGV M</u>
Mortalina	341	<u>TRMPKVQQTVDLFG-RAPSKAVNPDEAVAIGAAIQGGVLAG---DVTVDVLLLDVTPLSLGIETLGGV F</u>
DnaK	344	<u>TRMPMVQKKVAEFFG-KEPRKDVNPDEAVAIGAAVQGGVLTG---DVKDVLLLDVTPLSLGIETMGGV M</u>
		***** : *
Hsp70-1A	411	<u>TALIKRNSTIPKQTQIFTTYSNQPGLVQVYEGEGERAMTKDNLLGRFELSGIPAPRGVPQIEVTFDI</u>
Mortalina	406	<u>TKLINRNTTIPKKSQVFESTAAGQTQVEIKVCQGEREMAGDNKLLGQFTLIGIPAPRGVPQIEVTFDI</u>
DnaK	409	<u>TTLIAKNTTIPKHSQVFSTAEDNQSAVTIHLVQGERKRAADNKS LGQFNLDGINPAPRGMPQIEVTFDI</u>
		***** : *
Hsp70-1A	481	<u>DANGILNVTATDKSTGKANKITITNDKGRLSKEEIERMVQEAKEYKAEDVQRRERSAKNALESYAFNMK</u>
Mortalina	476	<u>DANGIVHVSADKDKGTRREQQIVIQSSGG-LSKDDIENMVKNAEKYAEEDRRKKERVEAVNMAEGIIHDTE</u>
DnaK	479	<u>DADGILHVSADKDKNGSKEKQITIKASSG-LNDEIQKMRDAEANAADRFEEVLQTRNQDHLHSTR</u>
		***** : *
Hsp70-1A	551	<u>SAVEDEGLKGKISEADKKKVLDKCQEVISWLDANTLAEKDEFEHKRKELEQVCNPIISGLYQAGGPGPG</u>
Mortalina	545	<u>TKMEEFKDQ--LPADCNKLKEEISKMRELLARKDSETGENIRQAASSLQQASLKL FEMAYKKMASEREG</u>
DnaK	548	<u>KQVEEAGDK--LPADDKTAIESALTALTKGEDK---AAIEAKMQELAQVSKLMEIAQQQHAQQQTA</u>
		***** : *
Hsp70-1A	621	<u>GFGAQGPKGGSGSGPTIEEVD----</u>
Mortalina	613	<u>SGSSGTGEQKEDQKEEKQ-----</u>
DnaK	613	<u>GADASANNAKDDVDVDAEFEEVKDKK</u>
		***** : *

Figura 4: Alinhamento das seqüências de aminoácidos da Hsp70-1A, da mortalina e da DnaK. Apresentação das seqüências de aminoácidos das proteínas humanas Hsp70-1A (código GenBank número BC002453) e mortalina (código GenBank número BC000478 – sem o peptídeo sinal para endereçamento da proteína para a mitocôndria) e da DnaK de *E. coli* (código GenBank número P04475) alinhadas pelo programa ClustalW (clustalw.genome.ad.jp/). A Hsp70-1A apresentou 65% e 60% de similaridade com a mortalina e com a DnaK, respectivamente. Estas últimas apresentaram similaridade de 75% entre si. A região sublinhada corresponde ao DLN, sendo o restante correspondente ao DLS. A Hsp70-1A possui 2 resíduos de triptofano (W90 e W580 – vermelho) enquanto a mortalina e DnaK possuem apenas um resíduo (W105 e W102, respectivamente – vermelho). Os aminoácidos que são importantes para classificação das Hsp70 em subfamílias (Brehmer *et al.*, 2001) estão apresentados em azul.

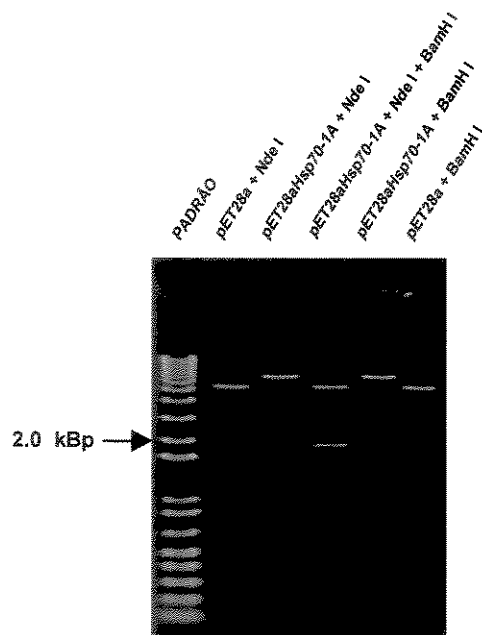


Figura 5: Verificação da clonagem do cDNA da Hsp70-1A no vetor pET28a. Gel de agarose 1% mostrando o perfil eletroforético da digestão dos vetores pET28aHsp70-1A e pET28a com as enzimas *Nde* I e *Bam*H I. A digestão dupla do vetor pET28aHsp70-1A levou à liberação de um fragmento de aproximadamente 1.900 pb, como determinado pela comparação com o padrão de *MM 1 Kb Plus DNA Ladder* (GIBCO BRL). A digestão individual dos vetores com cada enzima evidenciou a diferença do tamanho do vetor pET28aHsp70-1A em relação ao pET28a. A análise das seqüências das extremidades do SMC do vetor pET28aHsp70-1A, com o auxílio dos oligonucleotídeos *T7 promoter*, *T7 terminator* e por aqueles desenhados (descritos acima) para se anelarem na região central do cDNA da Hsp70-1A, confirmou a clonagem do cDNA correto da proteína Hsp70-1A (código GenBank número BC002453) em fase de leitura com o cDNA da cauda de polihistidina do vetor pET28a.

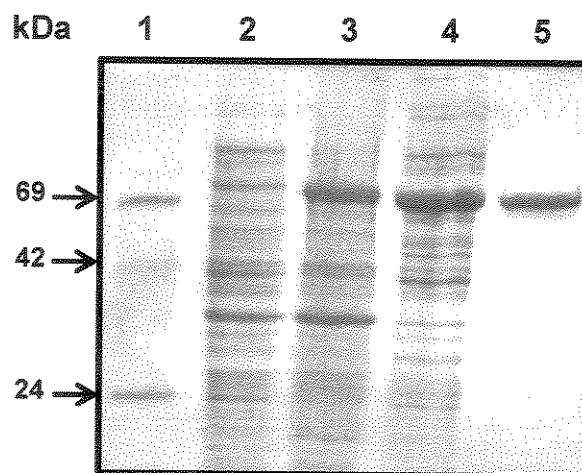


Figura 6: A Hsp70-1A foi altamente expressa e apresentou alto grau de pureza após o processo de purificação. Experimentos de SDS-PAGE 12% mostraram o perfil de expressão da Hsp70-1A recombinante e o grau de pureza da proteína após o processo de purificação. Canaleta 1: marcadores de *MM* em kDa (à esquerda); canaleta 2: sedimento bacteriano antes de indução; canaleta 3: sedimento bacteriano após 5 horas de indução com 0,4 mmol/L de IPTG; canaleta 4: sobrenadante do lisado bacteriano; canaleta 5: fração de 70 kDa correspondente a Hsp70-1A após a purificação por cromatografia de afinidade (veja item 4.1.4 de Resultados para detalhes). A Hsp70-1A foi expressa em alta quantidade, apresentando um rendimento de aproximadamente 30 mg/L de indução (dados não mostrados), e foi purificada a partir da fração solúvel.

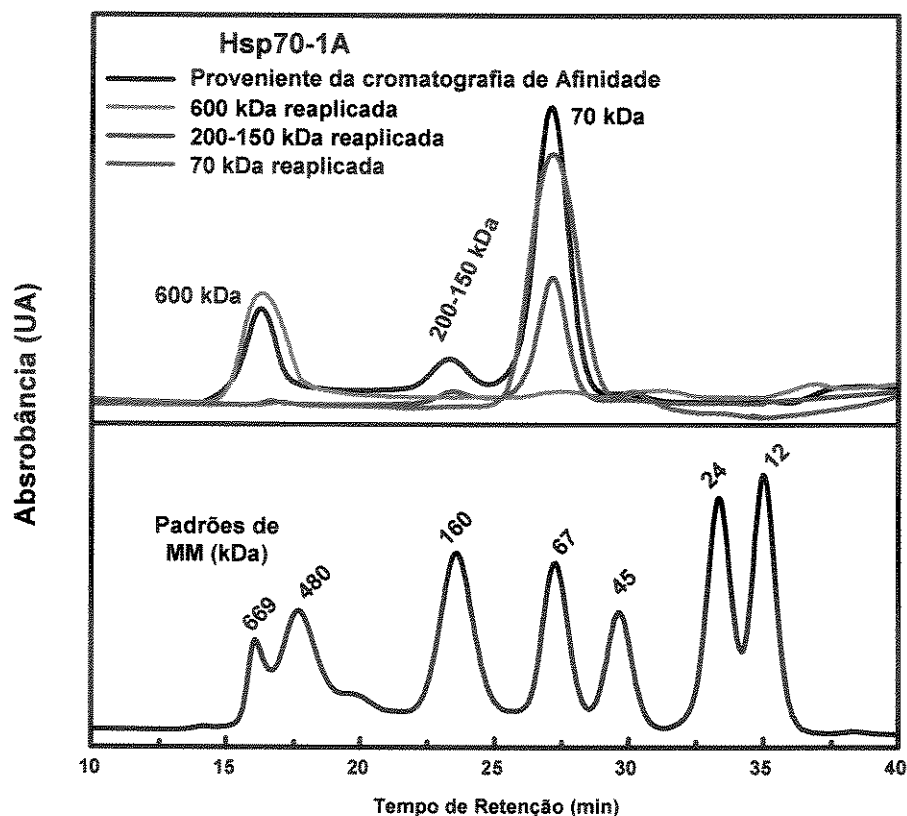


Figura 7: Cromatografia de exclusão molecular analítica para a proteína Hsp70-1A. O processo cromatográfico foi realizado em uma coluna *Superdex 200 HR 10/30* acoplada ao sistema *ÅKTA FPLC*. Cromatografia realizada em tampão Tris-HCl 25 mmol/L (pH 7,5), contendo NaCl 150 mmol/L e DTT 1 mmol/L, e em fluxo constante de 0,5 mL/min. a) Amostras de Hsp70-1A provenientes da cromatografia de afinidade e após diálise contra tampão Tris-HCl 25 mmol/L (pH 7,5), contendo NaCl 150 mmol/L e DTT 1 mmol/L. A estimativa da *MM* foi feita pela comparação do tempo de retenção de proteínas cujas *MM* são conhecidas (painel inferior). Também estão representados o perfil cromatográfico referente a reaplicação das frações correspondentes à *MM* de 600 kDa, 200-150 kDa e 70 kDa. A fração de 600 kDa manteve o mesmo perfil de eluição, enquanto a fração de 200-150 kDa apresentou um estado de equilíbrio com a fração de 70 kDa. A fração de 70 kDa quando reaplicada à CEM continuou apresentando um perfil de eluição de uma espécie monomérica de mesma *MM*.

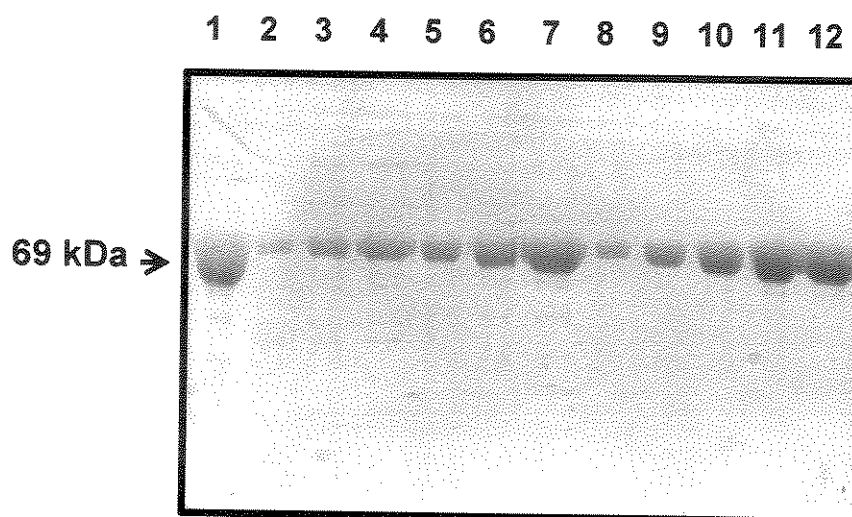


Figura 8: Análise das frações da Hsp70-1A eluídas pela CEM preparativa por SDS-PAGE. As frações obtidas na CEM preparativa (perfil similar ao representado na figura 7) foram aplicadas em concentrações crescentes em uma SDS-PAGE 10%. Canaleta 1: marcador de *MM* (BSA); canaleta 2-4: fração de 600 kDa com aproximadamente 0,5, 1,0 e 1,5 μg de proteína, respectivamente; canaletas 5-7: fração 200-150 kDa com quantidades aproximadas de 1,0, 2,0 e 3,0 μg de proteína, respectivamente; canaletas 8-12: fração de 70 kDa: aproximadamente 0,5, 1,0, 2,0, 3,0, 4,0 e 5,0 μg de proteína, respectivamente (veja item 4.1.4 de Resultados). Hsp70-1A foi purificada com mais de 95% de pureza e não sofreu processo de degradação.

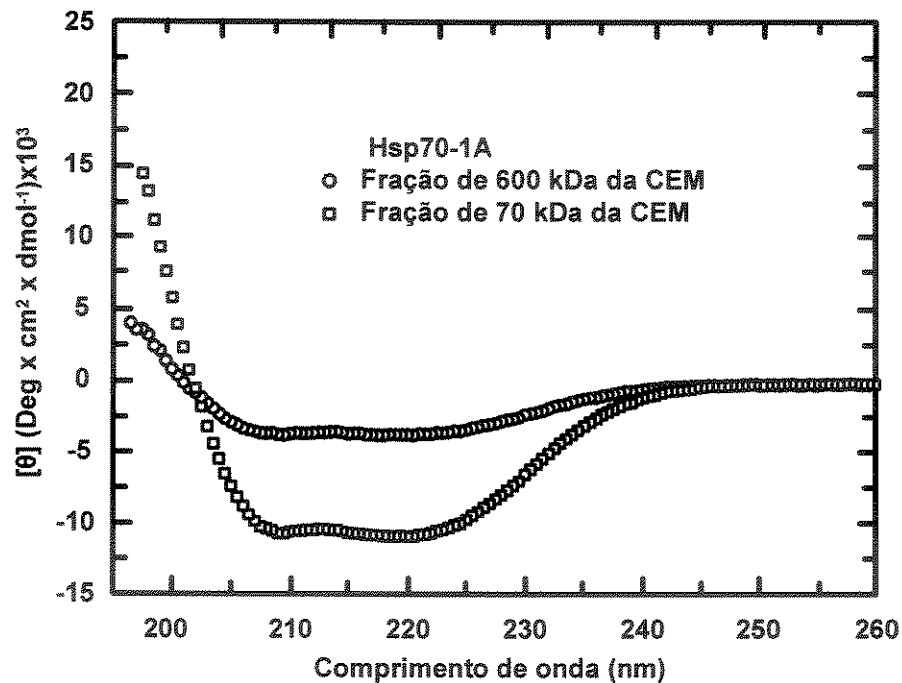


Figura 9: Experimentos de CD das frações da proteína Hsp70-1A purificadas pela CEM preparativa. Os espectros de CD das frações de 600 kDa e 70 kDa da Hsp70-1A, obtidas da CEM, foram coletadas entre 260-195 nm (tampão Tris-HCl 25 mmol/L, pH 7,5, contendo NaCl 150 mmol/L e DTT 1 mmol/L). A fração de 200-150 kDa apresentou um espectro (não mostrado) com o mesmo perfil daquele observado para a fração de 70 kDa da Hsp70-1A. Os dados sugerem fração de 600 kDa da Hsp70-1A apresentou um conteúdo de estrutura secundária diferente do que a fração de 70 kDa.

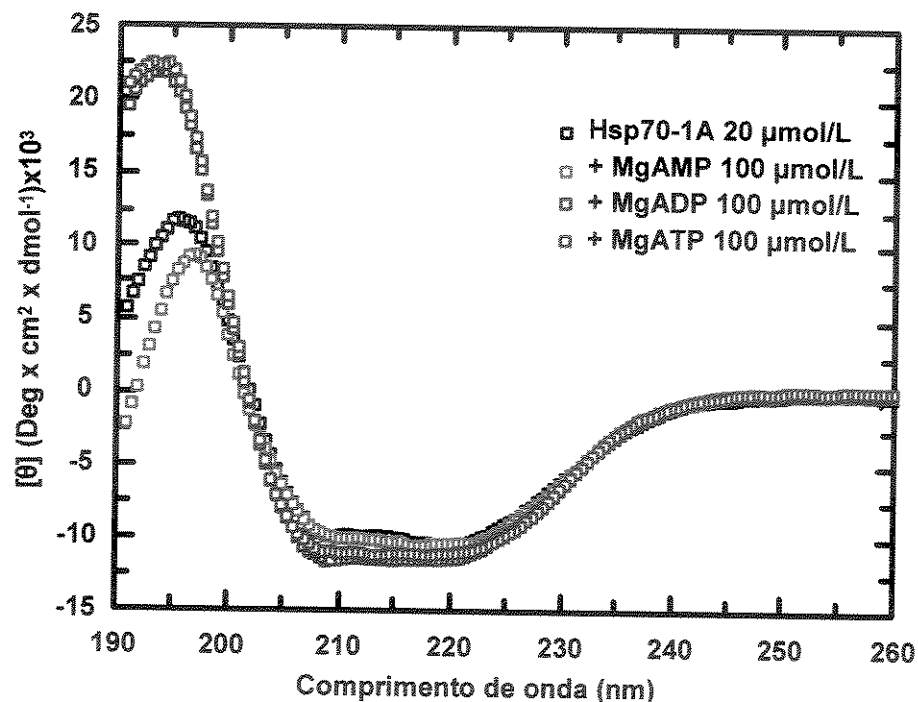


Figura 10: Experimentos de CD da proteína Hsp70-1A na ausência e na presença dos nucleotídeos adenosina. Os espectros de CD da Hsp70-1A (20 $\mu\text{mol/L}$) foram coletados entre os comprimentos de onda de 260 a 190 nm, no tampão TKP a 20 °C e em célula de quartzo de 0,1 mm. Os experimentos de CD da Hsp70-1A foram feitos na ausência e presença de nucleotídeos adenosina (AMP – verde, ADP – vermelho e ATP – azul), em uma concentração final de 100 $\mu\text{mol/L}$ na presença de 0,5 mmol/L de MgCl_2 . A quantidade de estrutura secundária para a Hsp70-1A na ausência de nucleotídeos, estimada pelo programa CDNN Deconvolution foi: 30% de α -hélices, 18% de folha β -pregueada, 17% de alças e 34% de estrutura randômica (erro de 10 %). Nas amostras contendo MgATP ou MgADP, a quantidade de estrutura estimada pelo programa foi aproximadamente 36% de α -hélices, 15% de folha β -pregueada, 16% de alças e 31% de estrutura randômica (erro < 10%).

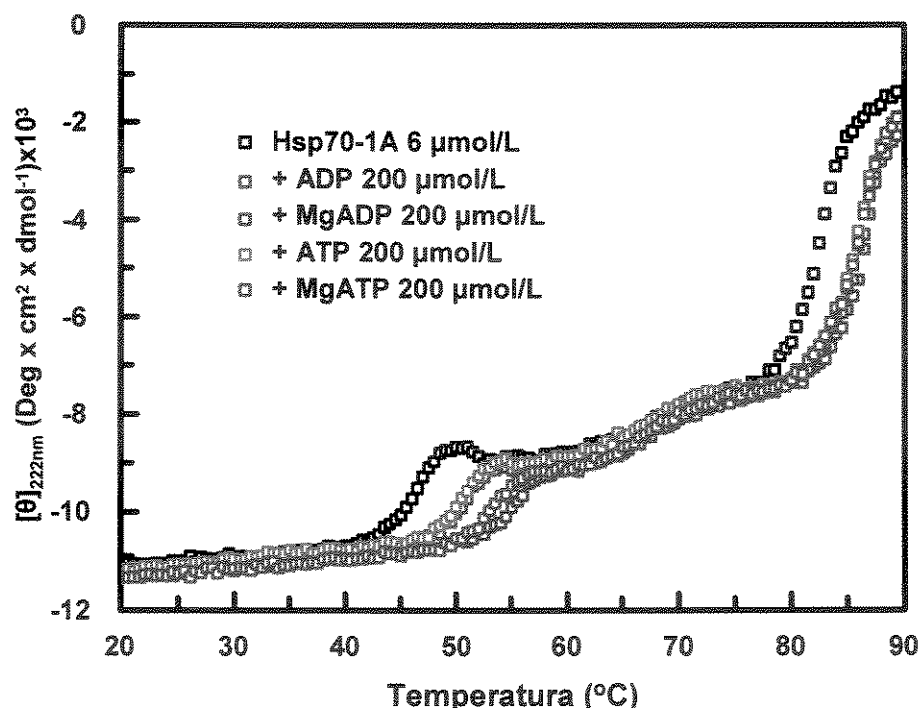


Figura 11: Ensaio de desnaturação térmica de 20 a 90 °C da Hsp70-1A na ausência e na presença de nucleotídeos adenosina. A proteína Hsp70-1A (6 µmol/L), preparada no tampão TKP, foi aquecida a 60 °C/hora, e sua desnaturação foi monitorada por CD no λ de 222 nm. Os ensaios foram realizados na ausência e na presença de 200 µmol/L de nucleotídeos adenosina (com e sem $MgCl_2$) e em cubeta de quartzo de 1 mm. Os dados apresentados nas curvas acima representam a média de 3 curvas de desenovelamento. A Hsp70-1A apresentou 3 Tms em aproximadamente 46, 68 e 82 °C. A desnaturação térmica da Hsp70-1A nestas condições foi irreversível mesmo se aquecida até 50 °C (dados não mostrados), sendo que a proteína sofreu agregação seguida de precipitação quando aquecida até 90 °C. A curva de desnaturação da Hsp70-1A na presença de AMP apresentou o mesmo perfil da curva referente à proteína sem aditivos (não mostrado). Porém, a presença de ADP ou ATP provocou um aumento de aproximadamente 4 °C na Tm_1 e Tm_3 . A presença de $MgCl_2$ não provocou mudanças no perfil da curva de desenovelamento da Hsp70-1A (não mostrado). Mas na presença de MgADP ou MgATP, foram observados aumentos no Tm_1 de 9 e 7 °C, respectivamente, sem mudanças no Tm_3 em comparação às amostras com apenas ADP ou ATP. Estes dados são apresentados na tabela II.

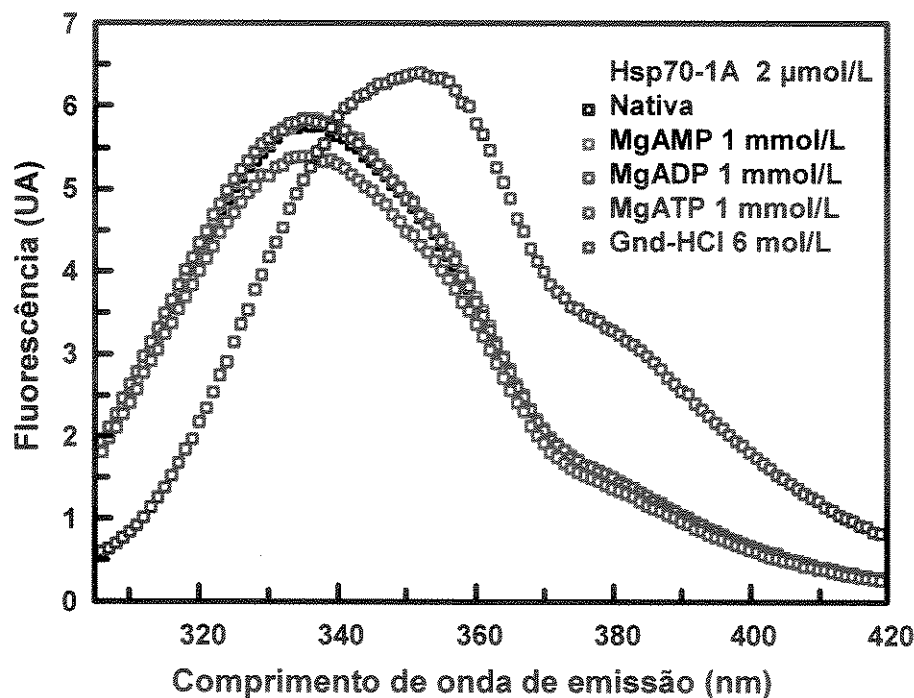


Figura 12: Experimentos de fluorescência da Hsp70-1A na ausência e na presença de nucleotídeos adenosina. Os espectros de emissão de fluorescência da Hsp70-1A (2 $\mu\text{mol/L}$ no tampão TKP) foram coletados em uma célula de quartzo de 1 cm x 1 cm e temperatura constante de 20 °C. Os experimentos foram realizados na ausência e na presença de 1 mmol/L de MgAMP, MgADP e MgATP e em 6 mol/L de Gnd-HCl. O λ de excitação utilizado foi de 295 nm para selecionar os resíduos de triptofano e o λ de emissão máximo de fluorescência em 335 nm (60% do total), com o equipamento configurado para a Hsp70-1A sem aditivos. O espectro de emissão de fluorescência da Hsp70-1A mostra um máximo de fluorescência em 336 nm e um “ombro” em 355 nm. A presença de Gnd-HCl deslocou o λ máximo de emissão de fluorescência da Hsp70-1A para 356 nm. A presença de MgAMP ou MgADP não provocou mudanças no espectro de emissão de fluorescência da Hsp70-1A. Mas, o MgATP reduziu a intensidade de emissão de fluorescência em aproximadamente 7% da área da integral da curva. O $\langle\lambda\rangle$ dos espectros de emissão de fluorescência da Hsp70-1a na presença e ausência de nucleotídeos adenosina foi em 345 nm. Não foram observadas mudanças no perfil dos espectros de emissão de fluorescência devido à ausência de MgCl_2 .

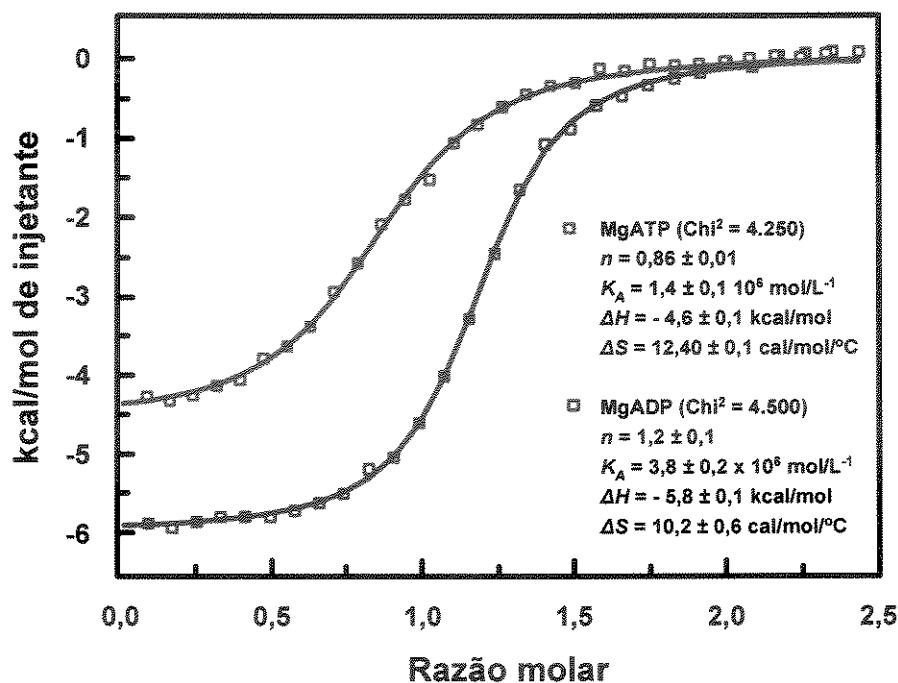


Figura 13: Determinação da K_A da Hsp70-1A por seus ligantes: MgATP ou MgADP. Os ensaios de ITC foram realizados para determinar as constantes termodinâmicas da interação entre a Hsp70-1A e seus ligantes: o MgATP e o MgADP. A amostra de Hsp70-1A (13-15 $\mu\text{mol/L}$ em tampão TKP) foi acondicionada na célula de 1,4 mL do microcalorímetro e alíquotas de 5 μL do titulante (MgATP ou MgADP preparadas no tampão TKP; entre 300-350 $\mu\text{mol/L}$) foram adicionados em intervalos de 6 minutos. O calor, em kcal/mol de injetante, envolvido na interação entre a Hsp70-1A e o respectivo ligante, foi integrado em função do tempo e avaliado em função da razão molar [nucleotídeo adenosina]/[Hsp70-1A]. A linha representa o melhor ajuste da curva de titulação, realizado por uma regressão linear, considerando um sítio de ligação do nucleotídeo ligante por molécula de Hsp70-1A. Os dados obtidos para a K_A mostram que a Hsp70-1A, nas condições testadas, tem maior afinidade por MgADP do que por MgATP.

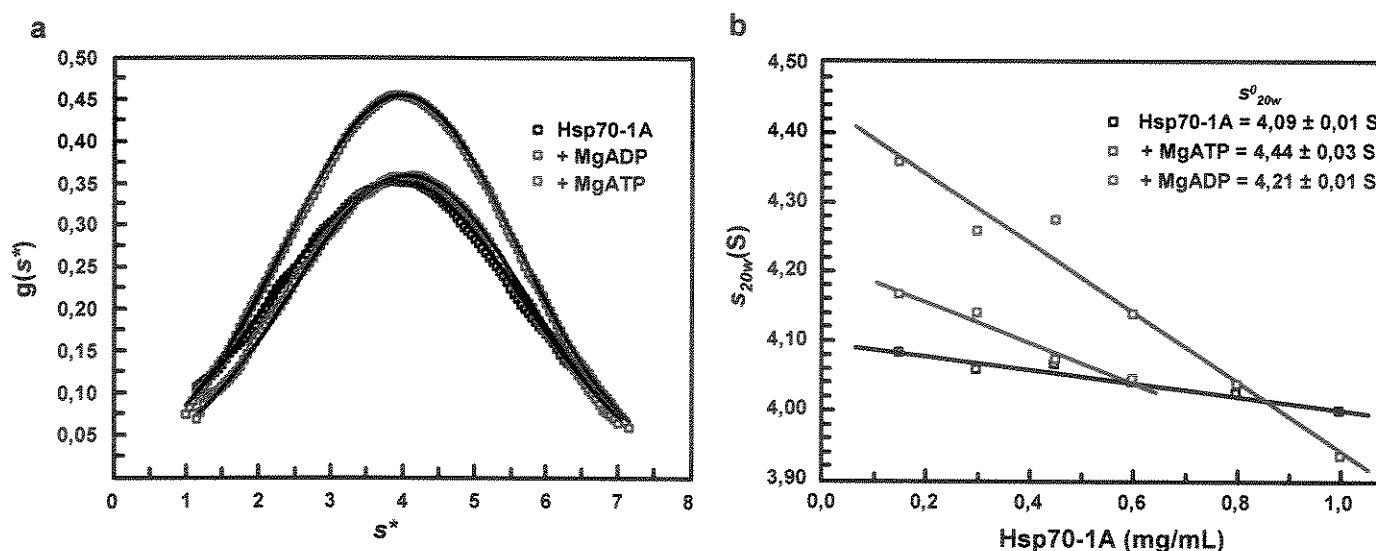


Figura 14: Experimentos de velocidade de sedimentação da proteína Hsp70-1A na ausência e na presença de MgADP e MgATP. Os experimentos de velocidade de sedimentação foram realizados a 20 °C e em velocidade constante de 25.000 rpm (rotor AN-60Ti). A monitoração da absorbância *versus* o raio foi feita em λ de 230 e 236 nm, para condições de baixa e alta concentração de proteína, respectivamente. a) conjunto de dados com a Hsp70-1A a 0,6 mg/mL em tampão TKP na ausência e na presença de nucleotídeos adenosina (MgATP e MgADP 200 μ mol/L). A distribuição da integral de $g(s^*)$ foi tratada utilizando o programa Origin® (Microcal) como um sistema de 1 gaussiana resultando no coeficiente de sedimentação aparente (s^*). Todas as análises foram feitas com o mesmo número de curvas obtidas (30 curvas). b) Estimativa do $s_{20,w}^0$. Gráfico dos valores de $s_{20,w}$ *versus* a concentração de Hsp70-1A mostrando a extrapolação, por regressão linear (programa Origin®), da curva em 0 mg/mL de concentração de proteína (Laue, 2001). Os últimos pontos das análises da Hsp70-1A na presença de MgADP foram descartados, devido ao ruído provocado pela presença da alta concentração protéica e também de nucleotídeos que inviabilizaram as análises.

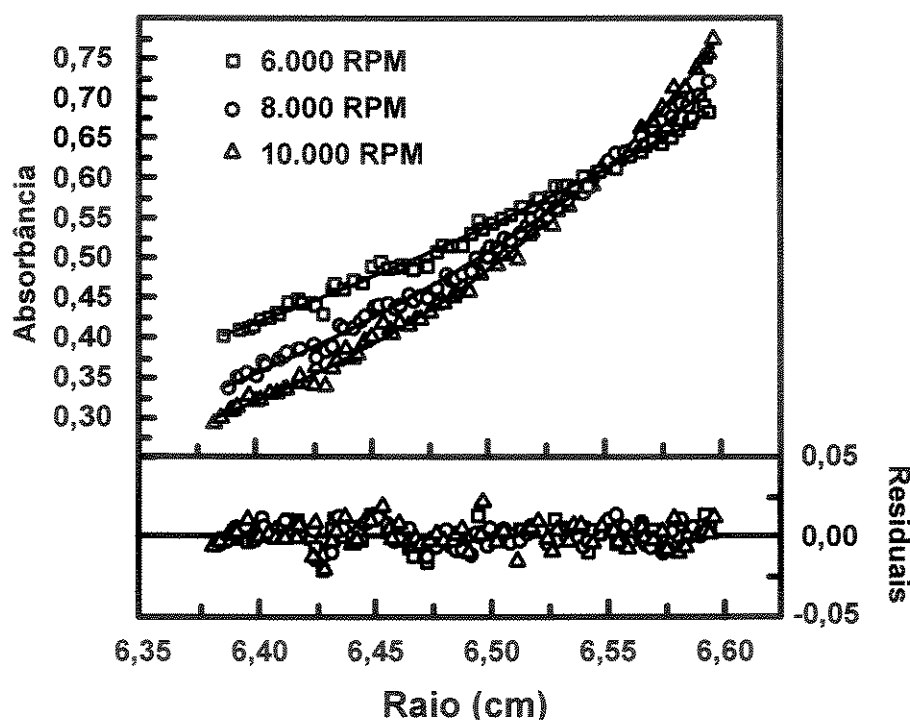


Figura 15: Experimentos de sedimentação em equilíbrio da proteína Hsp70-1A. Os experimentos de sedimentação em equilíbrio foram realizados em velocidades de 6.000, 8.000, e 10.000 rpm (rotor AN-60Ti) em temperatura constante de 20 °C. A coleta de dados de absorbância, em função do raio, foi realizada a 238 nm e em 3 diferentes concentrações de proteína: 100, 250 e 500 µg/mL. A figura representa as curvas experimentais e a interpretação da curva na concentração de 250 µg/mL de Hsp70-1A (veja item 4.1.9 de Resultados). A distribuição randômica dos residuais (painel inferior) mostra a qualidade do ajuste dos dados. Os dados são concordantes com o modelo da Hsp70-1A como um monômero com *MM* de 70 ± 2 kDa.

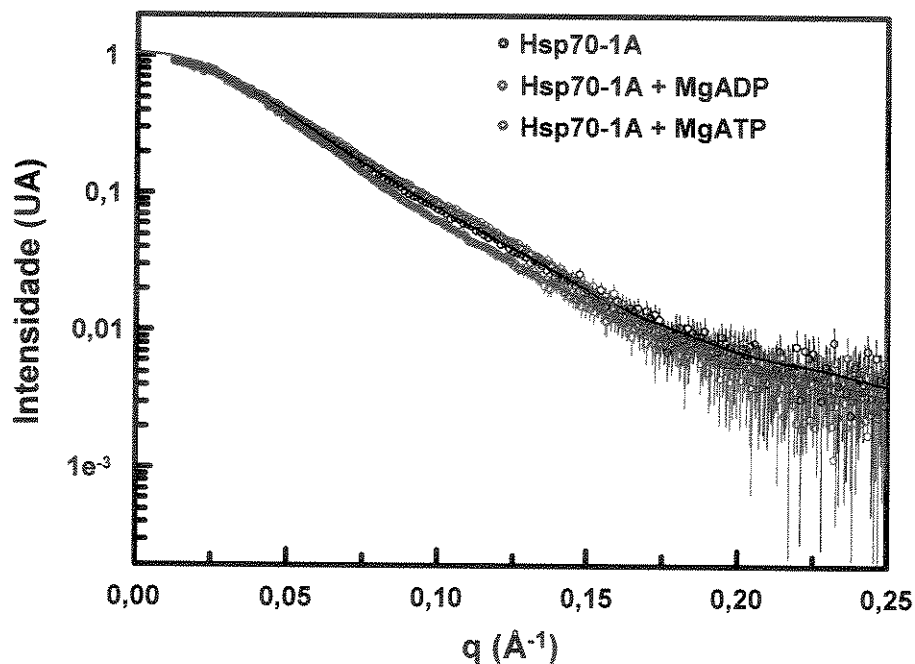


Figura 16: Curvas das $I(q)$ medidas por SAXS para a proteína Hsp70-1A na ausência e na presença de nucleotídeos adenosina. Os dados de SAXS foram coletados na linha SAS do LNLS em várias concentrações da proteína Hsp70-1A no tampão TKP (3,50, 6,70, 10,80 e 15,60 mg/mL) na presença e na ausência de MgATP ou MgADP 1 mmol/L. Após o tratamento dos dados e a normalização pela concentração, as curvas foram unidas para se obter dados de qualidade estatística em um grande intervalo de “ q ”. A linha contínua representa o ajuste realizado pelo programa GNOM e as curvas são apresentadas na forma normalizada pela intensidade espalhada para facilitar a comparação entre as mesmas. A presença de MgATP ou MgADP provocaram pequenas mudanças no perfil da curva de espalhamento que podem ser mais bem avaliadas pela análise das curvas das $p(r)$ resultante apresentadas adiante.

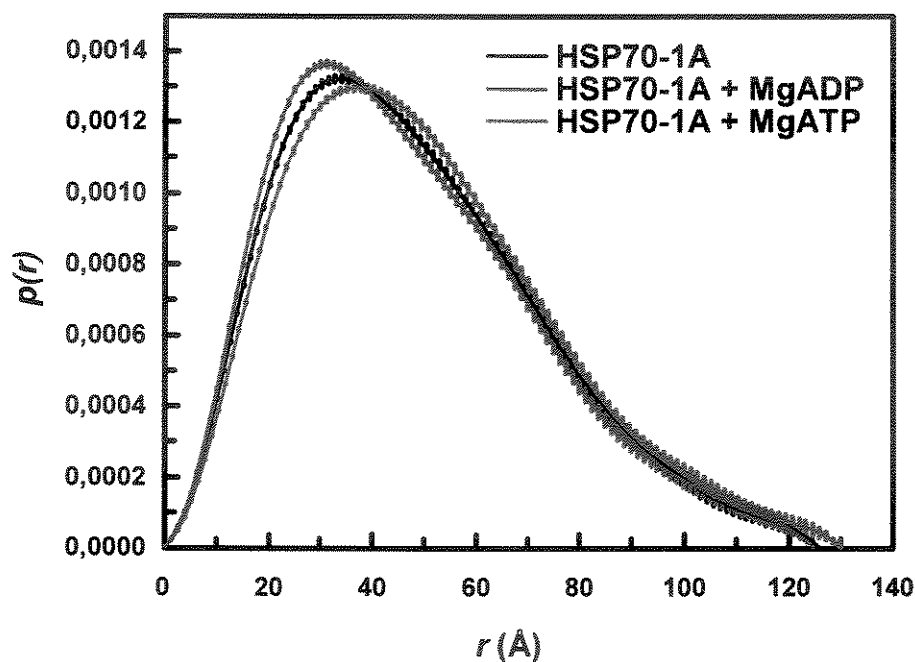


Figura 17: Análise das curvas das $p(r)$ da Hsp70-1A na ausência e na presença de nucleotídeos adenosina. As curvas das $p(r)$ apresentadas são resultantes das curvas das $I(q)$ calculadas pela transformada de Fourier indireta executada pelo programa GNOM. As curvas das $p(r)$ mostram que a Hsp70-1A possui um D_{max} de aproximadamente 125-130 Å. O perfil das curvas indica que a proteína Hsp70-1A possui uma forma prolata alongada nas diferentes condições experimentais. A presença de MgATP ou MgADP ocasionou pequenas mudanças conformacionais na Hsp70-1A avaliadas pelas modificações no perfil das curvas das $p(r)$.

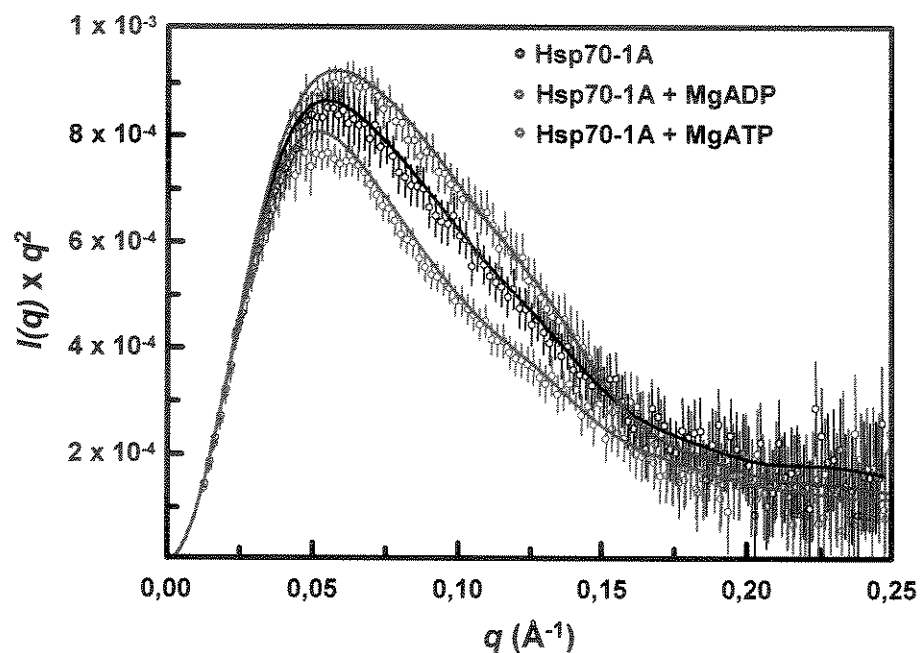


Figura 18: Gráfico de Kratky obtidos das curvas das $I(q)$ da Hsp70-1A medidas por SAXS. O perfil do gráfico de Kratky da Hsp70-1A, no tampão TKP e sem aditivos, sugere que esta partícula é relativamente compacta. A presença de um ombro na curva que pode indicar que a Hsp70-1A é uma proteína bi-lobular. Na presença de MgATP ou MgADP, o perfil do gráfico de Kratky foi similar ao da proteína sem aditivos, sugerindo que o grau de compactação da Hsp70-1A foi similar nas diferentes condições experimentais, ou seja, sem apresentar segmentos com alta flexibilidade.

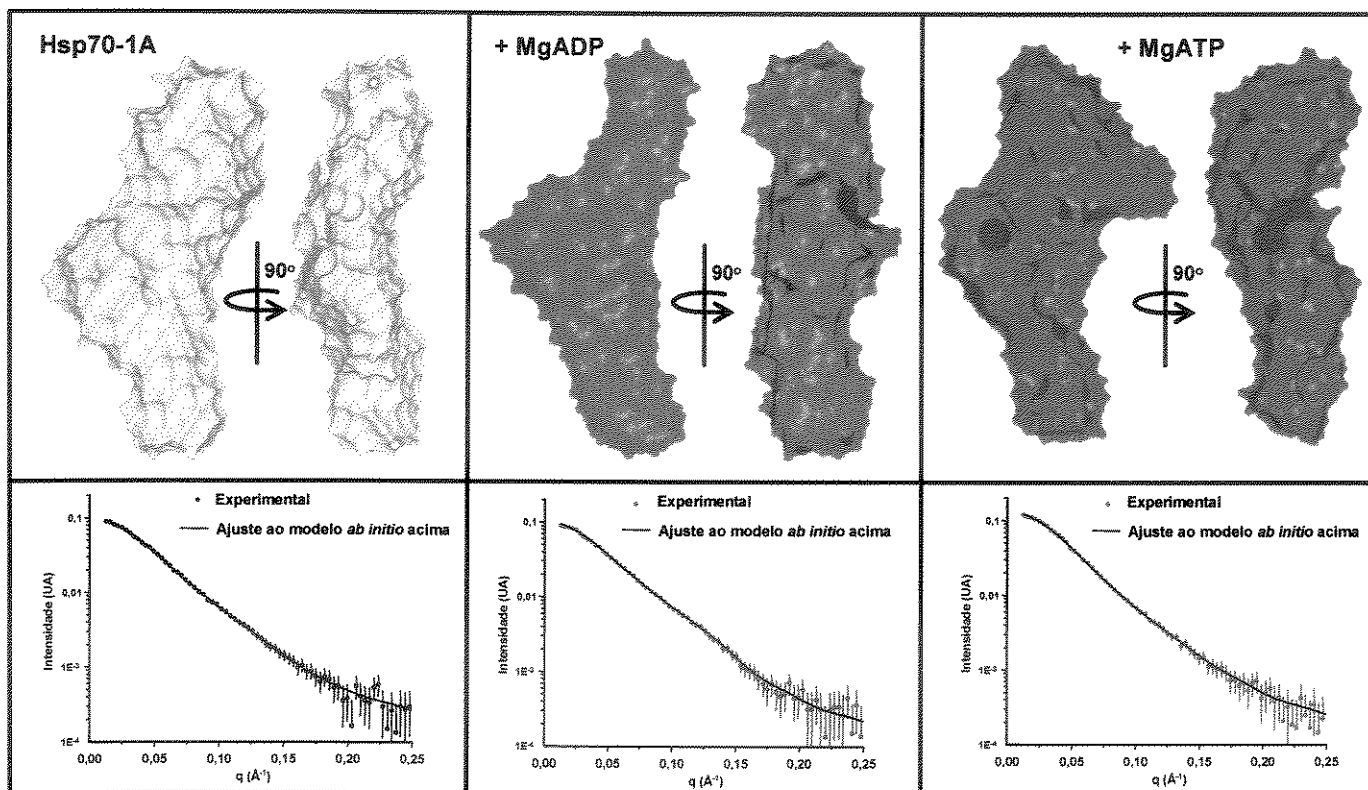


Figura 19: Modelos *ab initio* gerados para a proteína Hsp70-1A a partir dos dados de SAXS nas diferentes condições experimentais. Os modelos *ab initio* obtidos para a Hsp70-1A foram construídos pelo programa GASBOR e representam a média da superposição de 10 cálculos independentes utilizando o programa DAMAVER. Os modelos foram construídos para as 3 condições experimentais dos dados de SAXS da Hsp70-1A no tampão TKP e são apresentados pelas cores: amarelo: sem aditivos; vermelho: na presença de MgADP; e azul: na presença de MgATP. Os modelos são apresentados em duas posições em uma rotação de 90 graus no eixo axial. No painel inferior de cada modelo está apresentado a curva da $I(q)$ experimental e o ajuste com a curva da $I(q)$ do modelo *ab initio* apresentado acima. Os modelos sugerem que a Hsp70-1A possui um formato similar nas 3 condições, porém pequenas mudanças no preenchimento do espaço podem ser observadas, principalmente no modelo *ab initio* na presença de MgATP em relação aos demais.

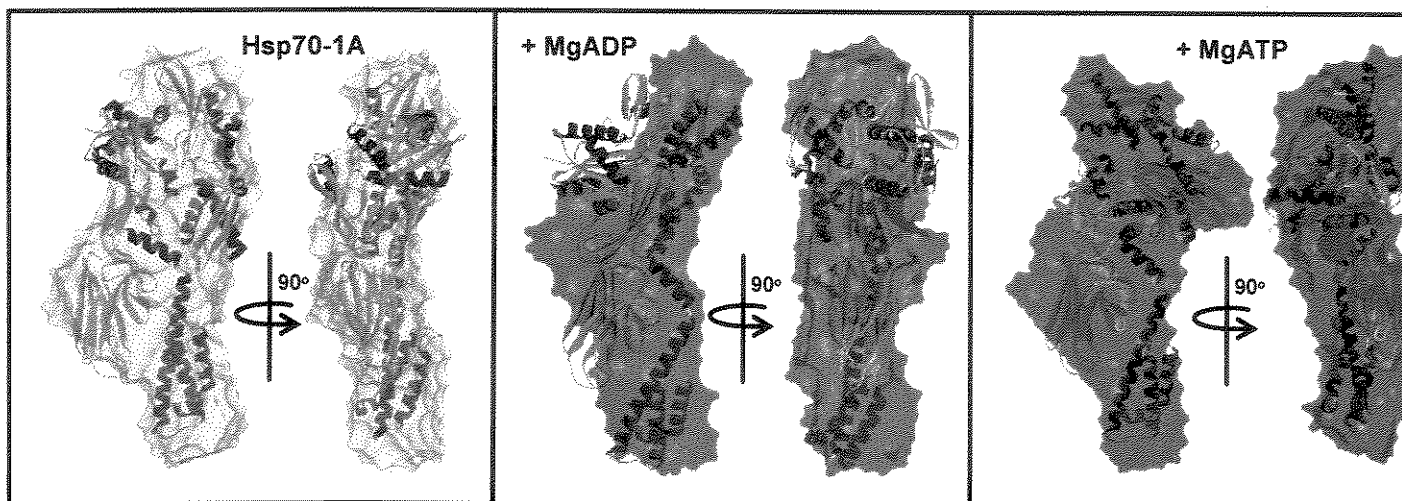


Figura 20: Superposição das estruturas em alta resolução determinadas para os domínios das Hsp70 com os modelos *ab initio* gerados para a proteína Hsp70-1A. O processo de superposição foi feito de maneira manual com o auxílio do programa WebLab ViewerLite (www.accelrys.com). O modelo em alta resolução do domínio DLN (código PDB número 1HJO – representado na parte superior da superposição) representa a estrutura de uma proteína que apresenta 100% de identidade na sequência de aminoácidos com a Hsp70-1A (entre os aminoácidos 3 a 382 da Hsp70-1A). O modelo em alta resolução do DLS foi proveniente da DnaK de *E. coli* (código PDB número 1DKX – representado na parte inferior da superposição). A superposição daqueles domínios no modelo *ab initio* da Hsp70-1A sem aditivos foi razoável (amarelo), como pode ser notado pela rotação de 90 graus no eixo axial da superposição. Para o modelo *ab initio* da Hsp70-1A com MgADP (vermelho) a superposição não foi adequada, pois houve a “sobra” da estrutura em alta resolução em relação ao envelope do modelo *ab initio*. A superposição do modelo *ab initio* para as condições da Hsp70-1A na presença de MgATP (azul) com os domínios em alta resolução foi adequada, como pode ser notado pela rotação de 90 graus da superposição. Os modelos de superposição sugerem que não ocorreu uma grande mudança conformacional, porém sugerem que a ligação do nucleotídeo induziu um rearranjo conformacional sem modificar o formato geral da Hsp70-1A.

4.1.2 Mortalina

A mortalina (mtHsp75 ou Grp75) é uma proteína mitocondrial que pertence à subfamília DnaK e apresenta 65% e 75% de similaridade, na sequência de aminoácidos, com a Hsp70-1A humana e a DnaK de *E. coli*, respectivamente (Figura 4). A mortalina apresenta em sua estrutura primária apenas um resíduo de triptofano na mesma posição na sequência de aminoácidos do que o resíduo de triptofano que a DnaK apresenta.

Pesquisa em banco de dados e obtenção do cDNA plasmidial da mortalina

A sequência de cDNA da proteína mortalina (*Homo sapiens, heat shock 70kD protein 9B (mortalin-2)*, – código GenBank número BC000478 – gene HSPA9B) foi utilizada para buscar e identificar seqüências EST similares no banco de dados GenBank divisão Human EST. Foram identificados dois clones EST (BE901434 e BE791812; Tabela I) que apresentavam similaridade com a mortalina e estavam disponíveis para a aquisição comercial junto ao sistema *IMAGE Consortium*. O DNA plasmidial do clone EST número BE901434 foi submetido ao seqüenciamento das extremidades internas do SMC do vetor pOTB7 e a análise destas seqüências indicou a presença do cDNA da proteína mortalina (código GenBank número BC000478).

Oligonucleotídeos para a PCR da mortalina

Para o desenho dos oligonucleotídeos foi considerado que a sequência de aminoácidos da mortalina apresenta um peptídeo sinal para a exportação da proteína para a mitocôndria. Considerando que tal peptídeo é clivado e não está na estrutura nativa da proteína, o oligonucleotídeo para a região 5' (5' TATGCATCACATATGATCAAGGGAGC 3') para a PCR da mortalina foi desenhado de modo a eliminar o cDNA que codifica os primeiros 46 resíduos de aminoácidos da proteína, a sequência do peptídeo sinal, como predito pelo programa MITOPROT II 1.0a4. O alinhamento da mortalina com a Hsp70-1A e DnaK, representado pela figura 4, também auxiliou na eliminação da sequência do peptídeo sinal. Para facilitar a clonagem do cDNA da mortalina no sistema pET, o códon referente ao aminoácido alanina 47 (A47) foi substituído pelo códon da metionina (A47M), permitindo utilizar a enzima de restrição *Nde* I para a clonagem no vetor de expressão pET28a. Para a região 3', o oligonucleotídeo foi desenhado para substituir bases de modo a criar um sítio de

restrição para a enzima *BamH* I (5' ATTTCTGGGATCCTTACTGTTTTCC 3') após o códon de terminação natural do cDNA da mortalina.

Clonagem do cDNA da mortalina em pET28a

As células do clone EST número BE901434 (Tabela I) foram crescidas, e após sedimentação, foram utilizadas para extrair o DNA plasmidial (via *QIAprep MiniPrep Kit* – Qiagen) que foi utilizado como molde para a PCR com os oligonucleotídeos descritos acima. A PCR forneceu um produto de aproximadamente 1.900 pb, tamanho esperado para o cDNA da mortalina. Este produto de PCR foi purificado do gel de agarose 1% e utilizado para a clonagem em pGEMT com o auxílio do kit *pGEMT easy*[®] conforme o protocolo do fabricante. Uma colônia de bactérias foi selecionada, pois continha vetores pGEMT que liberaram um fragmento de aproximadamente 1.900 pb pela digestão com as enzimas *Nde* I e *BamH* I. Este fragmento foi utilizado para posterior subclonagem do cDNA da mortalina no vetor pET28a como descrito no item 3.9 de Materiais e métodos.

A figura 21 representa o perfil eletroforético de vetores pET28aMOT, obtidos de uma colônia de bactérias previamente selecionada, que liberaram um fragmento de aproximadamente 1.900 pb a partir da incubação com as enzimas *Nde* I e *BamH* I. A análise das seqüências obtidas pelo seqüenciamento das extremidades internas do SMC daquele vetor pET28aMOT, com os oligonucleotídeos *T7 promoter* e *T7 terminator* (GIBCO BRL), confirmou a clonagem do cDNA da proteína mortalina (código GenBank número BC000478) no vetor pET28a, entre os sítios para *Nde* I e *BamH* I. O seqüenciamento também confirmou a ausência do peptídeo sinal descrito acima e a clonagem do cDNA da mortalina em fase de leitura com a seqüência de DNA da cauda de poli-histidina na extremidade amino-terminal do vetor pET28a. Assim, foi produzido o vetor pET28aMOT capaz de expressar a proteína recombinante mortalina, em fusão à cauda de poli-histidina na região amino-terminal, constituída de 650 resíduos de aminoácidos e *MM* correspondente a 70.794 g/mol.

Expressão e purificação da proteína mortalina

A figura 22 apresenta o cromatograma da purificação da mortalina por CEM, em uma coluna *Superdex 200 HR 10/30*, após a purificação inicial da proteína do sobrenadante do lisado bacteriano em uma cromatografia de afinidade por níquel (não mostrado). A análise do perfil cromatográfico mostra que uma grande quantidade de proteína de (95%) foi eluída em

uma fração que apresentou uma alta *MM* aparente (aproximadamente 600 kDa), como estimado pela análise do *ln* da *MM* versus o tempo de retenção de proteínas conhecidas. A CEM também apresentou a eluição de pequenas quantidades de proteína cujo tempo de retenção sugere possuírem *MM* de aproximadamente 100 kDa e 40 kDa. A figura 23 mostra um experimento de SDS-PAGE 10% referente às etapas de expressão, purificação e avaliação do grau de pureza da mortalina. A fração de alta *MM* aparente da mortalina obtida da CEM apresentou alta pureza e aproximadamente 70 kDa de *MM* (> 95% de pureza) em SDS-PAGE 10%.

Dicroísmo circular para a proteína mortalina

A figura 24 representa o espectro de CD da mortalina obtido após normalização em função da concentração. O perfil do espectro sugere que a proteína obtida possui estrutura secundária do tipo α -hélice. Porém, a estimativa de quantidade de estrutura secundária da mortalina, executada pelo programa CDNN Deconvolution, sugere um baixo conteúdo de estrutura secundária em α -hélice (8%) e alta quantidade de estrutura em folha β -pregueada (40%).

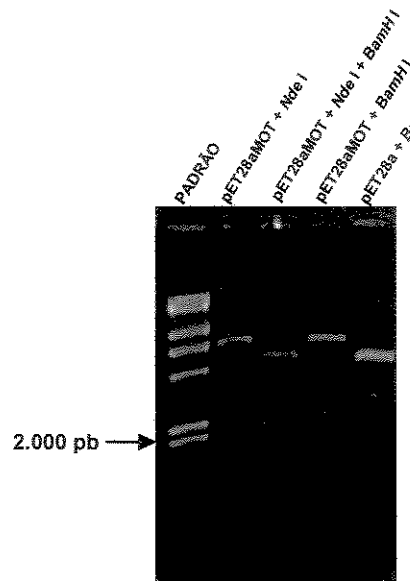


Figura 21: Clonagem do cDNA da mortalina no vetor pET28a. Gel de agarose 1% mostrando o perfil eletroforético da digestão dos vetores pET28aMOT com as enzimas *Nde* I e *Bam*H I. Também está mostrado a digestão de vetores pET28a apenas com a enzima *Bam*H I. A digestão dupla do vetor pET28aMOT levou à liberação de um fragmento de aproximadamente 1.900 pb, como determinado pela comparação com o padrão de *MM Lambda Hind III* (GIBCO BRL). A digestão individual dos vetores pelas enzimas isoladas evidenciou a diferença do tamanho do vetor pET28aMOT em relação ao pET28a. A análise das seqüências do cDNA, obtidas pelo seqüenciamento das extremidades internas do SMC do vetor pET28aMOT com o auxílio dos oligonucleotídeos *T7 promoter* e *T7 terminator*, confirmou a clonagem do cDNA da proteína mortalina (código GenBank número BC000478) sem o peptídeo sinal e em fase de leitura com o cDNA da cauda de poli-histidina do vetor pET28a.

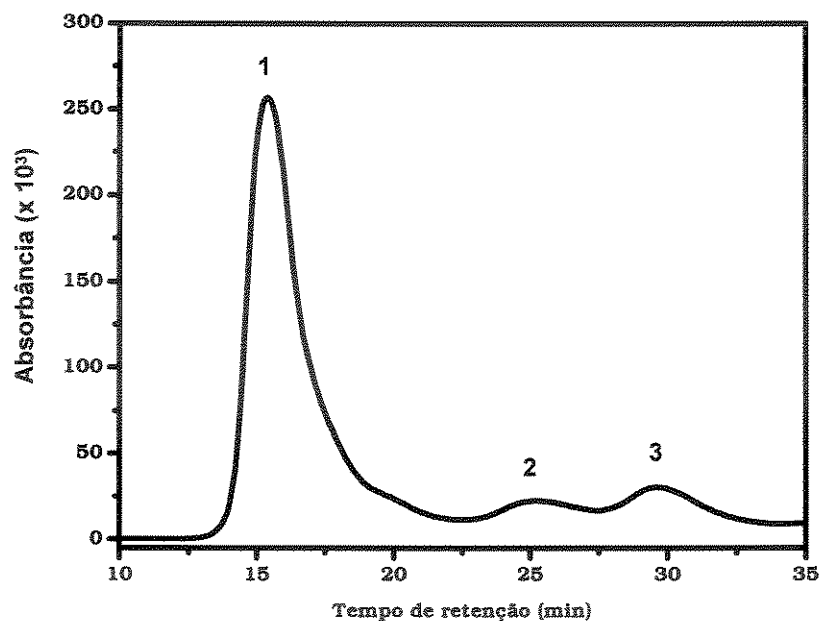


Figura 22: Cromatograma da CEM analítica para a proteína mortalina. A CEM foi realizada em coluna *Superdex 200 HR 10/30* acoplada ao ÄKTA FPLC. A amostra da proteína mortalina foi proveniente da cromatografia de afinidade e após diálise contra tampão Tris-HCl 25 mmol/L (pH 7,5), contendo NaCl 500 mmol/L e β -mercaptoetanol 1 mmol/L. A maior parte da proteína (aproximadamente 95%) foi eluída em um tempo de retenção correspondente a uma partícula de alta *MM* com aproximadamente 600 kDa (1). Também houve eluição de pequenas quantidades de proteína cujos tempos de retenção sugere proteínas de *MM* de aproximadamente 100 kDa (2) e 40 kDa (3), quando comparados com o tempo de retenção de uma mistura de proteínas de *MM* conhecidas (ver figura 7).

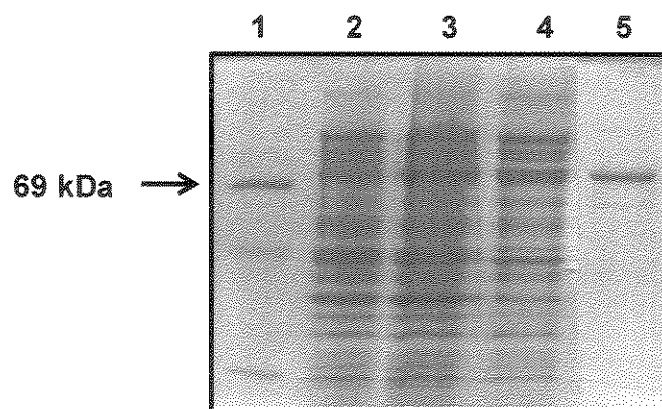


Figura 23: Etapas de expressão e purificação da mortalina analisada por SDS-PAGE.

Bactérias BL21(DE3) contendo o plasmídeo pET28aMOT foram induzidas à expressão pelo IPTG 0,4 mmol/L como descrito no item 3.13 de Materiais e métodos. Após indução por 5 horas, as células foram sedimentadas e lisadas como descrito no item 3.14.1 de Materiais e métodos. A purificação foi realizada através de cromatografia de afinidade por níquel seguida por uma CEM como descrito no item 3.14 de Materiais e métodos. Frações de cada etapa do processo de expressão e purificação foram coletadas e aplicadas em SDS-PAGE 10% juntamente com um padrão de MM. 1: *Protein Molecular Weight Standarts, High Range* (GIBCO BRL). 2: fração de bactérias BL21(DE3) em OD_{600nm} 0,6 sem a presença de IPTG; 3: fração de bactérias BL21(DE3) induzidas com IPTG 0,4 mmol/L por 5 horas; 4: fração do sobrenadante do lisado das bactérias; 5: fração de aproximadamente 600 kDa eluída na cromatografia de exclusão molecular. A proteína mortalina foi expressa em quantidade, apresentando um rendimento de aproximadamente 20 mg/L de indução, e a fração de 600 kDa da CEM da mortalina apresentou grau de pureza superior a 95%.

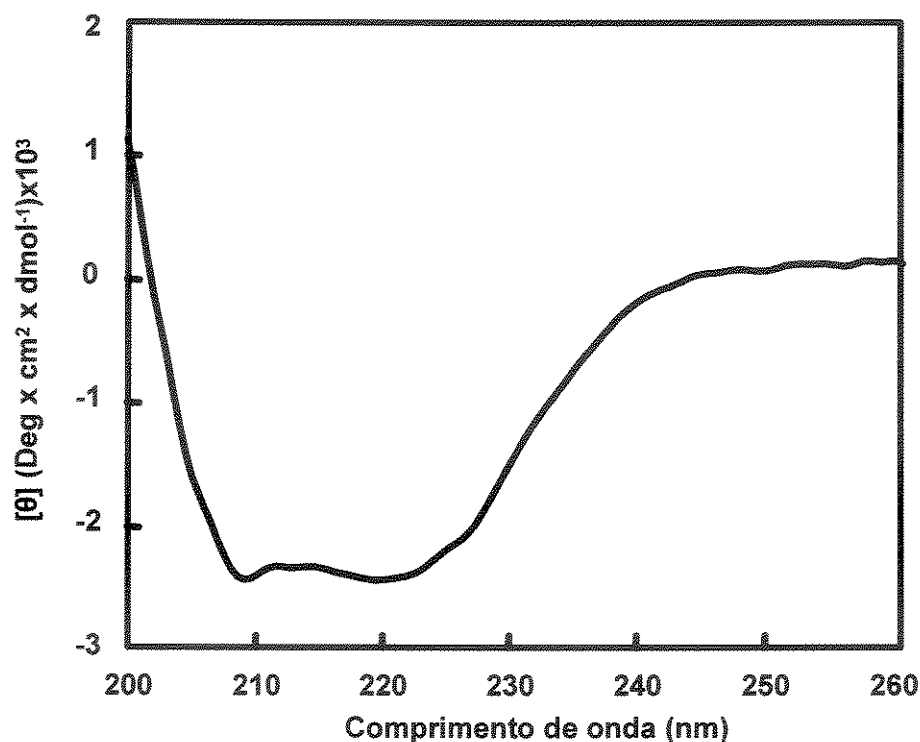


Figura 24: Caracterização da estrutura secundária da mortalina por CD. A amostra da fração de alta *MM* (aproximadamente 600 kDa) da mortalina (em tampão Tris-HCl 25 mmol/L, NaCl 500 mmol/L e β -mercaptoetanol 1 mmol/L – pH 7,5), obtida a partir da CEM, foi submetida à análise por CD. A análise do perfil do espectro de CD sugere que a proteína da fração de alta *MM* da mortalina é constituída de estrutura secundária do tipo de α -hélice. Entretanto, a quantidade de estrutura secundária da mortalina foi estimada pelo programa CDNN Deconvolution e indicou uma baixa quantidade de estrutura em α -hélices (8%) e alta quantidade de folha β -pregueada (40%). Os dados sugerem que a proteína não possui a quantidade de estrutura secundária esperada, principalmente quando comparada a Hsp70-1A (veja item 5.2 da Discussão para detalhes).

4.2 Hsp40 humana

Para os estudos das Hsp40 humanas, foram escolhidas duas proteínas desta família de chaperones. Uma pertencente à subfamília A (DjA1 – gene DNAJA1), que apresenta as principais regiões conservadas das Hsp40 (domínio J, região rica em glicinas e fenilalaninas, domínio rico em cisteínas e o domínio DnaJ_C). A outra proteína escolhida pertence à subfamília B (DjB4 – gene DNAJB4), que apresenta as regiões conservadas descritos acima, exceto aquela rica em cisteínas. Para analisar as propriedades oligoméricas da DjA1, foi produzido um mutante de deleção (DjA1₁₋₃₃₂) da região carboxi-terminal eliminando o sítio de dimerização daquelas proteínas. Abaixo são apresentados os resultados de obtenção do cDNA e a clonagem destes em vetores de expressão, os demais resultados são apresentados no Anexo II (Borges, J.C., Fischer, H., Craievich, A.F. & Ramos, C.H.I. Low-resolution structural study of two human Hsp40 chaperones in solution. DjA1 from subfamily A and DjB4 from subfamily B, have different quaternary structures. *J. Biol. Chem.*, submetido).

4.2.1 Hsp40 DjA1

Pesquisa em banco de dados e obtenção do cDNA plasmidial da Hsp40 DjA1

Utilizando a sequência de cDNA da proteína Hsp40 DjA1 (*Homo sapiens DnaJ (Hsp40) homolog, subfamily A, member 1 (DNAJA1), mRNA* – código GenBank número NM_001539 – gene DNAJA1), foram identificados seqüências ESTs no bando de dados do GenBank e dois clones foram adquiridos do sistema *IMAGE Consortium*. Estes clones estão descritos na tabela I (AW247277 e AI209061).

Oligonucleotídeos para a PCR da Hsp40 DjA1

O oligonucleotídeo 5' desenhado para a PCR da DjA1 substituiu bases na região 5' do cDNA molde (5' CCGGCAGGCTAGCATGGTGAAGAAACAAC 3') para a criação de um sítio de restrição para as enzimas *Nhe* I. Na extremidade 3', o oligonucleotídeo foi desenhado para substituir algumas bases e criar um sítio de restrição para a enzima *Xho* I após 3 bases do códon de terminação (5' TGAGTGTTATTCTCGAGTCATTAAGAGGTCTG 3'). Outros 2 oligonucleotídeos foram desenhados para se anelarem na região interna do cDNA com o

objetivo de obter a sequência interna do cDNA da DjA1 e verificar a veracidade da sequência de nucleotídeos (5' CTAAAGACCGGTGTAAAAGCTGCAACG 3' e 5' GCTTTTACACCGGTCTTTAGGACTGATCC 3').

Clonagem do cDNA da Hsp40 DjA1

As células do clone EST número AW247277 foram crescidas e sedimentadas para a extração do DNA plasmidial com o auxílio do kit *QIAprep MiniPrep Kit* (Qiagen). Este DNA plasmidial foi usado como molde para a PCR com os oligonucleotídeos descritos acima e foi obtido um produto de PCR de aproximadamente 1.200 pb, como analisado por gel de agarose 1% (dados não mostrados). Este produto de PCR foi purificado pelo kit *QiaEx II* (Qiagen) e utilizado para a clonagem no vetor pUC18 como descrito no item 3.9 de Materiais e métodos. Foi selecionada uma colônia contendo vetores pUC18, cuja análise por restrição com as enzimas *Nhe I* e *Xho I*, liberou um fragmento de aproximadamente 1.200 pb. Este fragmento foi utilizado para a subclonagem do cDNA da DjA1 no vetor pET28a como descrito nos item 3.9 de Materiais e métodos.

A figura 25 representa o perfil eletroforético da digestão de vetores pET28aDjA1, obtidos a partir de uma colônia previamente selecionada, que liberaram um fragmento de aproximadamente 1.200 pb, após incubação com as enzimas *Nhe I* e *Xho I*. A análise das sequências obtidas pelo seqüenciamento das extremidades internas do SMC daquele vetor, com o auxílio dos oligonucleotídeos *T7 promoter* e *T7 terminator* e com os oligonucleotídeos desenhados para o anelamento na região interna do cDNA da DjA1, confirmou a presença do cDNA completo para a DjA1 (código GenBank número NM_001539) clonado em fusão com a sequência de cDNA da cauda de poli-histidina. Houve apenas uma mutação sítio específico que provocou a inserção de um códon de terminação na região 3' do cDNA e reduziu o tamanho da proteína a ser expressa em 3 resíduos de aminoácidos. A análise da posição da mutação sugere que esta foi inserida no processo de PCR por um possível erro na síntese do oligonucleotídeo 3'. Devido à região da mutação se localizar na região carboxi-terminal com a eliminação de apenas 3 resíduos de aminoácidos, este vetor foi mesmo assim utilizado para a expressão da proteína DjA1. Deste modo, foi obtido o vetor pET28aDjA1 capaz de expressar a proteína recombinante DjA1, em fusão à cauda de poli-histidina na porção amino-terminal, constituída de 417 resíduos de aminoácidos e *MM* correspondente a 47.000 g/mol.

Outros ensaios realizados para a proteína Hsp40 DjA1

A DjA1 foi expressa em alta quantidade apresentando um rendimento de aproximadamente 30 mg/L de indução e os resultados de purificação, de CD, de UAL, de SAXS e dos ensaios de atividade chaperone da proteína Hsp40 DjA1 estão mostrados no manuscrito apresentado no Anexo II (Borges, J.C., Fischer, H., Craievich, A.F. & Ramos, C.H.I. Low-resolution structural study of two human Hsp40 chaperones in solution. DjA1 from subfamily A and DjB4 from subfamily B, have different quaternary structures. *J. Biol. Chem.*, submetido) e as informações adicionais para os experimentos estão contidas nas legendas das figuras.

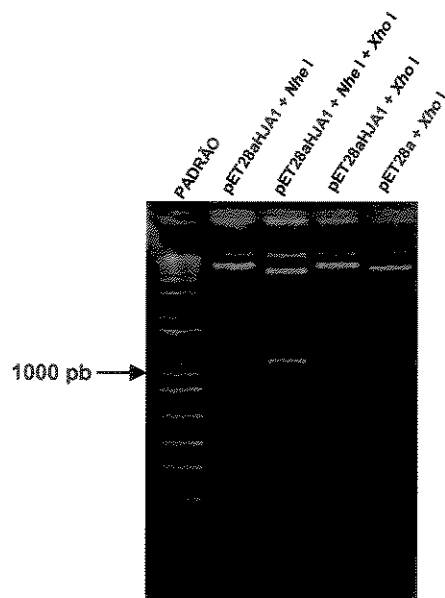


Figura 25: Verificação da clonagem do cDNA da Hsp40 DjA1 no vetor pET28a. Gel de agarose 1% mostrando o perfil eletroforético da digestão de vetores pET28aDjA1 com as enzimas *Nhe* I e *Xho* I. Na digestão dupla com as enzimas, o perfil de bandas evidencia a liberação de um fragmento de aproximadamente 1.200 pb, como determinado pela comparação com o padrão de *MM 1 Kb plus DNA ladder* (GIBCO BRL). A digestão individual dos vetores pelas enzimas evidenciou a diferença no tamanho do vetor pET28aDjA1 em relação ao pET28a. A análise das seqüências do cDNA, obtidas pelo seqüenciamento das extremidades internas do SMC do vetor pET28aDjA1, utilizando os oligonucleotídeos *T7 promoter*, *T7 terminator* e aqueles desenhados para se anelarem no interior do cDNA da DjA1 (descritos acima), confirmou a clonagem do cDNA da proteína Hsp40 DjA1 (código GenBank número NM_001539) em fase de leitura com o cDNA da cauda de poli-histidina do vetor pET28a.

4.2.2 Deleção carboxi-terminal da proteína Hsp40 DjA1 – DjA1₁₋₃₃₂

O DNA plasmidial utilizado para a produção da deleção da região carboxi-terminal da DjA1 foi o vetor pET28aDjA1 descrito no item 4.2.1 de Resultados. Abaixo são apresentados os resultados da construção do mutante e clonagem do cDNA da DjA1₁₋₃₃₂.

Oligonucleotídeos para a PCR da Hsp40 DjA1₁₋₃₃₂

Para criar o mutante de deleção da região carboxi-terminal da DjA1 e estudar as propriedades da forma monomérica da proteína DjA1, foi desenhado um oligonucleotídeo 3' (5' GAGAATGGCTAACTCGAGCCTGATAA 3') para se anelar na região da F333 na região carboxi-terminal da DjA1. Este oligonucleotídeo foi desenhado de forma a substituir o códon F333 e criar um códon de terminação no códon e também criar um sítio de restrição para a enzima *Xho* I após o códon F333. Tal deleção foi desenhada para manter a sequência total do domínio DnaJ_C da DjA1. O outro oligonucleotídeo utilizado para a PCR para a produção da DjA1₁₋₃₃₂ foi o oligonucleotídeo *T7 promoter* que se anela na região anterior ao sítio da enzima *Nhe* I do vetor pET28aDjA1 (descrito no item 4.2.1 de Resultados).

Clonagem do cDNA da proteína Hsp40 DjA1₁₋₃₃₂

O vetor pET28aDjA1 foi utilizado como molde para a PCR para a produção do cDNA da deleção da região carboxi-terminal da DjA1. Através da utilização dos oligonucleotídeos descritos acima, foi obtido um produto de PCR de aproximadamente 1.000 pb que foi purificado do gel de agarose 1% e submetido à digestão com as enzimas *Nhe* I e *Xho* I. Este fragmento foi submetido à reação de ligação ao vetor pET28a previamente preparado com as mesmas enzimas descritas acima, como descrito no item 3.9 de Materiais e métodos.

A figura 26 representa o perfil eletroforético da digestão enzimática de vetores pET28aDjA1 e pET28aDjA1₁₋₃₃₂, sendo que este último foi obtido de uma colônia de bactérias previamente selecionada do procedimento de clonagem. A digestão do vetor pET28aDjA1₁₋₃₃₂, com as enzimas *Nhe* I e *Xho* I, ocasionou a liberação de um fragmento de aproximadamente 1.000 pb. Este fragmento foi aproximadamente 200 pb menor do que aquele liberado pela digestão do vetor pET28aDjA1 com as mesmas enzimas. O seqüenciamento do DNA das extremidades internas do SMC do vetor pET28a, com os oligonucleotídeos *T7 promoter* e *T7 terminator*, confirmou a presença do cDNA da DjA1 clonado em fase de leitura com o cDNA da cauda de poli-histidina do pET28a, porém sem a

seqüência do cDNA correspondente aos últimos 65 resíduos de aminoácidos da região carboxi-terminal da DjA1. Deste modo foi obtido o vetor pET28aDjA1₁₋₃₃₂ capaz de expressar a proteína recombinante DjA1₁₋₃₃₂, em fusão à cauda de poli-histidina na região amino-terminal, composta por 354 resíduos de aminoácidos e *MM* igual a 39.552 g/mol.

Outros ensaios realizados para a proteína Hsp40 DjA1₁₋₃₃₂

A DjA1₁₋₃₃₂ foi expressa em alta quantidade apresentando um rendimento de aproximadamente 20 mg/L de indução e os resultados de purificação, de CD, de UAL, de SAXS e dos ensaios de atividade chaperone da proteína Hsp40 DjA1₁₋₃₃₂ estão mostrados no manuscrito apresentado no Anexo II (Borges, J.C., Fischer, H., Craievich, A.F. & Ramos, C.H.I. Low-resolution structural study of two human Hsp40 chaperones in solution. DjA1 from subfamily A and DjB4 from subfamily B, have different quaternary structures. *J. Biol. Chem.*, submetido) e as informações adicionais para os experimentos estão contidas nas legendas das figuras.

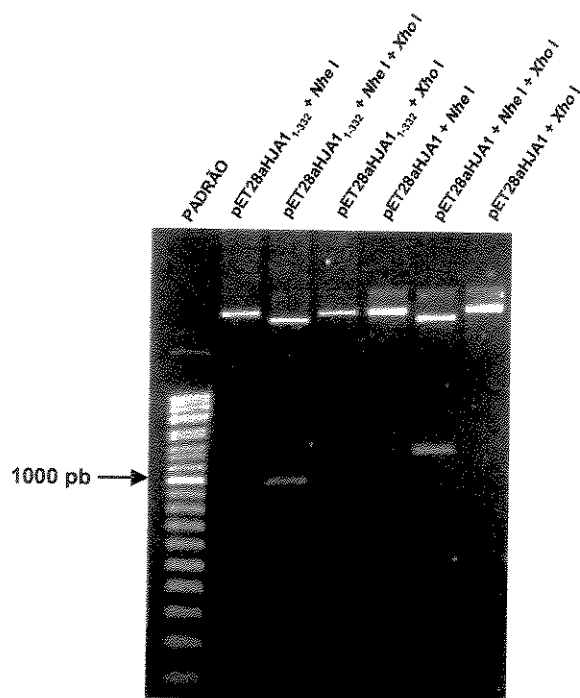


Figura 26: Verificação da clonagem do cDNA da deleção carboxi-terminal da Hsp40 DjA1 no vetor pET28a. Gel de agarose 1% mostrando o perfil eletroforético da digestão dos vetores pET28aDjA1₁₋₃₃₂ e pET28aDjA1 com as enzimas *Nhe* I e *Xho* I. A digestão dupla dos vetores pET28aDjA1₁₋₃₃₂ e pET28aDjA1 com as enzimas descritas levaram à liberação de fragmentos de aproximadamente 1.000 pb e 1.200 pb, respectivamente, determinados pela comparação com o padrão de *MM DNA ladder* (GIBCO BRL). A deleção da região carboxi-terminal da proteína Hsp40 DjA1 foi feita pela eliminação de uma sequência de 200 pb da região 3' do cDNA da DjA1. O vetor pET28aDjA1₁₋₃₃₂ foi submetido ao seqüenciamento das extremidades internas do SMC, com o auxílio dos oligonucleotídeos *T7 promoter* e *T7 terminator*, e a análise das seqüências obtidas confirmou a eliminação dos últimos 200 pb do cDNA da DjA1 em fase de leitura com o cDNA da cauda de poli-histidina do plasmídeo pET28a.

4.2.3 Hsp40 DjB4

Pesquisa em banco de dados e obtenção do cDNA plasmidial da Hsp40 DjB4

A partir da sequência de cDNA da proteína DjB4 (*Homo sapiens DnaJ (Hsp40) homolog, subfamily B, member 4 (DNAJB4), mRNA* – código GenBank número NM_007034 – gene DNAJB4), foram identificados dois clones ESTs no banco de dados GenBank que foram adquiridos do sistema *IMAGE Consortium* (Tabela I – AA081471 e AA464833).

Oligonucleotídeos para a PCR da Hsp40 DjB4

A partir da sequência de nucleotídeos da proteína DjB4 (código GenBank número NM_007034), foram desenhados dois oligonucleotídeos para a região 5' e 3' do cDNA da DjB4. O oligonucleotídeo 5' foi desenhado para substituir algumas bases e assim criar um sítio de restrição para a enzima *Nde* I (5' TCAAGGCATTCCATATGGGGAAAGACTATTA 3') em fase de leitura com o códon de iniciação ATG (metionina) do vetor pET28a. Para a região 3', o oligonucleotídeo desenhado substituiu algumas bases para a criação de um sítio de restrição para a enzima *Bam*H I após 5 bases do códon de terminação natural do cDNA da DjB4 (5' CAAAGTGGATCCTACTATGAGG 3').

Clonagem do cDNA da proteína Hsp40 DjB4

As células do clone EST número AA081471 (Tabela I) foram crescidas em meio líquido e sedimentadas para a extração do DNA plasmidial pelo kit *QIAprep MiniPrep Kit* (Qiagen). Este DNA plasmidial foi utilizado como molde para a PCR com os oligonucleotídeos descritos acima. O produto da PCR obtido apresentou aproximadamente 1.000 pb, mesmo tamanho que o cDNA completo para a proteína DjB4, e foi clonado em pUC18 como descrito no item 3.9 de Materiais e métodos. Uma colônia contendo vetores pUC18a, previamente selecionada, liberou um fragmento de 1.000 pb pela digestão com as enzimas de restrição *Nde* I e *Bam*H I. O seqüenciamento das extremidades internas do SMC daquele vetor confirmou a clonagem do cDNA da DjB4. O fragmento liberado pelas enzimas *Nde* I e *Bam*H I foi utilizado para a subclonagem em pET28a, como descrito no item 3.9 de Materiais e métodos.

A figura 27 representa o perfil eletroforético da digestão de vetores pET28aDjB4, obtidos de uma colônia de bactérias selecionada do procedimento de clonagem, que liberou um fragmento de aproximadamente 1.000 pb pela digestão com as enzimas *Nde* I e *Bam*H I. O vetor pET28aDjB4 foi submetido ao seqüenciamento das extremidades internas do SMC do pET28a, com os oligonucleotídeos *T7 promoter* e *T7 terminator*, e a análise das seqüências obtidas confirmou a presença da seqüência completa do cDNA da DjB4 (código GenBank número NM_007034) clonado em fase de leitura com o cDNA da cauda de poli-histidina do pET28a. Deste modo, foi construído o vetor pET28aDjB4 para a expressão da proteína recombinante DjB4, em fusão à cauda de poli-histidina na extremidade amino-terminal, composta por 357 resíduos de aminoácidos e *MM* igual a 39.970 g/mol.

Outros ensaios realizados para a proteína Hsp40 DjB4

A DjB4 foi expressa em alta quantidade apresentando um rendimento de aproximadamente 30 mg/L de indução e os resultados de purificação, de CD, de UAL, de SAXS e dos ensaios de atividade chaperone da proteína Hsp40 DjB4 estão mostrados no manuscrito apresentado no Anexo II (Borges, J.C., Fischer, H., Craievich, A.F. & Ramos, C.H.I. Low-resolution structural study of two human Hsp40 chaperones in solution. DjA1 from subfamily A and DjB4 from subfamily B, have different quaternary structures. *J. Biol. Chem.*, submetido) e as informações adicionais para os experimentos estão contidas nas legendas das figuras.

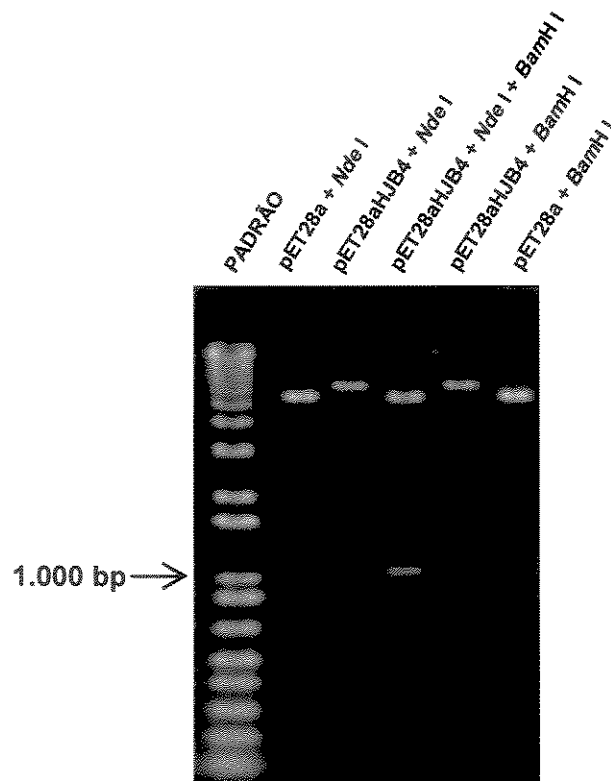


Figura 27: Verificação da clonagem do cDNA da proteína Hsp40 DjB4 no vetor pET28a.

Gel de agarose 1% mostrando o perfil eletroforético da digestão dos vetores pET28aDjB4 e pET28a com as enzimas *Nde* I e *Bam*H I. A digestão dupla do vetor pET28aDjB4 levou à liberação de um fragmento de aproximadamente 1.000 pb, como determinado pela comparação com o padrão de *MM 1 Kb Plus DNA Ladder* (GIBCO BRL). A digestão individual dos vetores com cada enzima evidenciou a diferença de tamanho do vetor pET28aDjB4 linear em relação ao pET28a linear. A análise das seqüências do cDNA, obtidas pelo seqüenciamento das extremidades internas do SMC do vetor pET28aDjB4, utilizando os oligonucleotídeos *T7 promoter* e *T7 terminator*, confirmou a clonagem correta do cDNA da proteína Hsp40 DjB4 (código GenBank número NM_007034) em fase de leitura com o cDNA da cauda de poli-histidina no vetor pET28a.

4.3 GrpE

A busca por seqüências de aminoácidos homólogas às proteínas GrpE resultaram na identificação de duas proteínas humanas preditas, por suas anotações, para serem precursores de proteínas mitocondriais. A seqüência de aminoácidos da EcGrpE foi utilizada para realizar a busca contra o banco de dados do GenBank, através da ferramenta BLASTp. As proteínas humanas identificadas foram: *GrpE protein homolog 1, mitochondrial precursor (Mt-GrpE#1) (HMGE)* (código GenBank número Q9HAV7) e *GrpE protein homolog 2, mitochondrial precursor (Mt-GrpE#2)* (código GenBank número Q8TAA5). Com a publicação do genoma humano em março de 2001, foi possível confirmar a presença de seqüências de DNA genômico que codificam as duas GrpEs humanas. Com o auxílio do programa tBLASTn, foram localizadas seqüências de nucleotídeos que codificam as GrpE humanas em 5 cromossomos diferentes (4, 5, 7, 13 e X).

O gene da GrpE#1 humana (*GrpE-like 1, mitochondrial (E. coli) [Homo sapiens]* – código GenBank número GRPEL1) é codificado pelo cromossomo 4 e possui 4 exons e 3 introns. Também foram encontrados seqüências similares à GrpE#1 no cromossomo 7 (em 3 diferentes regiões deste cromossomo de acordo com o seguinte padrão →←→ considerando a seqüência de nucleotídeos convencionada para o cromossomo 7) provavelmente oriundas de duplicações e/ou inversões. Uma destas seqüências poderia codificar uma terceira proteína similar à GrpE#1, porém esta seria formada pelos 80 primeiros resíduos de aminoácidos, correspondentes às longas hélices, e não apresentaria nem a segunda hélice do feixe de α -hélices e nem o domínio de interação com as Hsp70.

O gene da GrpE#2 humana (*GrpE-like 2, mitochondrial (E. coli) [Homo sapiens]* – código GenBank número GRPEL2) está codificado no cromossomo 5 (4 exons e 3 introns) e também apresenta seqüências com alta identidade (91%) nos cromossomos 7, 13 e X. Porém, estas seqüências apresentam vários códons de terminação na mesma fase de leitura que codifica corretamente a proteína e são idênticas entre si, na comparação na seqüência de aminoácidos, sendo assim consideradas pseudogenes.

As análises da estrutura primária as duas GrpEs humanas indicaram a presença de peptídeos sinais para a exportação das proteínas para a mitocôndria. Estes peptídeos sinais que provavelmente são clivados por uma protease no interior da mitocôndria e não devem estar presentes nas proteínas ativas (Matouschek, 2003). Assim, as GrpEs humanas foram clonadas sem o peptídeo sinal e os resultados obtidos são apresentados abaixo.

4.3.1 GrpE#1

Pesquisa em banco de dados e obtenção do cDNA plasmidial da GrpE#1

A sequência de nucleotídeos do cDNA da GrpE#1 humana (*Homo sapiens GrpE-like 1, mitochondrial (E. coli) (GRPEL1), mRNA* – código GenBank número NM_025196) foi utilizada para buscar e identificar clones ESTs no banco de dados GenBank, oriundos do sistema *IMAGE Consortium*. Dois clones EST contendo o cDNA completo da proteína GrpE#1 foram identificados e adquiridos (Tabela I – BE614754 e BE615384).

Oligonucleotídeos para a PCR da GrpE#1

Para o desenho do oligonucleotídeo 5' para a PCR da GrpE#1, sem o peptídeo sinal de exportação da proteína para a mitocôndria, foi considerado que a sequência correspondente aos primeiros 26 resíduos da sequência de aminoácidos da proteína GrpE#1 humana (código GenBank número Q9HAV7) são referentes ao peptídeo sinal para a exportação da proteína para a mitocôndria. Esta determinação foi feita levando-se em consideração o trabalho de Naylor *et al.* (1998) com a respectiva GrpE#1 de rato e com o auxílio do programa MITOPROT II 1.0a4. Para facilitar a clonagem do cDNA no sistema pET, foram substituídas algumas bases, no oligonucleotídeo 5' desenhado (5' TCTCCCGGCATATGTGCACAG 3'), na posição do códon referente ao aminoácido leucina 27 (A27), criando o códon da metionina (A27M) e assim utilizar a enzima de restrição *Nde* I para a clonagem no vetor de expressão pET23a. Para a região 3', foi utilizado o oligonucleotídeo comercial SP6, pois este se anelaria na região seguinte ao códon de terminação do cDNA da GrpE#1 presente no cDNA plasmidial do clone EST número BE614754. Entre o códon de terminação natural da GrpE#1 humana e o sítio de anelamento do oligonucleotídeo SP6 ao cDNA plasmidial pOTB7 (vetor do clone EST número BE614754) existia um sítio de restrição para a enzima *Xho* I. Esta enzima, juntamente com a enzima de restrição *Nde* I, foram utilizadas para a clonagem do cDNA da GrpE#1 amplificado pela PCR no vetor de expressão pET23a.

Clonagem do cDNA da proteína GrpE#1

As células do clone EST número BE614754 foram crescidas e o DNA plasmidial foi purificado do sedimento celular obtido pelo kit *QIAprep MiniPrep Kit* (Qiagen). Este DNA

plasmidial foi utilizado como molde para a PCR com o oligonucleotídeo 5' da GrpE#1 (descrito acima) e o oligonucleotídeo SP6. A PCR forneceu um produto de aproximadamente 600 pb correspondente ao esperado para o cDNA para a proteína GrpE#1, sem o peptídeo sinal, e com uma pequena sequência na região 3' do inserto correspondente ao vetor pOTB7. Este produto de PCR foi utilizado para a clonagem no vetor pGEMT com o auxílio do kit de clonagem *pGEMT easy*[®], conforme o protocolo do fabricante. Foi selecionada uma colônia contendo vetores pGEM que liberaram um inserto de aproximadamente 600 pb após a digestão deste com as enzimas *Nde* I e *Xho* I (dados não mostrados). Este fragmento foi submetido à reação de ligação ao vetor pET23a previamente preparado com as mesmas enzimas, como descrito no item 3.9 de Materiais e métodos.

A figura 28 representa o perfil eletroforético da digestão de vetores pET23aGrpE#1, obtidos de uma colônia de bactérias previamente selecionada, com as enzimas *Nde* I e *Xho* I. A digestão levou à liberação de um fragmento de DNA de aproximadamente 600 pb. O seqüenciamento das extremidades internas do SMC do vetor pET23a, utilizando os oligonucleotídeos *T7 promoter* e *T7 terminator* (GIBCO BRL), confirmou a clonagem do cDNA completo da proteína GrpE#1 (código GenBank número NM_025196), sem o peptídeo sinal e apresentando a mutação A27M, no vetor de expressão pET23a. Deste modo, foi construído o vetor pET23aGrpE#1 capaz de produzir a proteína recombinante GrpE#1 constituída por 191 resíduos de aminoácidos e *MM* igual a 21.467 g/mol.

Outros ensaios realizados para a proteína GrpE#1

A GrpE#1 foi expressa em alta quantidade apresentando um rendimento de aproximadamente 30 mg/L de indução e os resultados dos ensaios de purificação, de CD, de estabilidade térmica e química, de calorimetria, de UAL, de SAXS e de modelagem por homologia da proteína GrpE#1 estão descritos no artigo apresentado no Anexo IV (Borges, J.C., Fischer, H., Craievich, A.F., Hansen, L.D. & Ramos, C.H.I. (2003) Free human mitochondrial GrpE is a symmetric dimer in solution. *J. Biol. Chem.*, 278, 35337-35344.) e as informações adicionais para os experimentos estão contidas nas legendas das figuras. Alguns resultados de CD, de estabilidade térmica e química da GrpE#1, em alta força iônica, serão apresentados e discutidos em conjunto com aqueles obtidos para a GrpE#2.

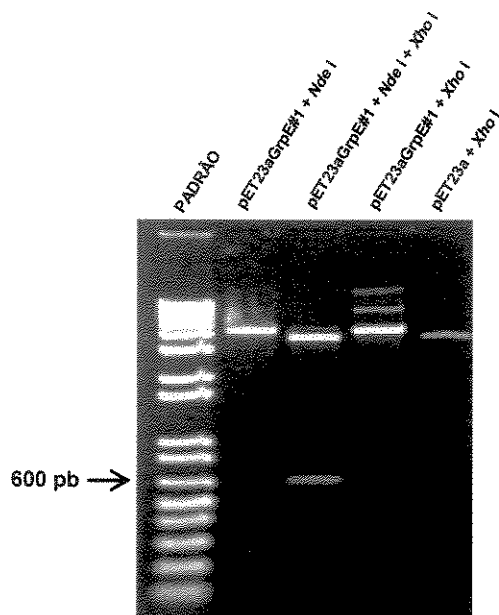


Figura 28: Clonagem do cDNA da proteína GrpE#1 no vetor pET23a. Gel de agarose 1% mostrando o perfil eletroforético de digestão do vetor pET23aGrpE#1 com as enzimas *Nde* I e *Xho* I. A digestão dupla do vetor pET23aGrpE#1 levou à liberação de um fragmento de aproximadamente 600 pb, como determinado pela comparação com o padrão de *MM 1 Kb plus DNA ladder* (GIBCO BRL). A digestão individual dos vetores por cada enzima evidenciou a diferença de tamanho do vetor pET23aGrpE#1 em relação ao pET23a. O vetor pET23aGrpE#1 foi submetido ao seqüenciamento de suas extremidades internas do SMC, com o auxílio dos oligonucleotídeos *T7 promoter* e *T7 terminator*. A análise das seqüências obtidas confirmou a clonagem do cDNA completo da proteína GrpE#1 humana (código GenBank número NM_025196) sem o peptídeo sinal para exportação para a mitocôndria no vetor pET23a.

4.3.2 GrpE#2

De forma similar àquela realizada para o cDNA da GrpE#1, o cDNA da proteína GrpE#2 também foi clonado sem a seqüência que codifica o peptídeo sinal para a exportação da proteína para a mitocôndria e sem a cauda de poli-histidina para auxiliar a purificação. A proteína GrpE#2 apresentou baixa solubilidade e baixa estabilidade em soluções de baixa força iônica, fatos que provavelmente inviabilizaram sua purificação por técnicas cromatográficas de troca iônica (dados não mostrados). Deste modo, o cDNA da GrpE#2 foi clonado em fusão à seqüência de nucleotídeos que codificam a cauda de poli-histidina do vetor pET28a e a proteína foi expressa com sucesso em baixas temperaturas. O processo de purificação foi feito por cromatografia de afinidade em alta força iônica e resultou na proteína estável e pura (Figura 30). Abaixo são apresentados os resultados referentes à caracterização biofísica e estrutural da GrpE#2. Também são apresentados alguns ensaios comparativos da proteína GrpE#2 em relação à GrpE#1 em alta força iônica.

Pesquisa em banco de dados e obtenção do cDNA plasmidial da GrpE#2

O cDNA da GrpE#2 humana foi identificado no banco de dados GenBank (*Homo sapiens GrpE-like 2, mitochondrial (E. coli) (GRPEL2), mRNA* – código GenBank número NM_152407) e utilizado para a busca de clones EST similares disponíveis para serem adquiridos comercialmente pelo sistema *IMAGE Consortium*. Foi obtido um clone EST (código GenBank número AI431866 – Tabela I) e o seqüenciamento do DNA das extremidades internas do SMC (vetor pT7T3D-Pac), com os oligonucleotídeos *M13 Forward* e *M13 Reverse*, mostrou a presença do cDNA completo da GrpE#2 humana (código GenBank número NM_152407).

Oligonucleotídeos para a PCR da GrpE#2

De maneira similar ao procedimento desenvolvido para a GrpE#1, o oligonucleotídeo 5' para a PCR da GrpE#2 foi desenhado de modo a excluir a seqüência de cDNA correspondente ao peptídeo sinal para a exportação da proteína para a mitocôndria. Para isto, foi considerado o trabalho de Naylor *et al.* (1998) com a respectiva GrpE#2 de camundongo, e a predição do peptídeo sinal realizada com o auxílio do programa MITOPROT II 1.0a4. Os primeiros 31 resíduos de aminoácidos da seqüência da GrpE#2 (código GenBank número Q8TAA5) foram considerados como sendo a seqüência do

peptídeo sinal de exportação da proteína para a mitocôndria. Para facilitar a clonagem do cDNA no sistema pET, o códon referente ao aminoácido fenilalanina 32 (F32) foi substituído no oligonucleotídeo 5' (5' ATGGCCGCTTCATATGAGCACTGC 3') para criar o códon da metionina (F32M) e assim possibilitar o uso da enzima de restrição *Nde* I para a clonagem no vetor de expressão. O oligonucleotídeo 3' foi desenhado para substituir algumas bases e desta forma criar um sítio de restrição para a enzima *Bam*H I (5' TCCTGATGGATCCATCACAGTCTTCTC 3') após o códon de terminação natural do cDNA da GrpE#2.

Clonagem do cDNA da proteína GrpE#2

As células do clone EST número AI431866 (Tabela I) foram crescidas e o sedimento celular obtido foi utilizado para a extração do DNA plasmidial pelo kit *QIAprep MiniPrep Kit* (Qiagen). Este DNA plasmidial foi usado como molde para a PCR da GrpE#2 com os oligonucleotídeos 5' e 3' descritos no tópico acima. Foi obtido um produto de PCR de aproximadamente 550 pb correspondente ao tamanho esperado para o cDNA da proteína GrpE#2. Este produto de PCR foi utilizado para a clonagem em pGEMT utilizando o kit de clonagem *pGEMT easy*[®] e conforme o protocolo do fabricante. Foi selecionada uma colônia de bactérias transformantes contendo vetores pGEM que liberaram um inserto de aproximadamente 550 pb após a digestão com as enzimas *Nde* I e *Bam*H I. O seqüenciamento parcial das extremidades internas do SMC do vetor pGEM, com os oligonucleotídeos *M13 Forward* e *M13 Reverse*, confirmou a clonagem do cDNA da proteína GrpE#2 humana. Este cDNA plasmidial foi utilizado para a subclonagem do cDNA da GrpE#2 no vetor de expressão pET28a utilizando as enzimas de restrição *Nde* I e *Bam*H I.

A figura 29 mostra o perfil eletroforético da digestão de vetores pET28aGrpE#2, obtidos de uma colônia de bactérias previamente selecionada, que liberaram um fragmento de aproximadamente 550 pb, a partir da incubação com as enzimas *Nde* I e *Bam*H I, sugerindo a clonagem do cDNA da GrpE#2 no vetor pET28a. A análise das seqüências, obtidas pelo seqüenciamento das extremidades internas do SMC do vetor pET28aGrpE#2, feita com o auxílio dos oligonucleotídeos *T7 promoter* e *T7 terminator* (GIBCO BRL), confirmou a clonagem do cDNA completo da proteína GrpE#2 (código GenBank número NM_152407). A análise das seqüências obtidas também mostrou a ausência da seqüência de cDNA que codifica o peptídeo sinal, para a exportação da proteína para mitocôndria, do

cDNA da GrpE#2 clonado no pET28a, em fusão ao cDNA da cauda de poli-histidina e a presença da mutação F32M. Assim, foi construído o vetor pET28aGrpE#2 capaz de produzir a proteína recombinante HisGrpE#2, expressa em fusão à cauda de poli-histidina na região amino-terminal, composta por 214 resíduos de aminoácidos e com *MM* igual a 24.003 g/mol.

Expressão e purificação da proteína GrpE#2

A figura 30 representa uma SDS-PAGE 12% que resume o processo de expressão e purificação da proteína GrpE#2 em fusão ao peptídeo contendo uma cauda de poli-histidina (HisGrpE#2). A HisGrpE#2 foi induzida à expressão pelo IPTG (compare canaletas 2 e 3 da figura 30). A proteína foi obtida na fração solúvel do lisado celular (canaleta 4), quando induzida à expressão em temperatura de 24 °C. O procedimento de lise celular realizado na presença de uma alta força iônica (500 mmol/L de NaCl) ajudou a manter a proteína HisGrpE#2 no sobrenadante. A HisGrpE#2 foi separada por dois passos cromatográficos: cromatografia de afinidade seguida por CEM preparativa, ambas realizadas na presença de NaCl 500 mmol/L (canaleta 5). A fim de se desenvolver estudos comparativos com a GrpE#1 humana, a cauda de poli-histidina foi clivada pela enzima proteolítica trombina (1 U/mg de HisGrpE#2, 16 horas e em temperatura de 25 °C) gerando a proteína GrpE#2. Após a clivagem pela trombina, a GrpE#2 foi novamente submetida à CEM para separação da proteína da cauda de poli-histidina e da trombina (canaleta 6). O controle da eficiência do processo de proteólise foi acompanhado por SDS-PAGE 12% (dados não mostrados). Após a proteólise da cauda de poli-histidina, a proteína GrpE#2 obtida é constituída por 197 resíduos de aminoácidos e tem *MM* de 22.121 g/mol.

Espectros de CD das GrpEs humanas

Ensaio de CD das proteínas GrpEs humanas foram realizados nas mesmas condições e os resultados são apresentados na figura 31. Os dados indicam que a GrpE#2 possui as duas bandas características para proteínas ricas em estruturas tipo α -hélice, como observado para a GrpE#1 (Anexo IV – Borges *et al.*, 2003). Entretanto, a GrpE#2 apresentou menor quantidade de estrutura secundária do que a GrpE#1. A quantidade de estrutura secundária das proteínas GrpE#1 e GrpE#2 estimada pelo programa CDNN Deconvolution, foram, respectivamente: 55% e 48% de α -hélice, 25% e 26% de folha β -pregueada e alças, e 20% e 26% de estrutura randômica (erros < 10%).

Ensaio de estabilidade térmica e química para as proteínas GrpEs humanas

A figura 32 apresenta os perfis de desenovelamento térmico das GrpEs humanas acompanhadas por CD. A GrpE#1 apresentou, entre as temperaturas de 20 a 90 °C, apenas uma T_m, enquanto a GrpE#2 apresentou duas T_ms. As primeiras T_ms das GrpEs humanas em 500 mmol/L de NaCl foram similares: 53 °C e 50 °C para a GrpE#1 e a GrpE#2, respectivamente. Porém, a GrpE#2 apresentou menor perda relativa de sinal do que a GrpE#1. O desenovelamento térmico da GrpE#1 acompanhado por CD sofreu influência da força iônica, pois em força iônica menor a GrpE#1 perdeu estrutura secundária na temperatura de 57 °C (Anexo IV – Borges *et al.*, 2003). A segunda T_m detectável da GrpE#2 foi em 78 °C e resultou na agregação seguida da precipitação da proteína (um precipitado esbranquiçado foi observado na cubeta). As proteínas GrpEs humanas apresentaram o desenovelamento térmico irreversível, mesmo se as soluções protéicas fossem aquecidas somente até temperaturas pouco superiores ao primeiro T_m (dados não mostrados – Anexo IV – Borges *et al.*, 2003). A GrpE#1 apresentou um pouco de estrutura secundária em temperaturas próximas a 90 °C, indicando que ela não foi completamente desenovelada pelo calor, enquanto a GrpE#2 foi completamente desenovelada, sofrendo agregação e precipitação, como apresentado acima. A tabela V sumariza os resultados obtidos para as GrpEs humanas e os compara com dados obtidos da literatura para a proteína EcGrpE.

Tabela V: Parâmetros biofísicos das GrpEs humanas e da proteína EcGrpE analisados por CD. Dados de T_m e EMR das GrpEs humanas determinadas neste trabalho e da EcGrpE obtidos de trabalhos da literatura científica.

Proteína	T _m (°C)		EMR _{222 nm} (Deg.cm ² .dmol ⁻¹) x 10 ³		
	1º	2º	20 °C	60 °C	90 °C
GrpE#1	53 ^{&}	ND	19	10	8
GrpE#2	50	78	16	10	2
EcGrpE [#]	50	76	20	10	6

[#] ensaios de desenovelamento realizados em soluções sem sal (Grimshaw *et al.*, 2001; Gelinas *et al.*, 2002); ND: não determinado. [&] T_m em 57 °C em baixa força iônica (Anexo IV – Borges *et al.*, 2003).

Os ensaios de desenovelamento das GrpEs humanas induzido por uréia, monitorado por CD em λ de 222 nm, são apresentados na figura 33. A GrpE#1 apresentou um perfil de desenovelamento induzido por uréia que se assemelha a uma sigmóide com uma C_m (concentração média de transição) em 2,0 mol/L. A GrpE#2 apresentou um perfil de desenovelamento induzido por uréia não-cooperativo com pelo menos duas transições. Estes dados sugerem a presença de intermediários na via de desenovelamento químico da GrpE#2 nas condições estudadas. Porém, as proteínas iniciaram seus processos de desenovelamento induzido por uréia em concentrações parecidas de uréia (aproximadamente 1 mol/L), sendo que o processo de desenovelamento foi irreversível para as duas proteínas (dados não mostrados).

Análise hidrodinâmica da proteína GrpE#2

As propriedades hidrodinâmicas da GrpE#2 foram avaliadas por UAL através da técnica de velocidade de sedimentação e os dados obtidos são apresentados na figura 34. Os experimentos foram realizados a 40.000 rpm, temperatura constante de 20 °C e em concentrações de 150 a 450 $\mu\text{g/mL}$. A análise dos dados pelo método de *Sedimentation Time Derivative* (Figura 34a) mostrou que a distribuição da integral $g(s^*)$ foi bem ajustada por duas gaussianas, na qual indicou a presença de uma pequena população de agregados. Esta população de agregados foi confirmada pela análise de *van Holde-Weischet* (Figura 34a interna), onde o comportamento não-ideal da partícula foi evidenciado devido ao desvio da curva para pequenos valores de s^* . A figura 34b mostra o $s_{20,w}$ em função da concentração da proteína GrpE#2 e a extrapolação da curva em 0 mg/mL (Laue, 2001), resultando em $s_{20,w}^0 = 2,80 \pm 0,01$ S. Estes dados juntamente com o $D_{20,w}^0$ (determinado por EDL em 3 diferentes concentrações: 1,0, 2,7 e 6,0 mg/mL) para a GrpE#2 (Tabela VI), sugerem que a proteína possui uma *MM* de aproximadamente 50 kDa, através a aplicação da equação 10. Os dados sugerem que a GrpE#2 pode não ser tão assimétrica quanto a GrpE#1 ou que a GrpE#2 pode ser mais flexível que a GrpE#1 (ver abaixo).

Ensaio de SAXS da proteína GrpE#2

Os experimentos de SAXS foram realizados na linha SAS do LNLS (Kellermann *et al.*, 1997) utilizando porta-amostra de mica, em duas distâncias 750 e 860 mm e λ de 1,488 Å. As medidas foram realizadas em duas concentrações da proteína GrpE#2 (3,5 e 9,3 mg/mL)

no tampão Tris-HCl 25 mmol/L (pH 7,5) contendo 500 mmol/L de NaCl e 1 mmol/L β -mercaptoetanol. Os dados foram tratados utilizando o programa TRAT1D (Oliveira, 2003) e normalizados pela concentração. As partes das curvas oriundas das diferentes concentrações de proteína foram unidas para se obter dados de qualidade estatística em um grande intervalo de " q ". A curva da $I(q)$ de SAXS para a GrpE#2 em alta concentração mostrou indícios de agregação e os dados a baixos q foram desprezados. A curva da $p(r)$ e o R_g da GrpE#2 foram calculados a partir das curvas corrigidas e normalizadas pela transformada de Fourier indireta realizada pelo programa GNOM (Semenyuk & Svergun, 1991). A figura 35 apresenta a curva da $I(q)$ para a proteína GrpE#2 com o ajuste realizado pelo programa GNOM (Figura 35a) e a curva da $p(r)$ resultante da transformada de Fourier indireta (Figura 35b). A curva da $p(r)$ da GrpE#2 sugere que a proteína apresenta uma conformação prolata alongada e um D_{max} de aproximadamente 145 Å. O sumário dos dados obtidos estão dispostos na Tabela VI. Os dados de SAXS foram utilizados para obter uma estimativa da MM da GrpE#2, pela comparação com a intensidade de espalhamento da BSA como padrão secundário de MM . A GrpE#2 apresentou uma MM de aproximadamente 42 kDa que está em concordância com o modelo da GrpE#2 na forma de dímeros (44 kDa). Estes dados também fornecem indícios de que não houve formação de agregados na solução nas amostras em baixa concentração.

A figura 36 apresenta o gráfico de Kratky obtido a partir da curva da $I(q)$ das proteínas GrpEs humanas. A análise do perfil das curvas obtidas pela relação $I(q) \times q^2$ versus q , mostra a ausência do perfil de sino e sugere a presença de porções flexíveis nas GrpEs humanas. As curvas do gráfico de Kratky das GrpEs humanas também sugerem que as porções com flexibilidade são mais evidenciadas nos dados da GrpE#2 em comparação com a GrpE#1.

Modelagem ab initio para a proteína GrpE#2

Os modelos *ab initio* obtidos para a GrpE#2 foram gerados pelo programa DAMMIM (Svergun, 1999). Como entrada de dados ao programa foi fornecido o D_{max} da partícula, obtido previamente pela curva da $p(r)$, e o programa construiu uma esfera preenchida de pequenas esferas cujo $r_{esf} \ll D_{max}$. O programa ajustou a configuração das pequenas esferas que melhor se ajustaram aos dados experimentais da curva da $I(q)$ da GrpE#2, eliminando as esferas dispersas no modelo. Como os dados experimentais não são em alta

resolução, o modelo final gerado corresponde à média de 20 cálculos independentes utilizando o programa DAMAVER.

O modelo *ab initio* da GrpE#2 está representado na figura 37a e a análise deste sugere que a proteína possui uma forma assimétrica. A figura 37b representa a superposição do modelo *ab initio* da GrpE#2 sobre o modelo *ab initio* da GrpE#1 (Anexo IV – Borges *et al.*, 2003). Os resultados sugerem que as GrpEs humanas possuem tamanho e forma geral parecidas, mas possuem conformações diferentes.

Tabela VI: Conjunto de parâmetros matemáticos, estruturais e hidrodinâmicos da proteína GrpE#2 obtidos nos experimentos de SAXS, UAL, EDL e para o modelo *ab initio*. Os dados experimentais de $s_{20,w}^0$, $D_{20,w}^0$, D_{max} e R_g são comparados com os dados obtidos para os modelos *ab initio*, gerados a partir dos dados de SAXS, utilizando o programa HydroPro (em condições padrão e assumindo a GrpE#2 como um dímero de 44 kDa).

Parâmetro		Método			
SAXS	Hidrodinâmicos	SAXS	Modelo <i>Ab initio</i>	UAL	EDL
NSC [#]	-	10	10	-	-
Discrepância χ	-	0,157	0,441	-	-
Resolução (Å)	-	31,4	31,4	-	-
MM (kDa)		42 ± 2	-	$50 \pm 1^{\$}$	60 ± 10
R_g (Å)		42 ± 2	$43,0 \pm 0,5$	-	-
D_{max} (Å)		~ 145	~ 150	-	-
-	$D_{20,w}^0$ (10^{-7} cm ² /s)	-	$4,8 \pm 0,1$	-	$5,3 \pm 0,1$
-	$s_{20,w}^0$ (S)	-	$2,30 \pm 0,03$	$2,80 \pm 0,01$	-
-	P^*	-	$1,42 \pm 0,03$	-	$1,74 \pm 0,08$

Parâmetros obtidos pelas seguintes técnicas: $s_{20,w}^0$ experimentos de velocidade de sedimentação; $D_{20,w}^0$ experimentos de EDL; $^{\$}MM$ calculada a partir dos dados $s_{20,w}^0$ e $D_{20,w}^0$ utilizando a equação 10; D_{max} e R_g oriundos da curva da $p(r)$; *P calculado para uma proteína de 44 kDa. $^{\#}NSC = D_{max} (q_{max} - q_{min})/\pi$ representa uma medida da quantidade de informação contida na curva de SAXS.

Análise hidrodinâmica dos modelos ab initio gerados para a proteína GrpE#2

As análises hidrodinâmicas dos modelos *ab initio* foram realizadas com o auxílio do programa HydroPro (García de la Torre *et al.*, 2000). Os parâmetros hidrodinâmicos foram estimados para cada um dos 20 modelos *ab initio* obtidos, de maneira independente pelo programa DAMMIM, e são representados pela média dos dados mais o desvio padrão da média. Para isto, foi utilizado a *MM* correspondente ao dímero, a viscosidade e a densidade em condições padrão (água e temperatura de 20 °C), e V_{bar} de 0,738 mL/mg estimado a partir da seqüência de aminoácidos da proteína GrpE#2. Estas características foram estimadas pelo programa Sednterp. A tabela VI sumariza os parâmetros hidrodinâmicos obtidos experimentalmente e os compara com os parâmetros estimados pelo programa HydroPro.

Superposição dos modelos construídos por modelagem molecular por homologia da proteína GrpE#1 com o modelo ab initio da proteína GrpE#2

A figura 38 representa a superposição dos modelos da GrpE#1 que foram construídos por modelagem molecular por homologia ao modelo *ab initio* da proteína GrpE#2 construído a partir dos dados de SAXS. A modelagem molecular por homologia da proteína GrpE#1 foi feita como descrito no Anexo IV (Borges *et al.*, 2003). O modelo da GrpE#1 construído a partir do monômero distal não apresentou um bom ajuste no domínio de interação das GrpEs com o DLN das Hsp70. Já o modelo gerado a partir do monômero proximal apresentou um melhor ajuste naquela região, porém ainda de maneira insatisfatória. Os dados sugerem que, em solução, a GrpE#2 apresenta uma outra conformação do que as observadas para os monômeros da EcGrpE na estrutura em alta resolução (código PDB número 1DKG).

Comparação da seqüência primária das proteínas GrpEs humanas com a EcGrpE

A figura 39 representa o alinhamento das seqüências de aminoácidos entre as duas GrpEs humanas e a EcGrpE. As principais regiões da proteína EcGrpE estão destacadas como descrito na figura 2: a região amino-terminal sem estrutura determinada (EcGrpE₁₋₃₉); as duas longas α -hélices (EcGrpE₄₀₋₈₈); o feixe de 4 α -hélices (EcGrpE₈₉₋₁₃₇) e o domínio carboxi-terminal estruturado em folha β -pregueada (EcGrpE₁₃₈₋₁₉₇). A tabela VII descreve a identidade e a similaridade total e também a similaridade por região, entre as proteínas

EcGrpE, GrpE#1 e GrpE#2. As seqüências de aminoácidos da GrpE#1 e da GrpE#2 apresentaram as mesmas percentagens de identidade e similaridade em relação à EcGrpE.

Na avaliação da similaridade entre as regiões descritas acima, as GrpEs humanas apresentaram maior similaridade entre elas do que com a EcGrpE, principalmente entre os resíduos 89-191. Entretanto, na região amino-terminal a GrpE#1 e a GrpE#2 são mais similares à EcGrpE do que entre si. A similaridade entre a EcGrpE e a GrpE#2 foi de aproximadamente 60% na região das longas α -hélices, enquanto que ela foi de aproximadamente 51% entre a EcGrpE e a GrpE#1 (Tabela VII). A similaridade entre a EcGrpE e a GrpE#1 foi de aproximadamente 58% na região do feixe de 4 hélices, enquanto que entre a EcGrpE e a GrpE#2 a similaridade na mesma região foi de 48% (Tabela VII). No domínio carboxi-terminal estruturado em folha β -pregueada a similaridade entre as GrpEs humanas e a EcGrpE foi de aproximadamente 52% (Tabela VII).

Tabela VII: Homologia na seqüência de aminoácidos entre as proteínas GrpE#1, GrpE#2 e EcGrpE. A identidade e a similaridade entre as proteínas foram calculadas com o auxílio do alinhamento múltiplo realizado pelo programa ClustalW. Foram considerados aminoácidos similares como sendo: aminoácidos polares; aminoácidos hidrofóbicos; aminoácidos carregados positivamente; aminoácidos carregados negativamente e aminoácidos aromáticos.

Proteína	Identidade (%)		Similaridade (%)		Similaridade por região da EcGrpE (%)			
	GrpE#1	GrpE#2	GrpE#1	GrpE#2	1-39	40-85	89-137	138-191
EcGrpE	30	29	51	52	-			
GrpE#1	-	44	-	66	40	51	58	52
GrpE#2	44	-	66	-	40	60	48	52
GrpE#1 x GrpE#2	-				30	66	75	77

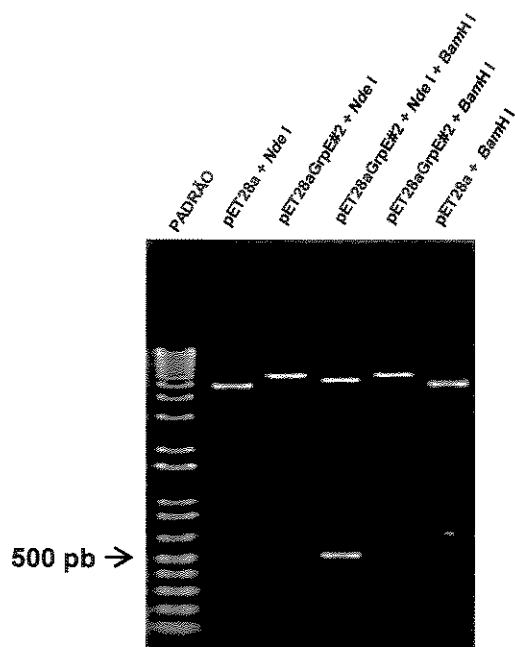


Figura 29: Clonagem do cDNA da proteína GrpE#2 humana no vetor pET28a. Gel de agarose 1% mostrando o perfil eletroforético de digestão dos vetores pET28aGrpE#2 e pET28a com as enzimas *Nde* I e *Bam*H I. A digestão do vetor pET28aGrpE#2 com as enzimas *Nde* I e *Bam*H I resultou na liberação de um fragmento de aproximadamente 550 pb referente ao cDNA da GrpE#2 humana, como determinado pela comparação com o padrão de *MM 1 Kb plus DNA ladder* (GIBCO BRL). A digestão individual dos vetores com cada enzima evidenciou a diferença no tamanho do vetor pET28aGrpE#2 linear em relação ao pET28a linear. A análise das seqüências do cDNA obtidas pelo seqüenciamento das extremidades internas do SMC do pET28aGrpE#2, utilizando os oligonucleotídeos *T7 promoter* e *T7 terminator*, confirmou a clonagem do cDNA completo da proteína GrpE#2 humana (código GenBank número NM_152407), sem o cDNA do peptídeo sinal para a exportação da proteína para a mitocôndria e em fusão ao cDNA da cauda de polihistidinas.

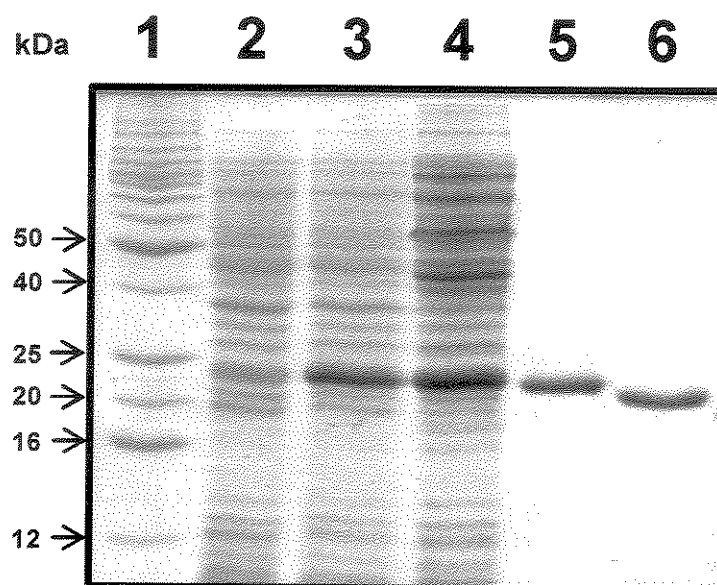


Figura 30. Análise do processo de indução e purificação da GrpE#2 analisada por SDS-PAGE 12%. A GrpE#2 foi expressa em fusão à cauda de poli-histidina para auxiliar o processo de purificação utilizando a cromatografia de afinidade seguida por CEM preparativa (não mostrado). A HisGrpE#2 foi expressa em alta quantidade apresentando um rendimento de aproximadamente 20 mg/L de indução. Para ensaios comparativos com a GrpE#1, a cauda de poli-histidina foi clivada pela enzima proteolítica trombina e a amostra foi reaplicada na CEM preparativa para separação da GrpE#2 da trombina e do peptídeo da cauda de poli-histidina. Canaleta 1: Marcador de *MM* em kDa (à esquerda); canaleta 2: sedimento bacteriano antes da indução; canaleta 3: sedimento bacteriano depois da indução com IPTG por 5 horas, sendo expressa em fusão com uma cauda de poli-histidina em meio LB a 24 °C; canaleta 4: após a lise das bactérias, a GrpE#2 foi localizada na sobrenadante; canaleta 5: HisGrpE#2 após os processo de purificação por cromatografia de afinidade e CEM preparativa; canaleta 6: GrpE#2, após proteólise com a enzima trombina (1 U/mg de HisGrpE#2, 16 horas e em temperatura de 25 °C).

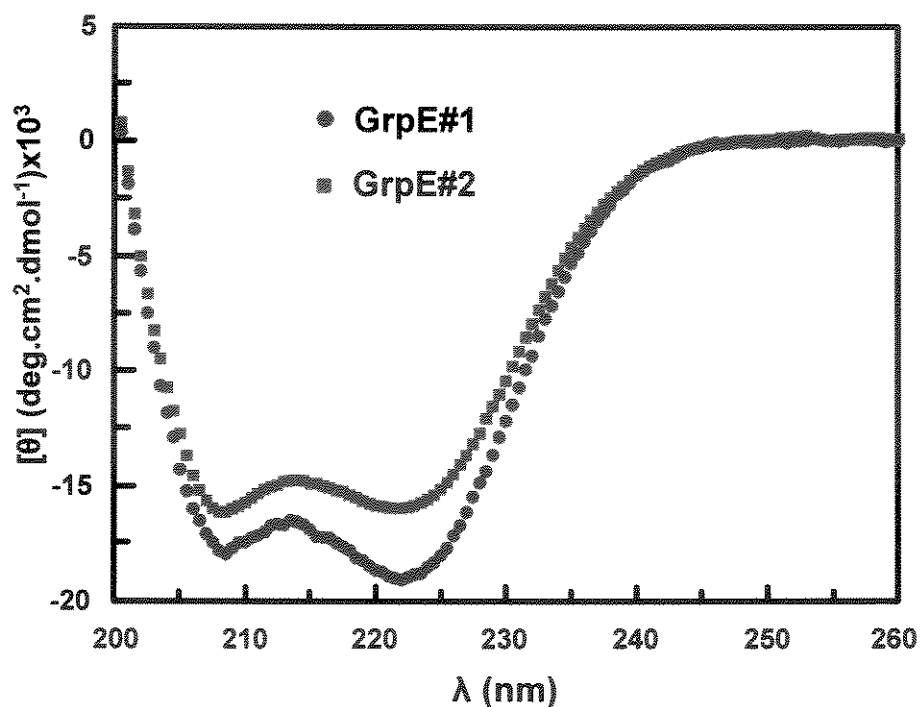


Figura 31. Análise de CD para as proteínas GrpEs humanas. Espectros da GrpE#1 e GrpE#2 foram medidos entre 200 e 260 nm em tampão Tris-HCl 25 mmol/L (pH 7,5), contendo 500 mmol/L de NaCl e 1 mmol/L de β -mercaptoetanol a 20 °C, usando uma cela de 0,1 mm de passo ótico e normalizados em *EMR*. A quantidade de estrutura secundária das GrpE#1 e GrpE#2 foi estimada pelo programa CDNN Deconvolution e foram, respectivamente: 55% e 48% of α -hélice, 25% e 26% folha β -pregueada e alças, e 20% e 26% estrutura randômica (erros < 5%).

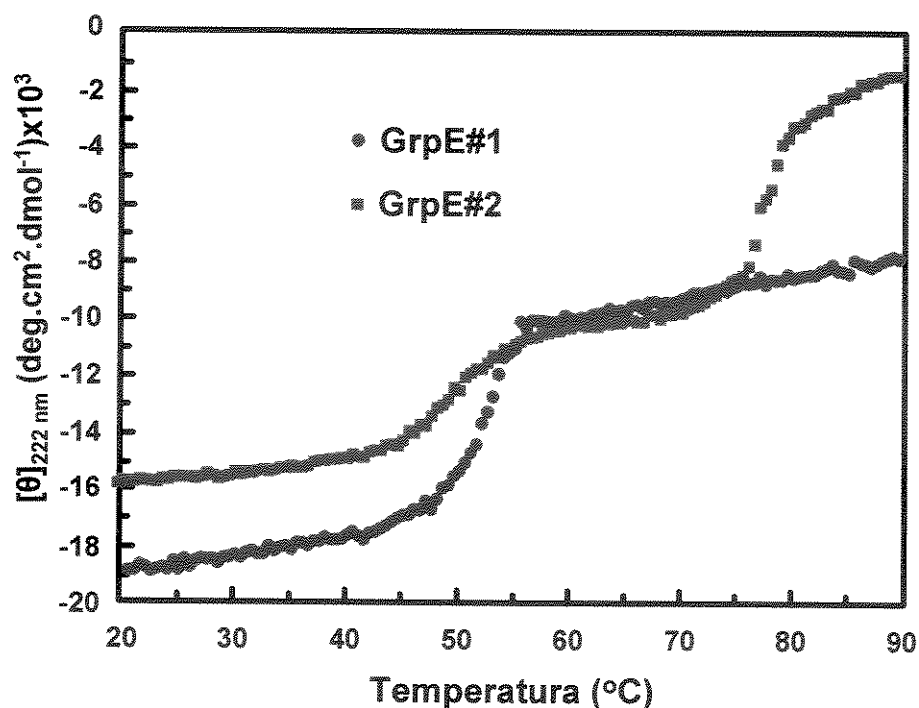


Figura 32: Ensaio de estabilidade térmica para as proteínas GrpE humanas. O desenovelamento térmico (20 a 90 °C) da GrpE#1 e GrpE#2 acompanhado por CD a 222 nm no tampão Tris-HCl 25 mmol/L (pH 7,5) contendo 500 mmol/L NaCl e 1 mmol/L β-mercaptoetanol. Os dados foram normalizados em *EMR* e estão apresentados em função da temperatura. A GrpE#1 apresentou o T_{m1} em 53 °C e a GrpE#2 em 50 °C. A GrpE#2 apresentou também um T_{m2} em 78 °C, que não foi observado para a GrpE#1. Mesmo a 90 °C, a GrpE#1 apresentou alguma estrutura secundária. Para ambas as GrpEs humanas, o desenovelamento térmico foi irreversível, mesmo se aquecidas a apenas a 55 °C (dados não mostrados – ver Anexo IV – Borges *et al.*, 2003).

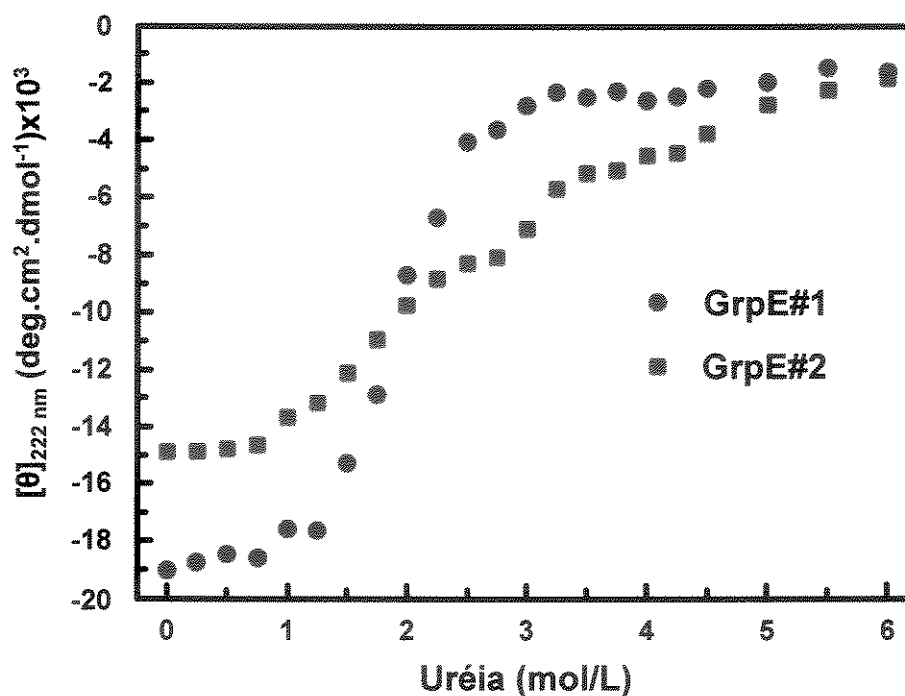


Figura 33: Desenovelamento das proteínas GrpE humanas por uréia. O desenovelamento das GrpE#1 e GrpE#2 foi induzido por uréia em tampão Tris-HCl 25 mmol/L (pH 7,5) na presença de 500 mmol/L de NaCl e 1 mmol/L de β-mercaptoetanol. Quantidade fixas das proteínas GrpE#1 e GrpE#2 (aproximadamente 5 – 10 μmol/L) foram diluídas em diferentes concentrações de uréia e o espectros de CD medidos após 45 minutos de incubação em temperatura de 22 °C. A leitura a 222 nm foi normalizada para *EMR* e é apresentada no gráfico acima em função da concentração de uréia. De modo geral, a GrpE#1 apresentou um perfil de desenovelamento sigmoidal, enquanto que a GrpE#2 apresentou espécies estáveis em concentrações superiores de uréia, sugerindo a presença de intermediários na via de desenovelamento por uréia da GrpE#2.

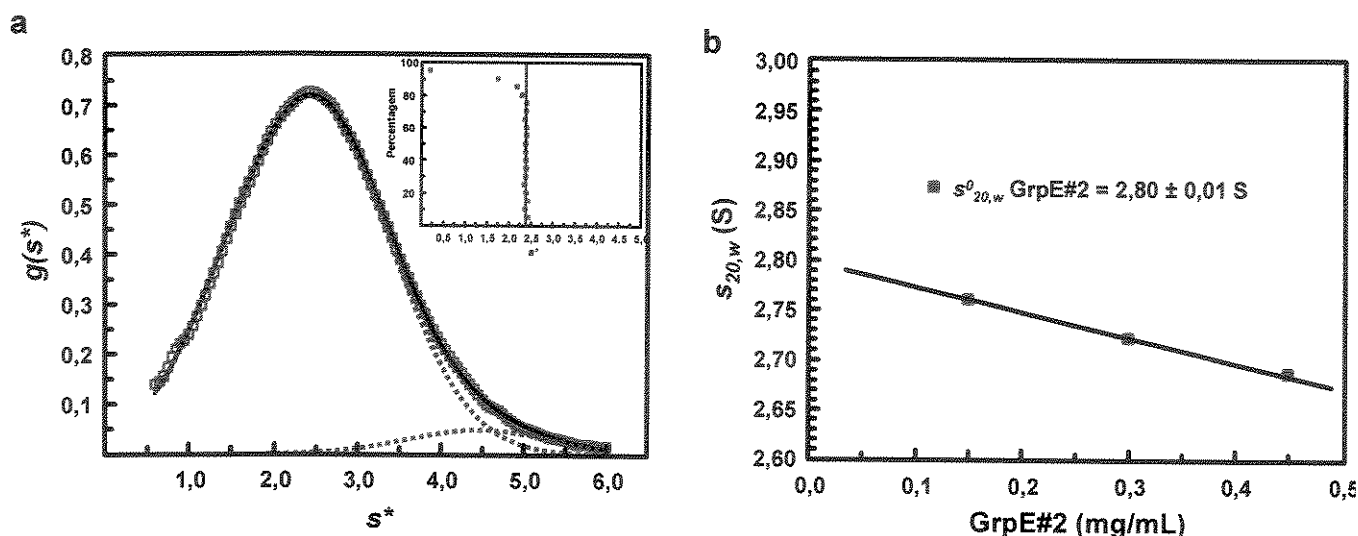


Figura 34: Experimentos de UAL: velocidade de sedimentação para a proteína GrpE#2.

Os experimentos de velocidade de sedimentação foram realizados a 20 °C, 40.000 rpm (AN-60Ti rotor) e a aquisição de dados em 230 nm. A análise dos dados foi feita pelo programa Origin® (MicroCal). A GrpE#2 foi preparada em tampão Tris-HCl 25 mmol/L (pH 7,5), contendo 500 mmol/L de NaCl e 1 mmol/L de β -mercaptoetanol e utilizada para os experimentos em três concentrações (de 150 a 450 $\mu\text{g/mL}$). A figura a) representa um experimento realizado com a GrpE#2, na concentração de 450 $\mu\text{g/mL}$ analisado pelo método de *Sedimentation Time Derivative*. A distribuição $g(s^*)$ mostra que duas curvas gaussianas resultam em um melhor ajuste do que apenas uma curva. A primeira Gaussiana mostra o s^* . O gráfico interno mostra uma análise de dados pelo método de *van Holde-Weischet* em que o ajuste mostra que a GrpE#2 apresentou um comportamento não-ideal. b) Estimativa do $s_{20,w}^0$. Gráfico dos valores de $s_{20,w}$ versus a concentração da GrpE#2 mostrando a extrapolação da curva em 0 mg/mL de concentração de proteína (Laue, 2001), resultando em uma estimativa para o $s_{20,w}^0$ da GrpE#2 de $2,80 \pm 0,01 \text{ S}$.

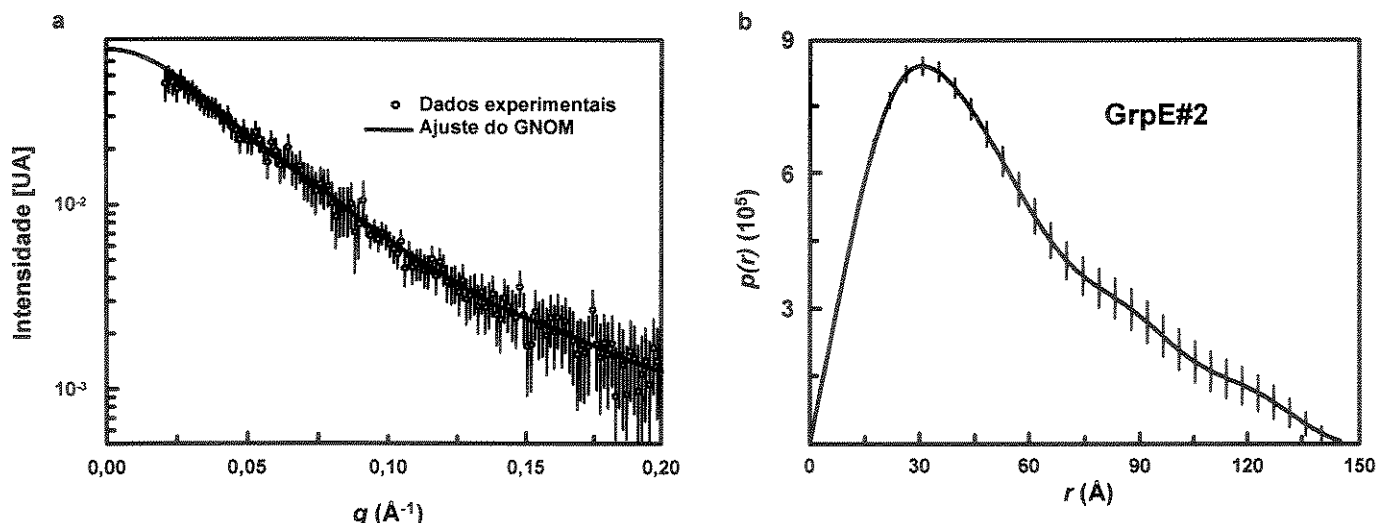


Figura 35: Dados de SAXS obtidos para a proteína GrpE#2. Os dados de SAXS foram coletados na linha de luz síncrotron SAS do LNLS em duas concentrações da proteína GrpE#2 (3,5 e 9,3 mg/mL) no tampão Tris-HCl 25 mmol/L (pH 7,5), contendo 500 mmol/L de NaCl e 1 mmol/L de β -mercaptoetanol. a) curvas de espalhamento $I(q)$: após tratamento dos dados e normalização pela concentração, as curvas foram unidas para obter dados de qualidade estatística em um grande intervalo de “ q ”. As curvas em alta concentração apresentaram indícios de agregação e os dados à baixo q foram desprezadas. A linha contínua representa o ajuste realizado pelo programa GNOM. b) Curva da $p(r)$ resultante da curva de espalhamento calculada pela transformada de Fourier indireta executada pelo programa GNOM. A curva da $p(r)$ mostra que a GrpE#2 possui um D_{max} de aproximadamente 145 Å e o formato da curva indica que a GrpE#2 é uma partícula prolata alongada. O R_g da GrpE#2 calculado a partir da curva da $p(r)$ foi de 42 ± 2 Å.

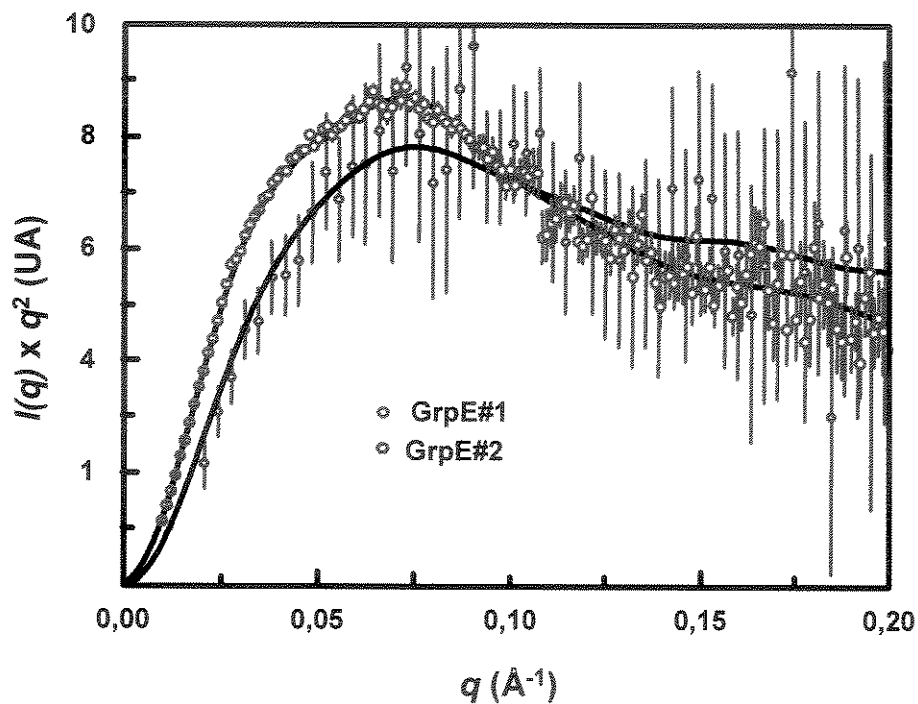


Figura 36: Gráfico de Kratky das proteínas GrpE humanas. A análise do perfil gráfico de Kratky fornece informações sobre a compactação da partícula espalhadora. Partículas compactas resultam em um perfil de sinusóide no gráfico de Kratky. Partículas que possuem segmentos flexíveis perdem o perfil de sinusóide. Os dados das GrpEs humanas indicam que estas proteínas possuem segmentos flexíveis em suas estruturas, sendo que na GrpE#2 o grau de flexibilidade foi mais evidenciado.

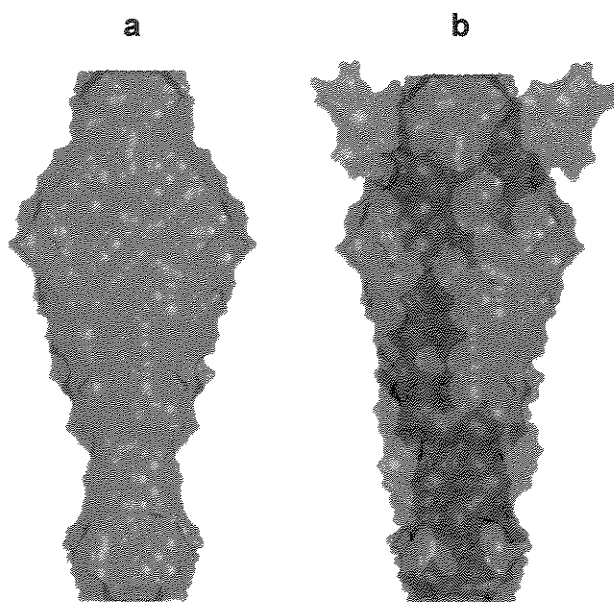


Figura 37: Modelos *ab initio* gerados para a proteína GrpE#2. Os modelos *ab initio* obtidos para a GrpE#2 foram gerados pelo programa DAMMIM e representa a média da superposição de 20 cálculos independentes, utilizando o programa DAMAVER. a) modelo *ab initio* da GrpE#2. b) superposição do modelo *ab initio* gerado para a GrpE#1 (azul – Anexo IV – Borges *et al.*, 2003) ao modelo *ab initio* da GrpE#2 (vermelho). A figura mostra que os modelos possuem dimensões aproximadas, mas com conformações diferentes de alguns de seus domínios.

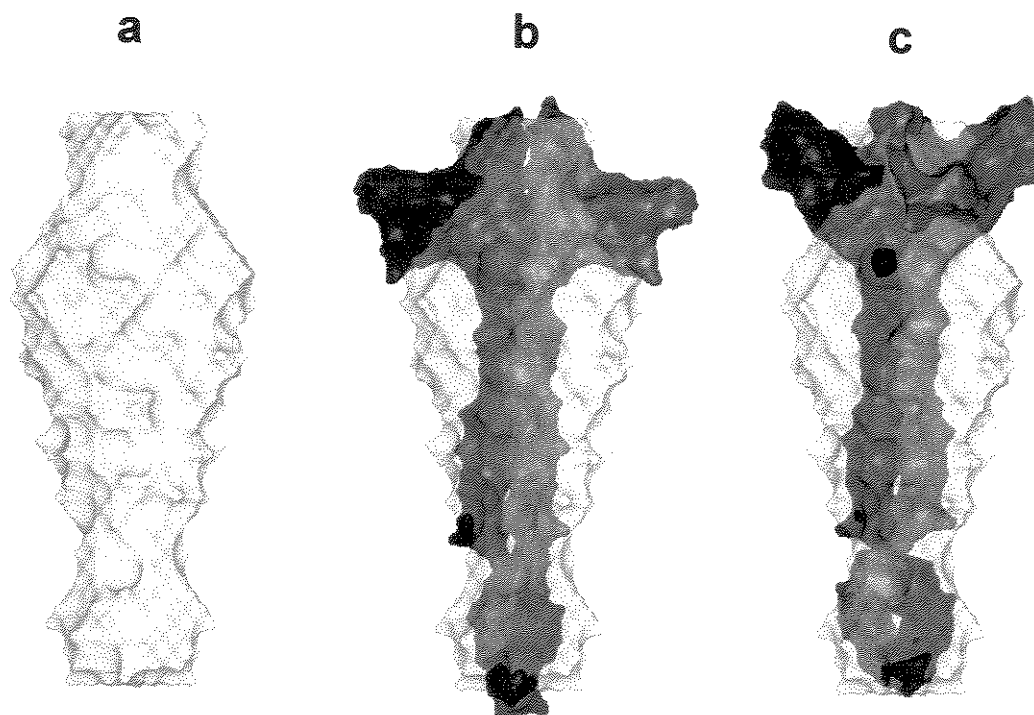


Figura 38: Comparação do modelo *ab initio* da GrpE#2 com os modelos da estrutura quaternária da GrpE#1. a) modelo *ab initio* gerado para a GrpE#2 a partir dos dados de SAXS; b) superposição do modelo da GrpE#1 gerado por homologia a partir do monômero proximal da EcGrpE com o modelo *ab initio* da GrpE#2 e c) superposição do modelo da GrpE#1 gerado por homologia a partir do monômero distal da EcGrpE com modelo *ab initio* da GrpE#2. A estrutura em alta resolução da EcGrpE, em complexos com o DLN da DnaK, mostra que apenas um monômero no dímero da EcGrpE está ligado ao DLN da DnaK (o monômero proximal) e que este assume uma conformação diferente daquele não ligado (o monômero distal). A principal diferença entre os monômeros está na angulação da abertura do domínio carboxi-terminal, que interage com o DLN da DnaK, em relação ao eixo axial da proteína EcGrpE (código PDB número 1DKG – Harrison *et al.*, 1997). O modelo da GrpE#1 construído a partir do monômero distal não apresentou um bom ajuste com o modelo *ab initio* da GrpE#2 principalmente no domínio de interação da EcGrpE com a DnaK. O ajuste deste domínio foi melhor na superposição do modelo *ab initio* da GrpE#2 com o modelo da GrpE#1 construído a partir do monômero proximal da EcGrpE, porém ainda de maneira insatisfatória. Os dados sugerem que em solução a GrpE#2 se apresenta em uma conformação outra que aquelas observadas na estrutura em alta resolução.

5 DISCUSSÃO

O objetivo geral deste trabalho foi estudar a relação da estrutura-função de proteínas humanas pertencentes à família das chaperones moleculares das Hsp70. Assim, foi necessária a aquisição dos cDNAs das proteínas a serem estudadas, as clonagens destes em vetores de expressão, a expressão e a purificação das proteínas para posterior caracterização biofísica. Considerando isto, cada uma das proteínas é discutida em separado. Para as proteínas Hsp40 (DjA1, DjA1₁₋₃₃₂ e DjB4) e GrpE#1, a discussão dos resultados está apresentada nos manuscritos em anexo (Anexo II e IV). Abaixo estão discutidos os resultados para as proteínas Hsp70-1A, mortalina e GrpE#2; esta última em análises comparativas com a GrpE#1.

5.1 Hsp70-1A

Este estudo descreve a obtenção e a caracterização da proteína humana Hsp70-1A, uma representante citoplasmática classificada como membro da subfamília Hsc70 das Hsp70, cuja expressão é constitutiva. A principal questão em relação às Hsp70 é a de compreender melhor quais são as mudanças conformacionais provocadas na proteína pela ligação do nucleotídeo adenosina, pela interação com as co-chaperones e com o substrato. Assim, a fim de aumentar a compreensão sobre o funcionamento dessa maquinaria, várias técnicas biofísicas de baixa resolução foram aplicadas para estudar a Hsp70-1A na presença e na ausência dos nucleotídeos adenosina (MgADP e MgATP). Estes nucleotídeos são capazes de deslocar o equilíbrio conformacional dos domínios da Hsp70, influenciando a ligação ou a liberação de substratos (revisão em Mayer *et al.*, 2001).

5.1.1 Clonagem do cDNA, expressão e purificação da proteína Hsp70-1A

O cDNA da proteína Hsp70-1A foi clonado no vetor de expressão pET28a, em fusão com a cauda de poli-histidina, a partir de um clone adquirido comercialmente e que continha o cDNA completo da proteína Hsp70-1A. A proteína Hsp70-1A foi expressa de maneira heteróloga, sendo purificada com um bom rendimento e grau de pureza. Como a Hsp70-1A liga nucleotídeos adenosina (ADP ou ATP), foi necessário purificar a proteína em condições de ausência destes ligantes. Theyssen *et al.* padronizou a utilização da enzima fosfatase alcalina ácida para a purificação da proteína DnaK livre de nucleotídeos adenosina porventura presentes nas amostras (Theyssen *et al.*, 1996). A aplicação desta metodologia

para a purificação da Hsp70-1A possibilitou a obtenção de amostras sem contaminações pelos nucleotídeos adenosina e os estudos conformacionais da proteína na presença de nucleotídeos adenosina. Os principais dados que indicaram a eficiência deste tratamento com a fosfatase alcalina ácida, para as amostras da Hsp70-1A, foram as mudanças na primeira e terceira Tms avaliadas pelos experimentos de desenovelamento térmico monitorados por CD. As amostras de Hsp70-1A contendo os nucleotídeos adenosina (ATP, ADP, MgATP ou MgADP) apresentaram aumentos de 4 a 9 °C na primeira e terceira temperaturas de desenovelamento. Estes aumentos não foram observados nas amostras da Hsp70-1A com adição de AMP e MgAMP. Os dados de ITC também corroboraram a eficiência na obtenção da Hsp70-1A na ausência de nucleotídeos adenosina, pois as K_{As} observadas na interação entre Hsp70-1A e MgATP ou MgADP foram diferentes entre si e compatíveis com os dados da literatura (ver adiante).

A segunda etapa de purificação foi feita através de uma CEM e as amostras da Hsp70-1A apresentaram poli-dispersidade, pois eluíram em 3 frações apresentando diferentes MM aparentes, porém com mesma MM como determinado por análises de SDS-PAGE. As análises das proteínas destas frações sugeriram que a primeira, eluída próximo ao volume morto da coluna, não se comportou em estado de equilíbrio com as demais frações, e ainda apresentou uma menor quantidade de estrutura secundária em análises por CD, e foi descartada. Benaroudj *et al.* (1996) também mostra, por CEM, que a primeira fração de amostras congeladas de uma Hsc70 de mamífero não está em equilíbrio com as demais frações.

A segunda fração das amostras de Hsp70-1A também apresentou poli-dispersidade e se mostrou em equilíbrio com a terceira e principal fração de 70 kDa. A fração de 70 kDa apresentou comportamento compatível a um sistema monodisperso e, após estocagem, começou a apresentar oligomerização dependente do tempo e da concentração. Este comportamento também é descrito para a DnaK e Hsc70 bovina, que mostram a formação de dímeros e trímeros (Schlossman *et al.*, 1984; Palleros *et al.*, 1991; Schönfeld *et al.*, 1995; Wilbanks *et al.*, 1995). Benaroudj *et al.* (1995; 1996; 1997) relatam a existência de um equilíbrio lento em amostras de Hsc70, entre espécies monoméricas e de MM maior, que é tempo e concentração-dependente. Porém, as amostras de Hsc70 utilizadas foram previamente dispensadas em alíquotas e armazenadas em - 80 °C (Benaroudj *et al.*, 1995), o que pode ter favorecido a oligomerização observada. O congelamento ou liofilização da

Hsp70-1A também resultou em diferentes espécies oligoméricas passíveis de separação por CEM (dados não mostrados).

Considerando o objetivo de estudar a influência dos nucleotídeos adenosina na relação estrutura-função da proteína Hsp70-1A humana através de técnicas como SAXS e UAL, por exemplo, e que a utilização de soluções monodispersas facilitariam as análises e interpretação dos resultados, foi necessário padronizar os procedimentos experimentais. Para isso, foi utilizada a fração de 70 kDa da proteína Hsp70-1A recém purificada por CEM, com controle do tempo de estocagem, para evitar a oligomerização. Como experimento de controle de qualidade, foi utilizada a técnica de EDL para avaliar a monodispersidade das soluções de Hsp70-1A.

Shi *et al.* sugerem que condições experimentais de alta força iônica podem induzir a oligomerização da DnaK (Shi *et al.*, 1996). Outros trabalhos indicam que determinados íons monovalentes (potássio e cloreto), divalentes (cálcio ou magnésio) e Pi se ligam à Hsp70, sendo esses importantes para a associação com os nucleotídeos e também para a hidrólise de ATP (revisado por Mayer *et al.*, 2001). Tais íons são encontrados nas estruturas em alta resolução do DLN das Hsp70 (Flaherty *et al.*, 1990; Wilbanks & McKay, 1995; Osipiuk *et al.*, 1999). É interessante ressaltar que a atividade de hidrólise de ATP das Hsp70 é dependente da concentração de íons monovalentes, sendo estimulada por íons potássio, mas não por íons sódio (O'Brien & McKay, 1995), e que os íons sódio não substituem os íons potássio na dissociação de proteínas desnaturadas das Hsp70 (Palleros *et al.*, 1993). A importância dos íons potássio para a função das Hsp70 pode ser ainda exemplificada pela redução da atividade de hidrólise de ATP por um fator de 4 vezes em 25 mmol/L de íons potássio em comparação a 75 mmol/L (Ha & McKay, 1994). Além disso, as Hsc70 necessitam de uma concentração mínima de 10 a 20 mmol/L de íons potássio para apresentarem atividade chaperone (Schlossman *et al.*, 1984) o que reforça a sua dependência deste íon, principalmente se for considerado que no meio intracelular existe uma alta concentração de íons potássio e baixa concentração de íons sódio.

Considerando os fatores discutidos acima, foi preparado um meio tamponante para a Hsp70-1A com o objetivo de controlar a oligomerização, com o mínimo de atividade de hidrólise de ATP, entretanto em condições para a perfeita ligação dos nucleotídeos adenosina. Assim, foi utilizado um tampão com baixa força iônica (NaCl 50 mmol/L) para evitar a oligomerização observada na CEM. O tampão também teve a suplementação de íons potássio em baixa concentração (5 mmol/L) para evitar hidrólise de ATP, visto que esta

é dependente de altas concentrações de potássio. O Pi foi adicionado para estabilizar a ligação de ADP (5 mmol/L) e também reduzir a hidrólise de ATP, pois a dissociação do Pi é uma etapa limitante para hidrólise de ATP (Ha & McKay, 1994). Nos experimentos, na presença de nucleotídeos adenosina, sempre foi adicionado íons magnésio (1 mmol/L) para a perfeita coordenação dos ligantes (ADP ou ATP) com a proteína Hsp70-1A. Desse modo, foi esperado controlar os efeitos de oligomerização observados para a Hsp70-1A, com perfeita ligação dos nucleotídeos, porém com baixa taxa de hidrólise de ATP em ADP e Pi.

5.1.2 MgADP e MgATP aumentaram a estruturação da proteína Hsp70-1A

Os experimentos de CD da Hsp70-1A mostraram que essa proteína apresenta predominância de estrutura secundária do tipo α -hélice. A análise pelo programa CDNN Deconvolution confirmou tal informação e mostrou que a Hsp70-1A é uma proteína constituída por estrutura secundária do tipo α -hélice e folha β -pregueadas. A quantidade de estrutura secundária observada nas estruturas em alta resolução do DLN da Hsp70 humana (código PDB número 1HJO) e do DLS da DnaK de *E. coli* (código PDB número 1DKX) são respectivamente: 37% e 39% de estrutura em α -hélice e 26% e 24% de estrutura em folha β -pregueada. Os dados experimentais observados para a Hsp70-1A estão compatíveis com os dados das estruturas em alta resolução, pois essas últimas não levam em consideração toda a estrutura da proteína e não são estruturas dinâmicas, como são os dados experimentais obtidos.

Ensaio de CD da Hsp70-1A na presença de nucleotídeos adenosina mostraram que o MgADP e o MgATP, mas não o MgAMP, induziram um pequeno aumento de intensidade de luz absorvida (à direita) próximo à região de 208 nm e um aumento maior (à esquerda) na região de 195 nm. Tais mudanças foram muito parecidas para ambos os nucleotídeos: MgADP e MgATP. Essas informações sugerem que pequenas mudanças conformacionais, avaliadas pelo aumento do sinal na estrutura secundária da proteína, foram provocadas pela interação da Hsp70-1A com os nucleotídeos adenosina ADP e ATP.

5.1.3 MgATP e MgADP aumentaram a estabilidade térmica da proteína Hsp70-1A

Ensaio de desnaturação térmica da Hsp70-1A foram monitorados por CD e mostraram a presença de 3 Tms para a proteína. Entretanto, o desenovelamento térmico da Hsp70-1A foi irreversível mesmo quando a amostra foi aquecida somente até 50 °C. Palleros

et al. (1991; 1992) relatam que a Hsc70 bovina também apresenta desenovelamento térmico irreversível quando aquecida a temperaturas superiores a 42 °C, diferentemente da DnaK que é parcialmente reversível (Grimshaw *et al.*, 2001). A adição de ADP ou de ATP nas amostras de Hsp70-1A provocaram aumentos de aproximadamente 4 °C na primeira e terceira Tms, sendo que a segunda Tm não sofreu alterações, e o controle utilizando AMP não mostrou diferenças. Em amostras da Hsp70-1A com a suplementação de íons magnésio aos nucleotídeos ATP e ADP, houve um incremento de 7 °C e 9 °C na Tm₁, respectivamente, porém, novamente sem mudanças na Tm₂. A Tm₃ também não sofreu alterações em relação àqueles aumentos observados apenas pela adição de ADP e ATP. Dados similares são relatados para a Hsc70 bovina na presença de MgATP e MgADP, sendo que o primeiro Tm está conexo à agregação da proteína em espécies oligoméricas (Palleros *et al.*, 1991; 1992). Os dados sugerem que os nucleotídeos ADP e ATP interagiram com a Hsp70-1A e provocaram mudanças conformacionais na estrutura da Hsp70-1A. Essa interação resultou em um aumento da estabilidade de um dos seus domínios, sendo que na presença de íons magnésio esse aumento foi mais evidenciado. Os íons magnésio podem ser necessários para a perfeita interação e coordenação dos nucleotídeos ADP e ATP com a Hsp70-1A, indicando que interações específicas entre a Hsp70-1A e os nucleotídeos na presença de magnésio. Considerando as diferenças observadas, a Tm₁ pode ser relacionada ao desenovelamento parcial do DLN que foi desnaturado completamente na Tm₃, visto que essas Tms sofreram influência da presença dos nucleotídeos ADP e ATP. A Tm₂ pode estar relacionada ao desenovelamento parcial do DLS, pois não sofreu mudanças pela presença dos nucleotídeos e íons magnésio. A mudança do sinal de CD referente à Tm₂ não foi muito intensa, o que também sugere que esta Tm pode estar relacionada ao desenovelamento do DLS1. Este subdomínio é rico em estrutura secundária do tipo folha β -pregueada e o sinal de CD pode ter sido “mascarado” no λ de 222 nm, devido a baixa intensidade de sinal de CD absorvido pelas estruturas do tipo β -pregueada neste λ .

A desnaturação por temperatura monitorada por DSC da DnaK e de seus fragmentos proteolíticos, também mostra a existência de 3 Tms na desnaturação térmica desta proteína (Montgomery *et al.*, 1999). O perfil do desenovelamento térmico da DnaK sugere que a primeira Tm está relacionada ao desenovelamento de grande parte do DLN; a segunda Tm com a maior parte do DLS, e a terceira Tm está relacionada a componentes de ambos os domínios (Montgomery *et al.*, 1999). Os dados observados para a desnaturação térmica da Hsp70-1A estão condizentes com as informações obtidas para a desnaturação térmica da

DnaK. Experimentos de DSC para a Hsp70-1A também foram realizados, mas foi observada uma extensa variação na capacidade calorífica provavelmente devido à agregação que, juntamente com a irreversibilidade do sistema, inviabilizaram as análises. Os efeitos observados no T_m da Hsp70-1A sugerem que a ligação do nucleotídeo adenosina provocou uma compactação do DLN da Hsp70-1A, o que aumentou a sua estabilidade térmica, sendo que esse efeito foi maior na presença de MgADP do que MgATP, e sugerindo que estes ligantes provocaram diferentes mudanças conformacionais na estrutura da Hsp70-1A.

5.1.4 Experimentos de fluorescência do triptofano da proteína Hsp70-1A

A presença de estrutura terciária da Hsp70-1A foi parcialmente mostrada por experimentos de fluorescência do triptofano. A Hsp70-1A possui dois resíduos de triptofano em sua estrutura primária e a figura 40 representa a estrutura em alta resolução do DLN de uma Hsp70 (código PDB número 1HJO) que possui 100% de identidade naquele domínio com a Hsp70-1A. O W90 da Hsp70-1A está localizado em uma alça enterrada no DLN1 e o resíduo de triptofano W580 da Hsp70-1A foi avaliado, por homologia, para estar localizado no subdomínio DLS2, em uma região de hélice-alça-hélice e relativamente exposto ao solvente (dados não mostrados). O espectro de emissão de fluorescência da Hsp70-1A mostrou um máximo de emissão de fluorescência em 336 nm e um tênue ombro localizado em 355 nm, mesmo máximo de emissão de fluorescência observado para a proteína em condições desnaturantes. Esses dados sugerem que um dos resíduos de triptofano da Hsp70-1A deve estar protegido do meio aquoso, enquanto o outro pode estar exposto ao solvente. A adição de MgATP na amostras da Hsp70-1A (mas não de MgADP ou de MgAMP) provocou uma redução na intensidade de emissão de fluorescência de aproximadamente 7%, considerando a área da integral sob a curva. Considerando a capacidade de ligação do MgATP ao DLN, juntamente com a estrutura em alta resolução representada na figura 40a (código PDB número 1HJO), é possível concluir que o resíduo W90 deve estar enterrado na estrutura da Hsp70-1A e foi sensível à ligação de MgATP.

Os dados de emissão de fluorescência do triptofano para a DnaK de *E. coli* mostram que, na presença de MgATP, a intensidade de emissão de fluorescência é reduzida em aproximadamente 15% quando comparado com a DnaK sem adição de nucleotídeos (Theyssen *et al.*, 1996; Montgomery *et al.*, 1996). Já na presença de MgADP, a intensidade do espectro de emissão fluorescência da DnaK é aumentada em aproximadamente 7%

(Theyssen *et al.*, 1996; Banecki *et al.*, 1992). Um outro dado observado para a DnaK é a mudança do λ do máximo de emissão de fluorescência induzido pela ligação de MgATP (de 345 para 342 nm) (Theyssen *et al.*, 1996). Os dados observados para a Hsp70-1A não mostraram mudanças no λ de emissão máxima de fluorescência ou no centro de massa espectral induzido pela ligação dos nucleotídeos adenosina. O resíduo W90 da Hsp70-1A deve estar suficientemente enterrado e protegido na estrutura do DLN para sofrer mais modificações no λ de emissão máxima de fluorescência. A adição de MgADP não induziu mudanças na intensidade de emissão de fluorescência nas amostras de Hsp70-1A, como é observado para a DnaK (Theyssen *et al.*, 1996; Banecki *et al.*, 1992). Uma possível explicação para esses dados aparentemente contraditórios é a de que na DnaK, o resíduo W102 (único resíduo de triptofano na DnaK) está localizado em uma folha β -pregueada e é parcialmente exposto para o ambiente hidrofílico. Enquanto o seu equivalente, sensível à ligação de MgATP na Hsp70-1A, o W90, deve estar em uma alça protegida no meio aquoso. Assim, as mudanças conformacionais induzidas pela ligação de ADP não surtiram efeitos na emissão de fluorescência da Hsp70-1A.

5.1.5 Calorimetria da interação da proteína Hsp70-1A com os nucleotídeos adenosina

A afinidade da proteína Hsp70-1A por MgADP ou MgATP

A presença de nucleotídeos adenosina influencia a afinidade das Hsp70 por substratos e, portanto, a afinidade das Hsp70 pelos nucleotídeos adenosina é um fator importante para a função das Hsp70 (Mayer *et al.*, 2001). Para determinar parâmetros termodinâmicos como a K_A , foram realizados ensaios de ITC para a Hsp70-1A com seus ligantes, MgATP ou MgADP, na presença de íons fosfato e potássio. Os dados de K_A obtidos indicam que a proteína Hsp70-1A apresentou alta afinidade pelos ligantes ADP e ATP e estão condizentes com a hipótese de que a interação com ATP não é o fator limitante do ciclo de ação das Hsp70. Em condições fisiológicas, as Hsp70 estão saturadas por ATP devido à alta concentração intracelular desse nucleotídeo (Mayer *et al.*, 2001). A K_A da Hsp70-1A por MgATP foi aproximadamente 3 vezes menor do que por MgADP e sugerem K_{Ds} da ordem submicromolar (se for considerado a relação $K_D = 1/K_A$ válida para a interação da Hsp70-1A com os nucleotídeos adenosina), como observado para outras Hsp70 (Mayer *et al.*, 2001). Porém, os dados de K_D para a Hsc70 bovina (que apresenta 95% de identidade com a Hsp70-1A) por MgATP e por MgADP, sugerem uma maior afinidade por MgATP do

que por MgADP (O'Brien & McKay, 1995). Essas disparidades podem ser justificadas pelas diferentes metodologias e temperaturas utilizadas para tais determinações. A afinidade da Hsc70 bovina pelos nucleotídeos foi feita por técnicas radioisotópicas, em temperatura de 37 °C e em condições de equilíbrio (O'Brien & McKay, 1995). A K_A da Hsp70-1A por ADP e ATP foi determinada utilizando a técnica de ITC a 20 °C, sendo que essa determinação não pode ser considerada em condições de equilíbrio. Também é preciso considerar que o agente tamponante utilizado, o Tris-HCl, apresenta uma alta energia de ionização que pode levar a uma superestimação das constantes termodinâmicas determinadas (Jelesarov & Bosshard, 1999). Deve-se avaliar se a interação entre a Hsp70-1A e os nucleotídeos provoca alguma ionização na proteína ou no ligante que pode ser tamponada pelo meio, excluindo a possível influência da energia de ionização do Tris-HCl. Um outro fator a ser considerado é a presença de íons fosfato, pois este é importante para reduzir a taxa de liberação de ADP pelo DLN, o que pode provocar um aumento na K_A das Hsp70 por ADP (Ha & McKay, 1994), principalmente em condições de saturação de íons fosfato. Palleros *et al.* (1991) também reportam uma maior afinidade da Hsc70 bovina por ADP do que por ATP, possivelmente porque as constantes de afinidade foram determinadas na presença de 20 mmol/L de íons fosfato. Ha & McKay (1994) estimam a K_D da Hsc70 bovina por íons Pi como sendo da ordem de 1 mmol/L e sugerem um mecanismo de dissociação do complexo Hsp70-ADP-Pi no qual o íon Pi se desligaria primeiro e em seguida o ADP se desligaria. Assim, com a afinidade das Hsp70 por íons Pi, existe a possibilidade desses também inibirem a ligação de ATP à Hsp70-1A, pois os experimentos foram realizados em condições de saturação de íons Pi.

Porém, os dados termodinâmicos obtidos nortearam as condições experimentais para os ensaios de UAL (a seguir), na qual a condição de saturação deve ser satisfeita com um mínimo de interferência dos nucleotídeos na absorbância da proteína, mesmo considerando pequenas taxas de contaminação das amostras de ATP com ADP e *vice versa*. Os dados também indicam que a K_A da Hsp70-1A é maior por MgADP, e que a hidrólise do ATP, nas condições experimentais utilizadas, foi inexistente ou desprezível, pois a ΔH de interação da Hsp70-1A com o ADP foi mais negativa do que a da interação da proteína com o ATP.

ΔH e ΔS da interação entre a proteína Hsp70-1A e MgATP ou MgADP

Mayer *et al.* (2000) propõem um mecanismo de ação das Hsp70, o qual envolve a modulação dos domínios da proteína, dependente do nucleotídeo adenosina ligado ao DLN, o que resultaria no deslocamento do equilíbrio conformacional do subdomínio DLS2. Uma análise dos dados de ΔH e ΔS obtidos para a interação entre a Hsp70-1A e os nucleotídeos adenosina indicam que estas constantes termodinâmicas podem estar condizentes com a hipótese proposta por Mayer *et al.* (2000). Se forem analisados conjuntamente, os dados obtidos para as titulações da Hsp70-1A com o MgADP e o MgATP, é possível observar que a interação com o MgADP apresentou ΔH mais negativo do que a interação com o MgATP e o inverso pode ser constatado para a ΔS . O mecanismo proposto por Mayer *et al.* (2000) sugere que, na presença de ADP, o equilíbrio conformacional do DLN2 estaria deslocado para a conformação fechada sobre o substrato. Rearranjos provocados pela mudança conformacional do DLS2 de “aberta” para “fechada”, induzida pelo ADP, poderiam levar a um empacotamento da molécula resultando na liberação de energia entálpica em detrimento à entrópica (Jelesarov & Bosshard, 1999). Os dados de desenovelamento térmico da Hsp70-1A, que sugerem um possível empacotamento maior da proteína na presença de MgADP do que de MgATP, e fortalecem essa hipótese. Na presença de ATP, como sugerido por Mayer *et al.* (2000), o DLS2 estaria na posição “aberta” apresentando uma maior mobilidade e favorecendo a entropia em detrimento à entalpia (Jelesarov & Bosshard, 1999).

É necessário salientar que os dados obtidos pela técnica de ITC representam o somatório de todos os eventos provocados pela interação da Hsp70-1A com os nucleotídeos adenosina, o que dificultou a avaliação das hipóteses apresentadas. Porém, uma análise utilizando a ΔC_p , que pode ser obtida a partir da titulação em diferentes temperaturas, pode auxiliar na interpretação desses dados. A ΔC_p está relacionada com o grau de exposição de superfície hidrofóbicas e hidratação das partículas (Pierce *et al.*, 1999; Jelesarov & Bosshard, 1999). Deste modo, os dados termodinâmicos obtidos para a Hsp70-1A sugerem a potencialidade da utilização de técnicas calorimétricas para fornecer mais informações sobre o mecanismo de ação de proteínas como as Hsp70, principalmente, na presença de um possível substrato. Assim, outros estudos se fazem necessários para aprimorar as discussões e verificar a veracidade da hipótese proposta acima.

5.1.6 MgATP e MgADP provocaram mudanças conformacionais na proteína Hsp70-1A

Análise hidrodinâmica da proteína Hsp70-1A

A técnica de UAL foi utilizada para avaliar o comportamento da Hsp70-1A em solução frente a processos de auto-associação e à mudanças conformacionais induzidas pelos nucleotídeos adenosina. Os ensaios de sedimentação em equilíbrio mostraram que a proteína se comportou como uma partícula de aproximadamente 70 kDa, sendo que os dados foram ajustados com qualidade para uma única partícula em solução nas condições experimentais testadas (tampão TKP). Os ensaios de velocidade de sedimentação também indicaram que a Hsp70-1A comportou-se como uma única partícula em solução. Schönfeld *et al.* descrevem a presença de duas espécies em ensaios de velocidade de sedimentação e também de sedimentação em equilíbrio para a proteína DnaK, podendo esses últimos dados ser também ajustados para um sistema de três espécies (Schönfeld *et al.*, 1995). Porém, Schönfeld *et al.* utilizaram uma concentração salina alta, na qual a oligomerização da proteína é favorecida (Shi *et al.*, 1996). A dependência da DnaK em relação ao tempo de estocagem e concentração de proteínas em equilíbrio com a fração de 70 kDa também é relatada por Schönfeld *et al.* (1995). Benaroudj *et al.* (1995; 1996; 1997) reportam o estudo de uma Hsc70 de mamífero por ultracentrifugação analítica mostrando a existência, em equilíbrio lento, de pelo menos 3 espécies da Hsc70: monômero, dímero e trímero. Porém, estes autores utilizaram amostras congeladas da Hsc70, condição que favoreceu a oligomerização da Hsp70-1A (dados não mostrados), e que a oligomerização foi concentração e tempo-dependente. Benaroudj *et al.* relatam que a fração de 70 kDa da Hsc70, purificada do estoque congelado por CEM, possui um $s_{20,w}$ de $4,46 \pm 0,01$ S, correspondente a uma partícula de aproximadamente 80 kDa (Benaroudj *et al.*, 1996). O $s_{20,w}^0$ de $4,09 \pm 0,01$ S determinado para a Hsp70-1A é inferior ao relatado por Benaroudj *et al.* (1996) e pode ser devido às condições de força iônica utilizada nas diferentes determinações. Nas amostras de Hsc70, Benaroudj *et al.* (1996) utilizaram 100 mmol/L de cloreto de potássio e na ausência de íons fosfato, enquanto para a Hsp70-1A foi utilizada uma condição de força iônica inferior aquela na presença de 5 mmol/L de íons K^+ e P_i .

Considerando o sistema analisado por Schönfeld *et al.* como de duas espécies, as amostras de DnaK apresentaram $s_{20,w}$ de 4,0 e 6,3 S, que correspondem a proteínas de aproximadamente 85 e 185 kDa (Schönfeld *et al.*, 1995). Os dados de velocidade de sedimentação obtidos para as amostras de Hsp70-1A sugerem um valor de $s_{20,w}^0$ de

aproximadamente 4,09 S, valor próximo ao observado para a DnaK, principalmente se for considerado a maior *MM* para a Hsp70-1A em relação à DnaK (72 *versus* 69 kDa, respectivamente). Considerando os valores de $D_{20,w}^0$ determinados pelos ensaios de EDL, o valor da *MM* calculado da relação *s/D* foi de aproximadamente 70 kDa e enfatiza a qualidade dos dados obtidos para a Hsp70-1A. Indiretamente, a não observação de outras espécies para a Hsp70-1A indicou que a força iônica utilizada e o controle do tempo de utilização das amostras foram efetivos em prevenir a poli-dispersidade e a oligomerização tempo-dependente observada nas CEM.

Como parcialmente discutido acima, os ensaios de ITC nortearam as condições de saturação pelos nucleotídeos adenosina que foram utilizados nos ensaios de velocidade de sedimentação. A concentração de nucleotídeos adenosina utilizada (200 $\mu\text{mol/L}$) foi aproximadamente 13 vezes maior que a concentração de Hsp70-1A nas amostras mais concentradas (1 mg/mL – 14 $\mu\text{mol/L}$), fornecendo uma boa condição de saturação. Nessas condições foi possível executar os experimentos de velocidade de sedimentação com uma relação sinal-ruído adequada, em comprimentos de onda entre 230 e 238 nm.

As análises dos experimentos de velocidade de sedimentação sugerem que a ligação dos nucleotídeos MgATP e MgADP na Hsp70-1A provocaram pequenas mudanças conformacionais que foram perceptíveis à técnica e que foi possível interpretá-las através do cálculo do $s_{20,w}^0$. O $s_{20,w}$ é dependente da concentração da proteína devido ao aumento da viscosidade do sistema provocado pelo aumento da concentração protéica (Laue, 2001). Os dados obtidos indicam que a presença de MgADP induziu pequenas mudanças conformacionais, resultando em um $s_{20,w}^0$ um pouco superior àquele observado na ausência do nucleotídeo ($s_{20,w}^0 \text{ ADP} = 4,21 \text{ S}$). Já na presença do nucleotídeo MgATP ($s_{20,w}^0 \text{ ATP} = 4,44 \text{ S}$), a mudança conformacional foi mais evidente em relação ao estado livre de nucleotídeos e na presença de MgADP. Benaroudj *et al.* (1995) também relatam que amostras de Hsc70 contendo ATP apresentam um *s* maior do que na ausência de ATP (4,9 S *versus* 4,3 S, respectivamente). Considerando os valores de $D_{20,w}^0$ determinados para a Hsp70-1A na presença de MgADP e MgATP e a relação *s/D*, os valores de *MM* determinados para a Hsp70-1A na presença de nucleotídeos estão próximos aos da partícula monomérica em solução (71 kDa). Os dados hidrodinâmicos obtidos sugerem a existência de pelo menos três diferentes estados conformacionais na presença de nucleotídeos adenosina, assumidos pela proteína Hsp70-1A: o estado nativo na ausência de nucleotídeos, um estado na presença de ADP e outro na presença de ATP. É bom enfatizar que esses dados

representam a média conformacional do sistema, pois a técnica de velocidade de sedimentação fornece dados referentes à média temporal das diversas conformações da proteína (Waxman *et al.*, 1994).

Análise por SAXS da proteína Hsp70-1A na ausência e na presença de MgATP ou MgADP

Os experimentos de SAXS para a Hsp70-1A sugerem, pela análise do perfil da curva da $p(r)$, que essa proteína possui uma forma prolata alongada (Feigin & Svergun, 1987). Entretanto, a adição dos nucleotídeos adenosina (MgATP ou MgADP) às amostras de Hsp70-1A não provocou grandes mudanças na forma da Hsp70-1A, como analisado pelo perfil das curvas da $p(r)$ (Feigin & Svergun, 1987). Esses dados não são condizentes com dados publicados para a Hsc70 bovina (Wilbanks *et al.*, 1995) e DnaK (Shi *et al.*, 1996). A Hsc70 bovina na presença de MgATP apresenta uma curva da $p(r)$ similar à obtida para a Hsp70-1A, possuindo uma forma prolata alongada (Wilbanks *et al.*, 1995). Para as amostras de Hsc70 na presença de ADP, é mostrada uma considerável modificação no perfil da curva da $p(r)$ em relação à amostra com ATP (Wilbanks *et al.*, 1995). Na presença de MgADP, a curva da $p(r)$ da Hsc70 bovina apresentou 2 máximos bem definidos, o que sugere uma partícula espalhadora com dois lóbulos relativamente separados (Wilbanks *et al.*, 1995; Svergun & Koch, 2003). Para os experimentos da DnaK (Shi *et al.*, 1996), apenas resultados para a proteína sem aditivos e com MgATP estão disponíveis e, nessas condições, a curva da $p(r)$ também apresentou um perfil multi-lobular com consideráveis diferenças provocadas pela presença de MgATP. Nos trabalhos de Wilbanks *et al.* (1995) e Shi *et al.* (1996), as principais diferenças em relação aos dados obtidos para a Hsp70-1A estão na constituição do solvente; o que pode justificar as diferenças observadas. Os dados apresentados por Wilbanks *et al.* (1995) para a Hsc70 bovina foram obtidos em tampão MOPS (10 mmol/L) contendo 150 mmol/L de KCl. Nessas condições, seria favorecida a hidrólise de ATP (Ha & McKay, 1994), que foi estimada pelos autores em aproximadamente 5% (Wilbanks *et al.*, 1995), e a oligomerização da proteína (Shi *et al.*, 1996). Wilbanks *et al.* (1995) considera a oligomerização como um fator para justificar as mudanças no perfil da curva da $p(r)$ na presença de MgADP. No trabalho de Shi *et al.* (1996) para a DnaK, as condições utilizadas são ótimas para controlar a oligomerização observada em alta força iônica. Assim, os dados apresentados para a DnaK, com o objetivo de monitorar o efeito da adição de ATP, foram obtidos em 1 mmol/L de Tris-HCl, e somente nessas condições Shi *et al.* (1996) considera a

proteína no estado monomérico. Para realmente ser possível usar a constituição do solvente para justificar as diferenças observadas no perfil da curva da $p(r)$ da Hsp70-1A, dever-se-ia realizar ensaios de SAXS em condições similares às aquelas realizadas pelos autores do trabalho de Wilbanks *et al.*, (1995) com a Hsc70 bovina, uma proteína com alta identidade com a Hsp70-1A. Porém, essas discrepâncias sugerem a importância dos íons na conformação da proteína Hsp70-1A.

Os valores do R_g determinados para a Hsc70 bovina (Wilbanks *et al.*, 1995) e DnaK (Shi *et al.*, 1996) sofrem variação devido à adição dos nucleotídeos adenosina nas amostras de proteínas. Entretanto, os valores do R_g calculados a partir das curvas das $p(r)$ para a Hsp70-1A não foram variáveis. O R_g está relacionado a mudanças conformacionais, pois ele está conexo ao centro de massa inercial dos centros eletrônicos da partícula espalhadora e assim ao perfil da curva da $p(r)$. Como as curvas das $p(r)$ da Hsp70-1A foram similares nas diferentes condições experimentais estudadas, variações no R_g da Hsp70-1A não foram observadas. Entretanto, mesmo considerando essas diferenças no perfil da curva da $p(r)$, o D_{max} observado para as Hsc70 bovina, DnaK e Hsp70-1A, foram entre 112 a 130 Å (Wilbanks *et al.*, 1995; Shi *et al.*, 1996). A importância dessa informação será mais bem discutida abaixo, juntamente com os dados das GrpEs humanas (item 5.4.4 de Discussão).

Os resultados de SAXS da proteína Hsp70-1A foram obtidos em condições de baixa força iônica para controlar a oligomerização tempo-dependente, em baixa concentração de íons K^+ para prevenir a hidrólise de MgATP, e na presença de íons fosfato (item 5.1.2 de Resultados). Nessas condições experimentais, a Hsp70-1A não sofreu processo de oligomerização ou agregação, em baixas concentrações de proteína, conforme avaliado por UAL (item 5.1.7 de Resultados). A hidrólise de ATP também foi controlada, como indiretamente observado pelos ensaios de desenovelamento térmico monitorado por CD e pelos experimentos de ITC (itens 5.1.4 e 5.1.6 de Resultados, respectivamente). Porém, os dados de SAXS obtidos para a Hsp70-1A na presença de MgATP apresentaram indícios de agregação que podem interferir nos resultados finais obtidos, tanto que a estimativa da MM da Hsp70-1A na presença de MgATP foi de aproximadamente 100 kDa contra 77 kDa nas demais condições (ausência de nucleotídeos e na presença de MgADP). Essa agregação foi provavelmente devido à alta concentração utilizada nos experimentos de SAXS.

Os modelos ab initio da Hsp70-1A nas diferentes condições experimentais

A partir dos dados de SAXS, foram construídos modelos *ab initio* nas diferentes condições experimentais. Os modelos foram construídos a partir da superposição de 10 modelos gerados de maneira independente, para assim, minimizar os erros inerentes ao processo de modelagem e ajuste da curva (Svergun & Koch, 2003). Esses erros podem ser resultantes de diferentes modelos que se ajustam à mesma curva da $I(q)$. A análise da curva da $I(q)$ dos modelos finais, pelo ajuste nas curvas das $I(q)$ obtidas experimentalmente, mostra a qualidade dos modelos obtidos.

A análise hidrodinâmica dos modelos pelo programa HydroPro também sugere que os modelos *ab initio* gerados para a Hsp70-1A na ausência de nucleotídeos e na presença de MgADP foram mais condizentes com os dados experimentais. Tanto na ausência de nucleotídeos adenosina quanto na presença de MgADP, os parâmetros hidrodinâmicos experimentais ($s_{20,w}^0$ e $D_{20,w}^0$) e os estimados para os modelos *ab initio* foram bem condizentes, o que enfatiza a qualidade desses modelos *ab initio*. O modelo *ab initio* para a Hsp70-1A na presença de MgATP apresentou discordância entre os parâmetros hidrodinâmicos ($s_{20,w}^0$ e $D_{20,w}^0$) avaliados pelo programa HydroPro e os dados obtidos experimentalmente. Esta discordância corrobora a hipótese de que a agregação prejudicou, em parte, a construção do modelo *ab initio* da Hsp70-1A na presença de MgATP. Os experimentos de SAXS para a proteína Hsp70-1A na presença de MgATP devem ser repetidos para avaliar os efeitos da suposta agregação.

Superposição dos modelos ab initio da Hsp70-1A com estruturas em alta resolução dos domínios das Hsp70

Os modelos *ab initio* construídos a partir dos dados de SAXS foram superpostos com as estruturas em alta resolução de domínios da proteína Hsp70 disponíveis no PDB. O modelo *ab initio* construído para a Hsp70-1A na presença de MgATP também foi utilizado, mesmo considerando que ele pode não ser um bom modelo e que existe a necessidade de novos experimentos de SAXS a fim de avaliar a influência da agregação observada.

A estrutura em alta resolução do domínio DLN que foi utilizada para sobrepor aos modelos *ab initio* da Hsp70-1A possui 100% de identidade, na seqüência de aminoácidos, com a Hsp70-1A na seqüência primária do DLN. Esta estrutura foi obtida a partir de um cristal do DLN formando complexos com ADP (código PDB número 1HJO – Osipiuk *et al.*,

1999). Também foi utilizada uma estrutura em alta resolução do DLS da DnaK de *E. coli* formando complexos com um peptídeo substrato (código PDB número 1DKX – Zhu *et al.*, 1996) e que apresenta somente 46% e 61%, de identidade e similaridade, respectivamente, entre os aminoácidos 411 a 635 com o DLS da Hsp70-1A. Existem modelos em alta resolução do DLS de proteínas oriundas de mamíferos e que possuem maior similaridade com a Hsp70-1A na região do DLS, entretanto o modelo do DLS da DnaK foi escolhido porque ele cobre uma seqüência de aminoácidos maior do que os demais modelos disponíveis.

A superposição do modelo *ab initio* obtido para a Hsp70-1A na ausência de aditivos acomodou satisfatoriamente as estruturas em alta resolução representadas pelos DLN e DLS. Se for considerado que os dados hidrodinâmicos experimentais e preditos para o modelo *ab initio* foram compatíveis, é possível concluir que este modelo *ab initio* representa uma das principais conformações da Hsp70-1A na ausência de aditivos. O modelo *ab initio* da Hsp70-1A na presença de MgATP, mesmo sem apresentar boa concordância com os dados hidrodinâmicos experimentais, apresentou uma superposição satisfatória com as estruturas em alta resolução. Entretanto, o modelo *ab initio* da Hsp70-1A na presença de MgADP não apresentou uma superposição adequada com tais modelos em alta resolução, principalmente na região do DLN e na região central do modelo *ab initio*. Essa superposição insatisfatória ocorreu, ainda que os modelos em alta resolução representem as estruturas na presença de ligantes (ADP e substrato) e que os dados hidrodinâmicos avaliados, para o modelo *ab initio* da Hsp70-1A na presença de MgADP, foram concordantes com os dados hidrodinâmicos experimentais. Uma possível explicação para tal fato, pode ser a influência da presença dos demais íons na conformação geral da proteína e a ausência do substrato ligado à Hsp70-1A. Como discutido acima, o perfil da curva da $p(r)$ da Hsc70 bovina sofreu uma grande mudança devido à presença de MgADP (Wilbanks *et al.*, 1995).

Apesar destas considerações, de forma geral, os modelos *ab initio* construídos para a Hsp70-1A acomodaram adequadamente os modelos em alta resolução dos domínios de proteínas Hsp70. Porém, dependendo da condição experimental, a posição relativa desses domínios foi modificada para melhorar a acomodação no modelo *ab initio*. Isso sugere que a ligação do nucleotídeo adenosina ao DLN provocou mudanças conformacionais na Hsp70-1A sem resultar em mudanças no formato geral da proteína. Como a resolução dos dados de SAXS foi de aproximadamente 25 Å, os detalhes atômicos dessas mudanças conformacionais não podem ser evidenciados. Também deve ser considerado que tais

modelos em alta resolução foram obtidos de domínios em uma única conformação espacial, como por exemplo, o DLS ligado a um substrato, o qual, talvez, não seja a melhor conformação para a superposição da proteína na presença de MgATP. Da mesma forma, o DLN foi cristalizado em complexo com o ADP e pode não representar a melhor conformação do modelo *ab initio* sem aditivos e com ATP. Um outro fator que deve ser levado em consideração é que os modelos em alta resolução (código PDB número 1HJO e 1DKX) não representam a proteína Hsp70-1A completa, pois faltam aminoácidos em ambas as extremidades amino e carboxi-terminal. Assim, é possível que exista densidade eletrônica em excesso representada no modelo *ab initio* da Hsp70-1A que não possui equivalente nos modelos em alta resolução. Apesar das considerações apresentadas, as análises da superposição dos modelos *ab initio* com os modelos em alta resolução sugerem que mesmo sem modificar o formato geral da Hsp70-1A, os nucleotídeos adenosina podem induzir uma reorientação dos domínios da proteína. Essa hipótese sugere que os nucleotídeos adenosina provocam uma série de pequenas mudanças conformacionais que são responsáveis pela reorganização dos domínios e pela ligação ou liberação do substrato.

5.1.7 Considerações finais acerca da proteína Hsp70-1A

Os resultados obtidos para a Hsp70-1A sugerem que a proteína foi expressa em quantidade e purificada como uma fração monomérica, sendo que todos os experimentos apresentados foram realizados com essa fração. A Hsp70-1A apresentou indícios de oligomerização tempo e concentração-dependente, sendo essa parcialmente controlada pela solução preparada e pelo tempo de utilização da amostra após armazenagem. A proteína apresentou estrutura secundária característica para proteínas α/β estruturadas, condizentes com os dados da literatura, e com indícios de ganho de estrutura devido à adição de MgATP ou MgADP. Os ensaios de desenovelamento térmico também mostraram um ganho de estabilidade térmica que pode ter sido ocasionada pelo ganho de estrutura secundária ou por uma maior compactação provocada pela ligação do nucleotídeo no DLN; sugeriram ainda, que o tratamento com a enzima fosfatase alcalina ácida proporcionou uma amostra de Hsp70-1A sem contaminação pelos nucleotídeos adenosina. Dados de ITC mostraram que a Hsp70-1A possui maior afinidade por MgADP do que por MgATP, nas condições testadas, sendo que as constantes de dissociação podem ser da ordem submicromolar, como mostrado para outras Hsp70. Indiretamente, tanto os experimentos de desenovelamento

térmico e de ITC sugerem a ausência da hidrólise de ATP (ou que esta foi desprezível) e a ausência de nucleotídeos contaminantes naquelas condições experimentais – uma condição necessária para o estudo conformacional da proteína em equilíbrio com os nucleotídeos. A maior afinidade da Hsp70-1A por MgADP pode ser explicada pela presença dos aditivos no tampão como, por exemplo, íons fosfato e potássio. Essas informações foram importantes para padronizar a quantidade de nucleotídeos nos experimentos de ultracentrifugação analítica e garantir uma condição de saturação dos nucleotídeos em relação à proteína Hsp70-1A. Os experimentos de UAL mostraram a ausência de oligomerização ou agregação nas condições experimentais utilizadas e que os nucleotídeos (MgADP ou MgATP) provocaram pequenas mudanças conformacionais na Hsp70-1A. Os experimentos de SAXS indicaram que os nucleotídeos MgATP e MgADP provocam pequenas mudanças conformacionais na Hsp70-1A, porém sem modificar o formato da partícula em solução. Esses dados sugerem que pequenas mudanças conformacionais ocasionadas pelos nucleotídeos adenosina resultam da reorganização dos subdomínios e domínios da Hsp70-1A.

Considerando o mecanismo funcional das Hsp70 proposto por Mayer *et al.*, (2000), apresentado na figura 1, os dados obtidos para a Hsp70-1A sugerem que a abertura ou fechamento da “tampa” (representada pelo DLS2) na presença de ATP e ADP, respectivamente, na realidade pode ser uma reorganização conformacional da proteína devido a uma reorientação do DLS2 que poderia expor o sítio de ligação de substratos das Hsp70. A ausência da “tampa” não prejudica a ligação de substratos ao DLS1 e nem a auto-regulação do DLN da DnaK (Pellecchia *et al.*, 2000) e também não muda o formato geral da Hsc70 bovina em relação à proteína completa (Wilbanks *et al.*, 1995). Estas informações enfatizam a necessidade de uma reorganização conformacional do DLS1 para ligar o substrato e dá suporte à hipótese de reorientação do DLS2 sem grandes mudanças conformacionais na estrutura das Hsp70.

Devido ao fato das Hsp70 serem proteínas modulares, constituídas por dois domínios que interagem e possuem um sistema de auto-regulação, é difícil obter um cristal protéico para estudar tais mudanças conformacionais em nível atômico (Mayer *et al.*, 2001). No entanto, ensaios de modelagem molecular direcionada por homologia e dinâmica molecular podem ajudar a obter um melhor ajuste entre um modelo em alta resolução construído por homologia e os modelos *ab initio* obtidos a partir dos dados de SAXS. A utilização de dados estruturais em baixa resolução, obtidos por técnicas como SAXS ou microscopia eletrônica,

podem auxiliar na construção de modelos *ab initio* que expliquem o formato geral da partícula. Estas informações também podem explicar algumas mudanças conformacionais sofridas por proteínas em solução que não são explicadas pelas estruturas em alta resolução (Svergun & Koch, 2003). Já estão disponíveis programas computacionais que realizam a superposição da estrutura em alta resolução no modelo *ab initio* gerado por SAXS ou microscopia eletrônica com um mínimo de discrepância (Tama *et al.*, 2004). Esses programas são capazes de flexibilizar a estrutura em alta resolução para melhor ajustar aos modelos em baixa resolução e podem ajudar a explicar muitas diferenças conformacionais vistas entre os modelos *ab initio* e modelos cristalográficos (Svergun & Koch, 2003). Deste modo, em conjunto com as informações de baixa resolução, poder-se-ia propor modelos que expliquem melhor as propriedades da Hsp70-1A e outras Hsp70. Também poderão auxiliar na identificação de aminoácidos na superfície de contato entre o DLN e DLS que fazem contatos para transferir as informações para a ligação e liberação dos substratos. Com esses modelos gerados por homologia também seria possível propor mutantes sítios-específicos que comprovassem e justificasse o melhor modelo funcional para essa importante maquinaria.

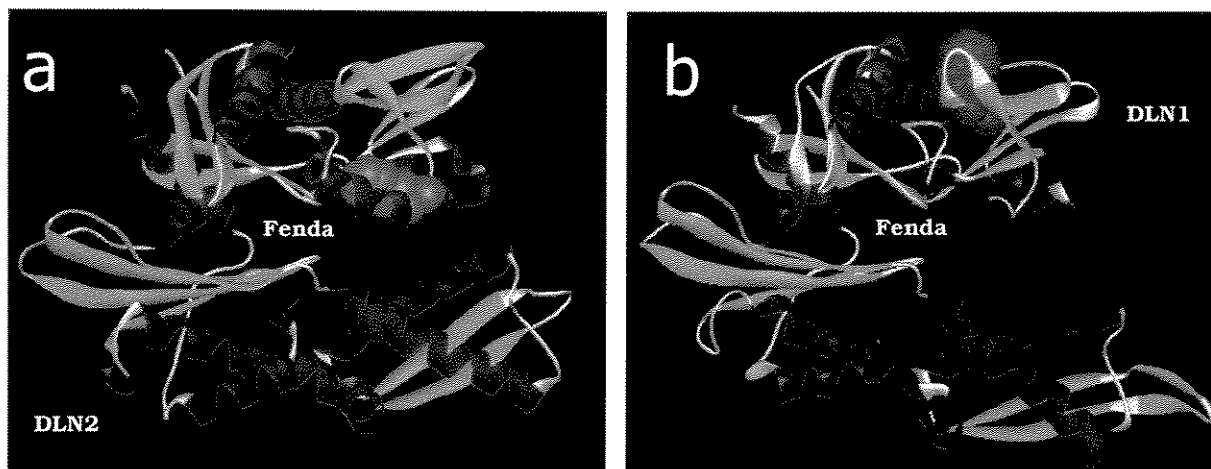


Figura 40: Posicionamento espacial relativo do triptofano localizado no DLN das proteínas Hsp70-1A e DnaK. a) Estrutura em alta resolução (código PDB número 1HJO) do DLN de uma proteína Hsp70 humana, que apresenta 100% de identidade nesse domínio com a Hsp70-1A em estudo, mostrando que seu resíduo de triptofano (W90), representado em azul, está localizado espacialmente em um ambiente protegido do meio. b) Respectivo DLN da proteína DnaK de *E. coli* (código PDB número 1DKG) com o seu único resíduo de triptofano (W102), representado em magenta, em um ambiente relativamente mais exposto ao solvente. São também mostrados os subdomínios DLN1 e DLN2 das Hsp70. O triptofano localizado no DLN da Hsp70-1A parece estar mais protegido do meio do que o seu correspondente na DnaK (veja item 5.1 da Discussão para detalhes). O cristal da proteína representada pela figura a (código PDB número 1HJO – Osipiuk *et al.*, 1999) foi obtido em complexos com ADP. A estrutura apresentada em b (código PDB número 1DKG) representa o DLN da DnaK formando um complexo cristalino com a EcGrpE (não mostrada) em uma conformação ótima para interação com o ATP (Harrison *et al.*, 1997).

5.2 Mortalina

As Hsp70 são classificadas em 3 subfamílias e as proteínas presentes em bactérias e mitocôndria de eucariotos são dependentes do auxílio da co-chaperone GrpE para sua reciclagem (Liberek *et al.*, 1991) devido a características da interface da fenda de ligação de nucleotídeos adenosina (Brehmer *et al.*, 2001). A principal motivação para obter uma representante humana da subfamília DnaK foi estudar a interação dessas com as GrpE humanas. Brehmer *et al.* (2001) mostra que as proteínas GrpE são capazes de interagir apenas com as Hsp70 da subfamília DnaK, influenciando, assim, a troca de ADP por ATP no DLN. Entretanto, Choglay *et al.* (2001) demonstra que a GrpE#1 humana é capaz de interagir fracamente com uma Hsc70 humana citoplasmática. No entanto, a proteína GrpE#1 humana utilizada por Choglay *et al.* (2001) apresenta o peptídeo sinal para endereçamento da proteína para mitocôndria, o qual pode interagir com as Hsp70 citoplasmáticas em sua via de transporte através da membrana (Matouschek, 2003). O sítio de interação das Hsp70 por substratos tem afinidade por regiões hidrofóbicas flanqueadas por resíduos de aminoácidos positivos (Rüdiger *et al.*, 1997), características apresentadas por peptídeos sinais de endereçamento para a mitocôndria (Matouschek, 2003). Os estudos da interação entre as GrpEs humanas e a mortalina, uma Hsp70 humana mitocondrial, são importantes para a caracterização das GrpEs e para compreender o porquê da existência de duas GrpEs em mitocôndria de organismos eucarióticos.

Outros fatores que também motivaram a obtenção da mortalina foi a possibilidade de estudos comparativos com a Hsp70-1A. Por exemplo, nos experimentos de fluorescência de triptofano da Hsp70-1A, que apresentou resultados diferentes aos observados para a DnaK. A mortalina possui apenas um resíduo de triptofano localizado no DLN e, na mesma posição, na sequência de aminoácidos, que o resíduo de triptofano da DnaK. A Hsp70-1A possui dois resíduos de triptofano em sua sequência de aminoácidos, sendo um localizado no DLN e outro no DLS. O resíduo de triptofano localizado no DLN está em uma posição na estrutura terciária diferente da qual está o único resíduo de triptofano da DnaK (Figura 4). Tais estudos comparativos também podem ser monitorados pelas demais técnicas utilizadas para os estudos da Hsp70-1A, como, por exemplo, a ITC, UAL e SAXS, ajudando a compreender o mecanismo de funcionamento dessa importante maquinaria no metabolismo de proteínas.

5.2.1 A proteína mortalina foi obtida como um agregado protéico

A mortalina foi clonada no vetor pET28a e expressa; sua purificação, porém, resultou em possíveis agregados protéicos, como sugerido pelas análises de CEM e pelos ensaios de dicroísmo circular. Diferentes condições de indução e tentativas de re-enovelamento, após desenovelamento por uréia, não resultaram em melhoras para a obtenção da proteína monomérica. Como resultados interessantes podem ser obtidos a partir do estudo da mortalina, novos experimentos devem ser realizados no intuito de obtê-la em sua estrutura nativa. A possibilidade de clonagem do cDNA da proteína mortalina em outros vetores para expressão em fusão com outras caudas para purificação por afinidade, ou na ausência dessas para purificação por troca iônica, poderão ser aplicados no futuro.

5.3 Hsp40

As Hsp40 são as principais co-chaperones das Hsp70. Essas proteínas são capazes de estimular a atividade ATPase das Hsp70 (Hendrick *et al.*, 1993; Laufen *et al.*, 1999) e também possuem atividade chaperone intrínseca, o que as tornam transportadores de substratos para as Hsp70 (Karzai & McMacken, 1996). As Hsp40 apresentam alta diversidade, sendo caracterizadas pela presença de um domínio altamente conservado chamado de domínio J (Cheetham *et al.*, 1998). As Hsp40 são classificadas em 3 subfamílias, segundo a presença de regiões conservadas em suas seqüências de aminoácidos (Cheetham *et al.*, 1998; Mayer *et al.*, 2001). Para estudar as proteínas Hsp40 de origem humana, foram obtidos os cDNAs de duas proteínas pertencentes à subfamília A e B, a DjA1 e a DjB4, respectivamente. Os cDNAs dessas proteínas foram clonados em vetores de expressão em fusão com a cauda de poli-histidinas localizada na extremidade amino-terminal. Como as Hsp40 das subfamílias A e B são proteínas encontradas na forma de dímeros, foi produzido um mutante de deleção da região carboxi-terminal da DjA1 para a caracterização de sua forma monomérica. Todas as proteínas foram bem expressas e os processos de purificação resultaram em proteínas com grau pureza superior a 95%. Os ensaios para a caracterização dessas proteínas e a discussão dos resultados obtidos estão apresentados no manuscrito apresentado no Anexo II (Borges, J.C., Fischer, H., Craievich, A.F. & Ramos, C.H.I. (2004) Low-resolution structural study of two human Hsp40 chaperones in solution. DjA1 from subfamily A and DjB4 from subfamily B, have different quaternary structures. *J. Biol. Chem.*, submetido). De maneira geral, as análises estruturais das Hsp40 humanas mostraram que as proteínas DjA1 e DjB4, das subfamílias A e B, respectivamente, possuem diferenças na estrutura quaternária e na atividade de chaperone molecular.

5.4 GrpE

As proteínas GrpE são fatores de troca de nucleotídeos adenosina do sistema chaperone Hsp70 (Liberek *et al.*, 1991; Szabo *et al.*, 1994; Harrison *et al.*, 1997), podendo agir também na dissociação dos substratos pela interação com o DLS das Hsp70 (Harrison *et al.*, 1997; Mally & Witt, 2001; Chesnokova *et al.*, 2003) ou prevenindo a ligação dos substratos (Brehmer *et al.*, 2004). Além disso, elas podem exercer uma função de termosensor para o sistema Hsp70 (Grimshaw *et al.*, 2001; Groemping & Reinstein, 2001; Gelinas *et al.*, 2002; Grimshaw *et al.*, 2003; Gelinas *et al.*, 2003). Os organismos procarióticos, como *E. coli* e *T. thermophilus*, possuem apenas uma isoforma das GrpE que são bem caracterizadas estrutural e funcionalmente (Harrison *et al.*, 2003). Entretanto, existem pelo menos duas isoformas de GrpE em organismos eucarióticos superiores, sendo que as GrpEs oriundas de mitocôndria de mamíferos apresentam funções e distribuição similares (Naylor *et al.*, 1998). No genoma humano foram identificados dois genes (genes GRPEL1 e GRPEL2) que codificam seqüências de aminoácidos similares às proteínas GrpE e também alguns pseudogenes, provavelmente originados por fenômenos de recombinação, duplicação e inversão gênica (Choglay *et al.*, 2001). Esses dados sugerem que os genes das GrpEs humanas são evolutivamente antigos no genoma humano. Porém, qual é a necessidade de duas proteínas homólogas com funções similares no interior de um mesmo compartimento intracelular? Com o objetivo de compreender e tentar responder esta pergunta, foram realizados experimentos para a caracterização da estrutura e da estabilidade das duas GrpEs encontradas na mitocôndria de humanos. De maneira geral, as GrpEs apresentam estruturas quaternárias similares, porém foram observadas algumas diferenças tanto na forma quanto na estabilidade dessas proteínas. Os dados sugerem que, apesar da similaridade na seqüência primária e quaternária, as GrpEs humanas podem apresentar diferentes funções e/ou atividades na troca de nucleotídeos adenosina das Hsp70 mitocondriais.

5.4.1 Clonagem do cDNA, expressão e purificação das proteínas GrpEs humanas

Os cDNAs completos das GrpEs humanas foram obtidos e clonados em vetores de expressão, porém, as primeiras tentativas de purificação da proteína GrpE#1 não foram bem sucedidas devido à sua agregação (dados não mostrados). A análise da seqüência de aminoácidos das GrpEs humanas mostrou a presença de peptídeos sinais responsáveis pelo

direcionamento das proteínas para a mitocôndria – fato condizente com o observado para as GrpEs de outros mamíferos (Naylor *et al.*, 1998). Após novo procedimento de clonagem do cDNA da GrpE#1, no qual a seqüência de cDNA correspondente ao peptídeo sinal foi excluída da seqüência de nucleotídeos da GrpE#1, a proteína foi expressa e obtida na forma solúvel e com estrutura secundária compatível à esperada para proteínas ricas em estrutura secundária em α -hélices, como a EcGrpE (Harrison *et al.*, 1997). A GrpE#1 foi expressa e purificada na forma solúvel, mesmo na ausência de sal, e a discussão dos resultados obtidos para a GrpE#1 estão descritos no artigo publicado e apresentado no anexo IV (Borges *et al.*, 2003).

De forma similar à GrpE#1, o cDNA da GrpE#2 foi também clonado sem a seqüência de nucleotídeos responsável pela codificação do peptídeo sinal de exportação para a mitocôndria. No entanto, diferente da GrpE#1, a GrpE#2 apresentou baixa solubilidade e estabilidade em soluções com baixa força iônica, fatos que provavelmente inviabilizaram os procedimentos de purificação da proteína por troca iônica. Assim, o cDNA da GrpE#2 foi clonado em fusão à cauda de poli-histidina, expressa e as células lisadas em alta força iônica, resultando na proteína solúvel e estável. Assim, o processo de purificação pôde ser realizado por cromatografia de afinidade seguida de uma CEM preparativa, ambas em alta força iônica, resultando na proteína HisGrpE#2 na forma pura e estável. A cauda de poli-histidina da HisGrpE#2 foi clivada através da proteólise limitada com a enzima trombina e foi obtida a GrpE#2 com aproximadamente 22 kDa, analisada por SDS-PAGE. A GrpE#2 apresentou uma *MM* menor em aproximadamente 2 kDa do que a *MM* apresentada pela HisGrpE#2, enfatizando a eficiência do processo de proteólise limitada. Como discutido acima, a GrpE#2 apresentou estabilidade apenas em condições de alta força iônica (500 mmol/L de NaCl), enquanto a GrpE#1 foi solúvel mesmo na ausência de sal (Anexo IV – Borges *et al.*, 2003). Portanto, foi necessário caracterizar a estabilidade da GrpE#1 nas mesmas condições experimentais que a GrpE#2, com o intuito de estudos comparativos entre essas duas proteínas.

5.4.2 As proteínas GrpEs humanas possuem diferentes quantidades de estrutura secundária

Os ensaios de CD das GrpEs humanas sugerem que essas proteínas possuem alto conteúdo de estrutura secundária do tipo α -hélices. Porém, a GrpE#2 apresentou menor

quantidade de sinal de CD a 208 e 222 nm e, portanto menor quantidade de estrutura secundária quando comparada à GrpE#1. Comparando o perfil do espectro de CD das GrpEs humanas com a EcGrpE, é possível concluir que a GrpE#1 possui uma maior similaridade com o espectro apresentado pela EcGrpE (Grimshaw *et al.*, 2001; Gelinas *et al.*, 2002) do que a GrpE#2. Esses dados sugerem que a GrpE#2 pode ser uma proteína com menor quantidade de estrutura secundária, podendo apresentar seguimentos mais flexíveis, do que a GrpE#1.

5.4.3 As proteínas GrpEs humanas possuem diferentes estabilidades

Os experimentos de desenovelamento térmico das GrpEs humanas sugerem que essas proteínas sofreram uma tênue perda de estrutura secundária entre as temperaturas de 20 a 40 °C, sendo que essa perda foi mais acentuada para a GrpE#1 do que para a GrpE#2. A EcGrpE também apresenta um perfil de desenovelamento térmico similar entre 20 a 40 °C, com tênue perda de sinal de CD, que está relacionada à perda da estrutura das duas longas α -hélices e, assim, com a função de termosensor da EcGrpE (Grimshaw *et al.*, 2001; Grimshaw *et al.*, 2003; Gelinas *et al.*, 2003). Em temperaturas superiores a 40 °C, a perda de estrutura secundária das GrpEs humanas foi mais acentuada e resultou no desenovelamento parcial das proteínas com a detecção de Tms bem definidas. A primeira Tm foi maior para a GrpE#1 em comparação à GrpE#2 (53 °C *versus* 50 °C, respectivamente) nas condições testadas. Os dados da literatura também mostram que a EcGrpE possui uma Tm em 50 °C (Grimshaw *et al.*, 2001; Gelinas *et al.*, 2002). Esses dados são interessantes, pois na região das longas α -hélices a similaridade entre a EcGrpE e a GrpE#2 é maior do que entre a EcGrpE e a GrpE#1.

O perfil de estabilidade similar entre a GrpE#2 e a EcGrpE também foi notado pela análise da presença de uma segunda Tm em aproximadamente 76 a 78 °C. Entretanto, a GrpE#2 sofreu agregação seguida de precipitação, enquanto a EcGrpE é parcialmente reversível nas condições testadas (Grimshaw *et al.*, 2001; Gelinas *et al.*, 2002; Grimshaw *et al.*, 2003; Gelinas *et al.*, 2003; Brehmer *et al.*, 2004). A GrpE#1 não perdeu totalmente o sinal de CD em λ de 222 nm, mesmo quando aquecida a 90 °C, mas sofreu agregação em temperaturas superiores ao primeiro Tm, como avaliado por DSC (Anexo IV – Borges *et al.*, 2003). A EcGrpE também apresenta sinal da presença de estrutura secundária mesmo quando aquecida a 90 °C (Grimshaw *et al.*, 2001; Gelinas *et al.*, 2002; Grimshaw *et al.*, 2003;

Brehmer *et al.*, 2004). Estas informações sugerem que a GrpE#1 e a EcGrpE também compartilham estabilidades térmicas similares entre alguns de seus domínios. A GrpE#1 apresenta maior similaridade – do que a GrpE#2 – com a EcGrpE na região correspondente ao feixe de 4 hélices, região que é responsável pela estabilidade do dímero da EcGrpE (Gelinas *et al.*, 2002). Um exemplo da importância dos domínios e das regiões conservadas na estabilidade térmica e química da EcGrpE é apresentado no trabalho de Brehmer *et al.* (2004). A deleção dos primeiros 36 resíduos de aminoácidos da EcGrpE ocasiona a irreversibilidade do reenovelamento térmico e químico da proteína quando monitorado por CD.

Os ensaios de desnaturação induzida por uréia também mostraram que a GrpE#1 e a GrpE#2 possuem diferentes estabilidades. A GrpE#1 apresentou um perfil de desenovelamento cooperativo sugerindo que seus domínios possuem estabilidades similares a agentes químicos. A GrpE#2 foi desenovelada com pelo menos duas transições em concentrações de uréia superiores do que a GrpE#1, indicando que seus domínios possuem diferentes estabilidades. Considerando as características de estabilidade frente a desnaturantes químicos, a GrpE#2 apresentou resultados mais similares aos da EcGrpE, pois esta proteína também foi desenovelada em múltiplas etapas (Gelinas *et al.*, 2002). Entretanto, a EcGrpE suporta condições de baixa força iônica (Grimshaw *et al.*, 2001; Gelinas *et al.*, 2002), condição em que a GrpE#2 foi instável, mas que a GrpE#1 apresentou estabilidade (Anexo IV – Borges *et al.*, 2003). Estas informações sugerem que a EcGrpE e a GrpE#1 também possuem similaridades em relação à estabilidade química.

Os estudos de estabilidade térmica mostraram que a GrpE#1 possui maior resistência à desnaturação térmica do que a GrpE#2 e a EcGrpE. A menor estabilidade térmica da GrpE#2, em comparação a GrpE#1, também pode ser avaliada pelo fato de que ela foi obtida, em quantidades adequadas no sobrenadante do lisado celular, quando as células foram induzidas à expressão em baixas temperaturas. A GrpE#1 foi obtida em condições de indução em temperaturas superiores. Contudo, a GrpE#1 apresentou menor estabilidade nos ensaios de desnaturação induzida por uréia do que a GrpE#2 e a EcGrpE. Os dados de DSC, obtidos para a GrpE#1 e apresentados no Anexo IV (Borges *et al.*, 2003), mostraram que essa proteína não apresentou dependência da força iônica do meio no seu desenovelamento térmico. Entretanto, na avaliação da desnaturação térmica monitorada por CD, a alta força iônica reduziu em 4 °C a primeira T_m da GrpE#1 quando comparado aos dados obtidos em condições de baixa força iônica (Anexo IV – Borges *et al.*, 2003). Os dados

de desenovelamento térmico e químico da GrpE#1 sugerem a influência da força iônica na estabilidade dessa proteína. No entanto, a inviabilidade experimental de realizar experimentos com a GrpE#2 em condições de baixa força iônica e a ausência de informações disponíveis sobre a influência da força iônica na estabilidade térmica da EcGrpE, inviabilizaram as análises comparativas e restringiram as conclusões. Devido à irreversibilidade dos processos de desenovelamento térmico e químico das GrpEs humanas, as análises mais apuradas foram prejudicadas, o que dificultou a impetração mais adequada em relação à EcGrpE. Talvez, tentativas de experimentos de desenovelamento térmico em outras condições de força iônica, acidez e agentes de controle de pH possam satisfazer a condição de reversibilidade e auxiliar na interpretação dos dados.

5.4.4 As proteínas GrpEs humanas possuem estruturas quaternárias alongadas

Os ensaios de SAXS e os modelos ab initio para as proteínas GrpEs humanas

Foram realizados experimentos de SAXS para as GrpEs humanas para avaliar a estrutura quaternária e o formato destas proteínas em solução. Foram utilizadas duas concentrações da GrpE#1 (5,0 e 19,4 mg/mL) para coletar dados de SAXS, em baixa concentração de sal e não foram observados indícios de agregação (Anexo IV – Borges *et al.*, 2003). Os dados da GrpE#2 foram obtidos em duas concentrações (3,5 e 9,3 mg/mL) e em alta força iônica, sendo que foram observados indícios de agregação nos dados da GrpE#2 na amostra em alta concentração (9,3 mg/mL). Estas diferenças de comportamento das GrpEs humanas em alta concentração também evidenciam as dissimilaridades na estabilidade dessas proteínas. As análises das curvas de espalhamento e das curvas das $p(r)$ também evidenciaram a melhor qualidade estatística dos dados de SAXS obtidos para a GrpE#1 em comparação à GrpE#2. Esta diferença na qualidade dos dados foram provavelmente às limitações impostas pela agregação da amostra da GrpE#2 em alta concentração. Contudo, os dados de SAXS mostraram que as GrpEs humanas possuem *MM* compatíveis à estrutura quaternária dessas proteínas na forma de dímeros (aproximadamente 42-44 kDa). Estes dados sugerem que mesmo com os problemas de agregação, os dados de SAXS da GrpE#2 apresentaram qualidade adequada para as demais análises.

Os experimentos de SAXS das GrpEs humanas mostraram que estas proteínas são anisotrópicas e possuem uma forma geral prolata alongada, como analisado pelo perfil das

curvas das $p(r)$ (Feigin & Svergun, 1987). Os dados de SAXS foram utilizados para construir modelos *ab initio* e as análises destes mostraram que a GrpE#1 e a GrpE#2 apresentam estrutura quaternária parecidas. Entretanto, possuem importantes diferenças no posicionamento do domínio carboxi-terminal estruturado em folha β -pregueada (Figura 2 – representado por EcGrpE₁₃₈₋₁₉₇). A GrpE#1 apresentou esse domínio na conformação correspondente ao do monômero distal da EcGrpE (Anexo IV – Borges *et al.*, 2003); enquanto a GrpE#2 apresentou tal domínio mais parecido à conformação correspondente ao do monômero proximal, porém, ainda, de maneira insatisfatória. Através da modelagem do domínio carboxi-terminal da GrpE#1 dentro do envelope do modelo *ab initio* da GrpE#2, ele pode ser ajustado satisfatoriamente dentro do envelope do modelo quando o domínio carboxi-terminal foi posicionado em paralelo às duas longas α -hélices (dados não mostrados). Os dados sugerem que a GrpE#2 não se apresentou, nas condições estudadas, em nenhuma das conformações dos monômeros observadas na estrutura em alta resolução determinado para a EcGrpE.

Os dados de SAXS obtidos para as GrpEs humanas mostram que as regiões amino-terminal apresentaram alguma estruturação e também que tais proteínas possuem D_{max} de aproximadamente 145-155 Å. Com este tamanho, as GrpEs humanas são grandes o suficiente para interagirem com ambos os domínios das Hsp70 em um mesmo ciclo de interação, como proposto por Harrison *et al.* (1997), Mehl *et al.* (2001), Chesnokova *et al.* (2003) e Brehmer *et al.* (2004). Os dados para as Hsp70 sugerem que elas possuem diâmetro máximo de aproximadamente 110 a 135 Å nas diferentes condições experimentais analisadas, como mostrado para a Hsp70-1A (item 4.1.1 de Resultados e 5.1.7 de Discussão), DnaK (Shi *et al.*, 1996) e Hsc70 bovina (Wilbanks *et al.*, 1995).

As análises hidrodinâmicas das proteínas GrpEs humanas

As análises hidrodinâmicas por UAL também corroboraram a hipótese de que as GrpEs humanas apresentam diferentes propriedades biofísicas em solução e possuem assimetria conformacional. A GrpE#1 apresentou um $s_{20,w}^0$ de 2,34 S que foi concordante com os dados estimados para o modelo *ab initio* gerado a partir dos dados de SAXS e sem problemas de agregação (Anexo IV – Borges *et al.*, 2003). No entanto, a GrpE#2 apresentou um $s_{20,w}^0$ de 2,80 S que foi dissimilar com os dados estimados pelo programa HydroPro, cujo valor de $s_{20,w}^0$ estimado para o modelo *ab initio* da GrpE#2 foi de 2,30 S.

O cálculo da *MM* da GrpE#1 a partir dos dados hidrodinâmicos foi concordante com os dados esperados para a forma dimérica da proteína (Anexo IV – Borges *et al.*, 2003). Mas, a *MM* estimada para a GrpE#2 a partir da relação *s/D* foi maior ao valor calculado a partir da sequência de aminoácidos da proteína (50 kDa *versus* 44 kDa, respectivamente) e também ao valor que foi calculado a partir dos dados de SAXS (aproximadamente 42 kDa). Os dados hidrodinâmicos sugerem que a GrpE#2 pode não ser tão assimétrica quanto a GrpE#1 ou que a GrpE#2 pode ser mais flexível que a GrpE#1. Alguns autores discutem que partículas assimétricas não rígidas ou volumosas sofrem uma distorção no cálculo do valor de *s* devido à flexibilidade de segmentos da partícula em solução que reduz a razão friccional quando comparado a corpos rígidos (Laue *et al.*, 1992). Na verdade, os experimentos de velocidade de sedimentação medem a média temporal das partículas na solução, sendo que a razão friccional está relacionada à média temporal do formato da molécula. Assim, não é possível distinguir, utilizando UAL somente, partículas com alta assimetria e flexíveis (ou volumosas) de moléculas rígidas com baixa assimetria (Waxman *et al.*, 1994). Os modelos *ab initio* gerados a partir dos dados de SAXS podem ser considerados como sendo corpos rígidos, porém, foram obtidos por uma técnica que também coleta informações correspondentes à média temporal da densidade eletrônica da partícula espalhadora. No entanto, se forem analisados conjuntamente apenas os dados hidrodinâmicos determinados para as GrpEs humanas e EcGrpE, cujo valor de $s_{20,w}$ determinado foi de 2,7 S (Schönfeld *et al.*, 1995; Schönfeld & Behlke, 1998), a GrpE#2 foi mais similar à EcGrpE do que a GrpE#1.

As proteínas GrpEs humanas são proteínas flexíveis

Os dados de SAXS obtidos para as GrpEs humanas, analisados pelo perfil do gráfico de Kratky, sugerem que tais proteínas apresentam porções flexíveis e que a GrpE#2 possui regiões mais flexíveis do que a GrpE#1. Proteínas compactas apresentam a curva do gráfico de Kratky na forma de sino e proteínas com segmentos flexíveis diferem deste perfil (Doniach *et al.*, 1995). A técnica de SAXS também é baseada em uma média temporal da densidade eletrônica da partícula espalhadora. Assim, os dados de SAXS e de UAL mostraram que a GrpE#2 possui em sua estrutura regiões com flexibilidade e que o corpo rígido representado pelo modelo *ab initio* pode representar uma das principais conformações da partícula.

5.4.5 A similaridade entre as proteínas GrpEs humanas e a proteína EcGrpE

A análise da similaridade na sequência aminoácidos entre as GrpEs humanas e a EcGrpE sugere que em alguns domínios a proteína GrpE#1 é mais similar à proteína EcGrpE, enquanto que, em outro domínio, a proteína GrpE#2 é mais similar. Os dados de estrutura secundária, desenovelamento térmico e comportamento em soluções contendo baixa força iônica, sugerem que a GrpE#1 compartilha mais características similares com a EcGrpE do que a GrpE#2. Entretanto, em ensaios de desenovelamento térmico e químico e nas propriedades hidrodinâmicas, a GrpE#2 apresentou mais características similares com a EcGrpE do que a GrpE#1. Os experimentos de SAXS sugerem que ambas as GrpEs humanas possuem um formato alongado, porém apresentam possíveis diferenças na conformação do domínio de interação com as Hsp70.

5.4.6 A relação entre a forma e a termodinâmica das proteínas GrpEs humanas

Dados termodinâmicos de mutantes da EcGrpE, avaliados por DSC, mostram a existência de 3 temperaturas de transição e que o sinal do desenovelamento térmico das regiões correspondentes ao feixe de 4 hélices e rico em folha β -pregueada estão superpostos na proteína nativa (Gelinas *et al.*, 2003). O desenovelamento do domínio carboxi-terminal rico em estrutura em folha β -pregueada da EcGrpE possui uma T_m de aproximadamente 77 °C e o feixe de 4 hélices da EcGrpE sofre o desenovelamento em temperaturas superiores a 90 °C (Gelinas *et al.*, 2003). Os dados apresentados por Gelinas *et al.* (2003) também sugerem que essas duas regiões da EcGrpE estão interligadas termodinamicamente, pois as mutações realizadas em determinadas regiões estabilizam um dos domínios e/ou desestabilizam o outro. Considerando as informações obtidas para as GrpEs humanas, é possível concluir que estas proteínas apresentam diferentes estabilidades, principalmente nessas duas regiões, como discutido abaixo.

A GrpE#2, que possui menor similaridade na região do feixe de 4 hélices com a EcGrpE, pode ter sofrido a desnaturação térmica do domínio rico em folha β -pregueada em 76 °C. Essa desnaturação levou à agregação da GrpE#2, possivelmente porque contatos entre a região rica em estrutura β -pregueada, que poderiam estabilizar o feixe de 4 hélices, não existem mais. Entretanto, os modelos *ab initio* construídos para a GrpE#2 sugerem que os domínios da região carboxi-terminal da GrpE#2 podem estar em uma conformação diferente das observadas para cada monômero da estrutura em alta resolução da EcGrpE

(Harrison *et al.*, 1997). Uma possibilidade é que os domínios carboxi-terminal do dímero da GrpE#2 assumissem uma conformação paralela às duas longas α -hélices. Nesta conformação eles seriam desenovelados em temperaturas próximas a T_m relacionada às duas longas α -hélices.

A GrpE#1 ainda apresentou sinais da presença de estrutura secundária em 90 °C, sem apresentar uma segunda T_m característica, possivelmente devido a maior estabilidade do feixe de 4 hélices. Essa região apresentou maior similaridade com a EcGrpE do que a GrpE#2. A EcGrpE apresenta uma pequena perda de sinal de CD que foi associada ao desenovelamento da região rica em estrutura β -pregueada (Gelinas *et al.*, 2002) e esse sinal não foi identificado para a GrpE#1. Esse dado pode ser um indicativo de que os domínios carboxi-terminal da GrpE#1 foram desenovelados, devido aos respectivos contatos com o feixe de 4 hélices, em temperaturas superiores a 90 °C. A outra possibilidade é a de que a região rica em estrutura β -pregueada da GrpE#1 sofreu o desenovelamento juntamente com a região das duas longas α -hélices, como previamente discutido no Anexo IV (Borges *et al.*; 2003). Entretanto, os dados termodinâmicos obtidos para as GrpEs humanas não apresentam resolução suficiente para uma melhor intercorrelação das T_m s observadas e as regiões das proteínas que foram desenoveladas. Mas, direcionam as possibilidades de novos experimentos a serem realizados, como por exemplo, mutantes de deleção das regiões conservadas das GrpEs humanas.

5.4.7 As proteínas GrpEs humanas compartilham similaridades e diferenças com a EcGrpE

De maneira geral, os resultados obtidos sugerem que as GrpEs humanas compartilham diferenças e similaridades com a EcGrpE, tanto na sequência primária, conformação, estabilidade térmica e química. Essas diferentes características podem indicar que essas proteínas humanas possuem diferentes funções na resposta mitocondrial ao estresse térmico e/ou químico. O primeiro estudo comparativo entre as GrpE#1 e GrpE#2 de mamíferos não mostram consideráveis diferenças entre as GrpEs homólogas de rato e camundongo (Naylor *et al.*, 1998), pois esse estudo não avalia a estabilidade e estrutura daquelas proteínas. Naylor *et al.* (1998) mostra que as proteínas GrpE#1 e GrpE#2 de rato e camundongo, respectivamente, possuem a capacidade de interagir com as Hsp70 mitocondriais e que as funções dessas proteínas podem estar relacionada à regulação pós-

transcricional e/ou pós-traducional. A baixa similaridade entre as GrpEs humanas na região amino-terminal pode indicar que elas apresentam diferentes funções. Os primeiros 34 aminoácidos da EcGrpE, que estão ausentes na estrutura em alta resolução (Figura 2 – Harrison *et al.*, 1997), são importantes para a interação com o DLS da DnaK e para aumentar a dissociação de substratos com aquele domínio (Harrison *et al.*, 1997; Mally & Witt, 2001; Chesnokova *et al.*, 2003), ou, por um mecanismo diferente, prevenindo a interação de novos substratos com o DLS das Hsp70 (Brehmer *et al.*, 2004).

Uma forma de compreender como essas diferentes propriedades biofísicas estão ligadas com a função das GrpEs, é a caracterização funcional dessas proteínas em relação à interação com as Hsp70 mitocondriais. Entretanto, a ausência de uma representante das Hsp70 de origem mitocondrial impossibilitou a realização, neste trabalho, de um teste funcional das GrpEs humanas. Testes que caracterizem a interação das GrpEs humanas com as Hsp70 mitocondriais poderão resultar em conclusões mais adequadas sobre a função das GrpEs humanas na mitocôndria. E assim, melhorar o entendimento do funcionamento do sistema chaperone Hsp70 e entender o porquê da existência de duas proteínas homólogas às GrpEs localizadas em um mesmo compartimento intracelular.

5.5 Discussão geral

Este trabalho de doutoramento descreve estudos de caracterização parcial da relação estrutura-função, por várias técnicas biofísicas, de uma das maquinarias mais importantes para o enovelamento e metabolismo protéico: as chaperones moleculares do sistema Hsp70. Um total de 7 cDNAs de proteínas humanas foram clonados em vetores de expressão e expressas com sucesso. Os procedimentos de purificação resultaram na obtenção de 6 proteínas em suas formas solúveis e com *MM* compatíveis com as estruturas oligoméricas correspondentes já descritas na literatura. Para a caracterização das proteínas em solução, foram aplicadas técnicas de biofísica molecular de baixa resolução como: dicroísmo circular, fluorescência do triptofano, calorimetria de titulação isotérmica, calorimetria diferencial de varredura, espalhamento dinâmico de luz, ultracentrifugação analítica e espalhamento de raios X a baixos ângulos. Essas técnicas possibilitaram a caracterização da estrutura secundária, estrutura terciária e quaternária das proteínas; também possibilitaram a caracterização da estabilidade frente à desnaturação térmica e química de algumas das proteínas e também características em relação à interação com ligantes. Todas essas informações foram úteis para a interpretação da relação estrutura-função das proteínas estudadas.

Algumas atividades de caracterização necessárias para a interpretação dos resultados, não foram determinadas por motivos técnicos, como, por exemplo, a caracterização da atividade chaperone da Hsp70-1A. A necessidade de uma padronização das técnicas, para a determinação dessas atividades, aliadas ao pouco tempo, dificultou esses experimentos, de modo que essas características deverão ser realizadas em outra oportunidade. A modelagem molecular por homologia da Hsp70-1A e GrpE#2 não foram realizadas pelo pouco tempo disponível. A modelagem das proteínas Hsp40 por homologia também não foi realizada devido ao mesmo motivo e também pela presença, na sequência dessas proteínas, de uma região com baixa predição de estrutura secundária (a região rica em glicinas e fenilalaninas).

O sistema chaperone molecular Hsp70 e as suas co-chaperones Hsp40 e GrpE agem em conjunto através de interações intermoleculares que podem ser: transiente estável ou instável, sendo ambos dependentes de aditivos (revisado por Mayer *et al.*, 2001). Essas informações mostram a necessidade de estudos de interação entre tais proteínas para confirmar as interações e também entender como elas acontecem. Porém, devido ao estágio

de implementação da linha de pesquisa no grupo, foi necessário compreender melhor qual o comportamento de cada uma das proteínas em solução, qual a estrutura apresentada por elas e a potencialidade das técnicas implementadas para a caracterização biofísica. Assim, foi empregada a filosofia de “dividir para conquistar” a fim de obter informações estruturais de cada uma das proteínas separadas em solução. Compreendendo melhor a relação estrutura-função das proteínas sozinhas, poder-se-ia estudá-las melhor em conjunto.

Experimentos para estudar a interação das GrpEs humanas com uma Hsp70 de origem mitocondrial seriam avaliados, porém, essa última (mortalina) não foi obtida em sua forma nativa. Porém, estes estudos se fazem necessários e serão realizadas novas tentativas em uma outra oportunidade. Para avaliar a interação entre as Hsp70 e Hsp40, que é transiente instável (Pierpaoli *et al.*, 1998; Laufen *et al.*, 1999) e que dificulta a obtenção de complexos estáveis, estão sendo construídos mutantes sítio-específico, nos quais os aminoácidos responsáveis pela interação entre as Hsp40 e Hsp70 (Greene *et al.*, 1998; Suh *et al.*, 1998) foram modificados por cisteínas. A intenção é obter uma forma oligomérica estável entre a Hsp70 e Hsp40, passível de ser estudada por técnicas como SAXS, AUC e também, talvez por difração de raios-X através da obtenção de cristais das proteínas formando complexos com os ligantes. Desse modo, poderão ser mostrados dados estruturais dessas proteínas formando complexos protéicos. E assim, levar a comunidade científica a compreender melhor como é o mecanismo de interação entre as Hsp70 e suas co-chaperones da família das Hsp40 e, principalmente, compreender de que forma estrutural elas interagem. No entanto, o estudo individual de cada proteína foi bem sucedido e como “o tempo todo o tempo passa”, pouco tempo restou para os estudos de interação proteína-proteína.

6 CONCLUSÕES

6.1 Conclusões gerais

As mudanças conformacionais da Hsp70-1A induzidas pelos nucleotídeos adenosina foram pequenas. Entretanto, elas podem ser o resultado de uma reorganização estrutural dos subdomínios e domínios da proteína.

As análises estruturais das Hsp40 humanas mostraram que as proteínas das subfamílias A e B, DjA1 e DjB4, respectivamente, possuem diferentes formas em solução e possuem diferente atividade chaperone.

As GrpEs humanas possuem estrutura quaternária similar, porém diferem no posicionamento relativo de alguns domínios e na estabilidade. Esses dados sugerem que essas proteínas possuem diferenças em suas funções e/ou atividades.

6.2 Conclusões específicas sobre a proteína Hsp70-1A

O cDNA da Hsp70-1A foi clonado em vetor de expressão, a proteína foi expressa e purificada em alta quantidade, apresentando-se com mais de 95% de pureza.

Foi preparado um tampão para os ensaios biofísicos da Hsp70-1A a fim de controlar os efeitos observados para a Hsp70-1A em relação à oligomerização, com perfeita ligação dos nucleotídeos, porém com baixa taxa de hidrólise de ATP.

Os experimentos de CD da Hsp70-1A mostraram que essa proteína possui predominância de estrutura secundária do tipo α -hélice e que, a adição de MgADP ou MgATP provocou pequenas mudanças no espectro de CD.

O desenovelamento térmico da Hsp70-1A mostrou que os ligantes MgATP ou MgADP provocaram uma compactação da proteína que aumentou a resistência da proteína à desnaturação térmica. Indiretamente, esses dados mostram que a Hsp70-1A foi purificada sem a presença de nucleotídeo adenosina ligado à mesma.

Experimentos de fluorescência mostraram que MgATP provocou pequenas mudanças estruturais na proteína na região do triptofano que resultou em uma menor intensidade de fluorescência.

A Hsp70-1A apresentou maior afinidade por MgADP do que por MgATP. Foi observada uma maior ΔH para a interação da Hsp70-1A com MgADP do que por MgATP que, indiretamente, mostrou a ausência de hidrólise de ATP.

Os dados de afinidade da Hsp70-1A, determinados para seus ligantes MgATP e MgADP por calorimetria de titulação isotérmica, nortearam os experimentos de velocidade de sedimentação na presença desses aditivos em condições de saturação.

Os dados de velocidade de sedimentação mostraram que a Hsp70-1A possui *MM* de aproximadamente 70 kDa, mesmo na presença dos nucleotídeos adenosina. Os dados de sedimentação em equilíbrio estão concordantes com o modelo da Hsp70-1A como um monômero com *MM* de 70 ± 2 kDa.

Os dados de velocidade de sedimentação mostraram que os ligantes MgATP ou MgADP provocaram pequenas mudanças conformacionais na Hsp70-1A.

Os dados de SAXS da Hsp70-1A mostraram que os nucleotídeos adenosina provocam pequenas mudanças no perfil da curva da função de intensidade de espalhamento e na curva da função de distribuição de distâncias.

As amostras da proteína Hsp70-1A na presença e na ausência dos nucleotídeos adenosina apresentaram D_{max} e modelos *ab initio* similares.

Os modelos *ab initio* construídos a partir dos dados de SAXS, de forma geral, acomodaram bem as estruturas em alta resolução dos domínios das Hsp70 disponíveis no PDB.

Os dados experimentais sugerem que pequenas mudanças conformacionais são ocasionadas pela ligação do nucleotídeo e podem resultar da reorganização dos subdomínios e domínios da Hsp70-1A com a possível exposição do sítio de ligação dos substratos.

6.3 Conclusões específicas sobre a proteína mortalina

O cDNA da proteína mortalina, sem o peptídeo sinal para exportação da proteína para mitocôndria, foi clonado em fase de leitura com a cauda de poli-histidina do DNA plasmidial pET28a.

A mortalina foi expressa, porém foi purificada com uma partícula que eluiu com uma alta *MM* aparente nas cromatografias de exclusão molecular.

Os dados de CD mostraram que a mortalina não possui a quantidade de estrutura secundária esperada, o que inviabilizou a utilização desta proteína para os ensaios comparativos com a Hsp70-1A e os estudos de interação proteína-proteína com as GrpEs.

6.4 Conclusões específicas sobre as proteínas Hsp40

Os cDNAs das Hsp40 humanas, DjA1, DjB4 e um mutante de deleção da região carboxi-terminal da DjA1, foram clonados em vetor de expressão; as proteínas foram bem expressas e purificadas na forma estável para os ensaios biofísicos e apresentaram estrutura secundária compatível com o esperado.

As Hsp40 apresentaram atividade chaperone na prevenção da agregação da luciferase, porém, com diferentes atividades para a DjA1 e DjB4. A DjA1₁₋₃₃₂ foi menos ativa em prevenir a agregação da luciferase em relação à DjA1, sugerindo uma dependência da estrutura quaternária.

Os dados de UAL e SAXS mostraram que as Hsp40, DjA1 e DjB4 são dímeros, e a DjA1₁₋₃₃₂ é um monômero.

Os dados de SAXS mostraram que a DjA1 e a DjB4 possuem diferentes formas. Modelos *ab initio* foram gerados para as Hsp40 e para a deleção DjA1₁₋₃₃₂.

Dados de velocidade de sedimentação indicam que a DjA1₁₋₃₃₂ pode ser mais globular do que o esperado a partir do modelo *ab initio*, sugerindo que esta proteína possui flexibilidade em sua estrutura.

O modelo *ab initio* da DjA1₁₋₃₃₂ superpôs adequadamente à metade do modelo *ab initio* do dímero da DjA1.

Os modelos *ab initio* das DjA1, DjA1₁₋₃₃₂ e DjB4 acomodaram bem os modelos em alta resolução de domínios das Hsp40 e sugeriram como as Hsp70 podem interagir com as Hsp70.

6.5 Conclusões específicas sobre as proteínas GrpEs humanas

Os cDNAs das GrpEs humanas (GrpE#1 e GrpE#2) foram clonados em vetores de expressão; as proteínas foram expressas e purificadas na forma estável para os ensaios biofísicos.

As GrpE#1 e GrpE#2 humanas apresentaram estrutura secundária similar à EcGrpE, entretanto, sendo a primeira mais similar.

As GrpEs humanas apresentam diferentes perfis de desenovelamento térmico e químico, sendo que a GrpE#2 é mais similar a EcGrpE do que a GrpE#1. No entanto, a GrpE#1 também apresentou similaridades com a EcGrpE, pois mesmo a 90 °C ambas apresentaram sinal estrutura secundária, enquanto a GrpE#2 precipitou.

As análises hidrodinâmicas mostraram a GrpE#2 é mais similar à EcGrpE do que a GrpE#1.

Os dados de SAXS mostraram que as GrpEs humanas são proteínas alongadas, com regiões flexíveis, possuem simetria axial e que a região amino-terminal possui estrutura.

Os dados de SAXS e UAL mostraram que a GrpE#2 possui regiões mais flexíveis do que a GrpE#1 e que o corpo rígido representado pelo modelo *ab initio* pode representar uma das principais conformações da partícula.

Os dados de D_{max} obtidos a partir dos dados de SAXS mostraram que a GrpE#1 e GrpE#2 são grandes o suficiente para interagir com ambos domínios das Hsp70.

Os modelos *ab initio* da GrpE#1 e GrpE#2 apresentaram diferenças entre si na forma geral da partícula, principalmente na região do domínio carboxi-terminal.

Dados de SAXS e UAL possibilitaram a modelagem da GrpE#1 a partir da estrutura em alta resolução parcial da EcGrpE.

A modelagem da GrpE#1 humana a partir do monômero distal apresentou uma melhor superposição com o modelo *ab initio*, sugerindo sua possível conformação em solução.

A modelagem da GrpE#1 humana a partir do monômero proximal apresentou uma melhor superposição com o modelo *ab initio* da GrpE#2, porém, ainda insatisfatória. Esses dados sugerem que a GrpE#2 em solução possui uma outra conformação que aquelas mostradas para cada monômero na estrutura em alta resolução da EcGrpE.

7 REFERÊNCIAS BIBLIOGRÁFICAS

- Anfisen, C.B. (1973) Principles that govern the folding of protein chains. *Science*, 181, 223-230.
- Banecki, B., Zylicz, M., Bertoli, E. & Tanfani, F. (1992) Structural and functional relationships in DnaK and DnaK756 heat-shock proteins from *Escherichia coli*. *J. Biol. Chem.*, 267, 25051-25058.
- Banecki, B., Liberek, K., Wall, D., Wawrzynow, A., Georgopoulos, C., Bertoli, E., Tanfani, F. & Zylicz, M. (1996) Structure-function analysis of the zinc finger region of the DnaJ molecular chaperone. *J. Biol. Chem.*, 271, 14840-14848.
- Beissinger, M. & Buchner, J. (1998) How chaperones fold proteins. *Biol. Chem.*, 379, 245-259.
- Benaroudj, N., Batelier, G., Triniolles, F. & Ladjimi, M.M. (1995) Self-association of the molecular chaperone HSC70. *Biochemistry*, 34, 15282-15290.
- Benaroudj, N., Triniolles, F. & Ladjimi, M.M. (1996) Effect of nucleotides, peptides, and unfolded proteins on the self-association of the molecular chaperone HSC70. *J. Biol. Chem.*, 271, 18471-18476.
- Benaroudj, N., Triniolles, F. & Ladjimi, M.M. (1996) The COOH-terminal peptide binding domain is essential for self-association of the molecular chaperone HSC70. *J. Biol. Chem.*, 272, 8744-8751.
- Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. & Bourne, P.E. (2000) The Protein Data Bank. *Nucleic Acids Research*, 28, 235-242.
- Böhm, G., Muhr, R. & Jaenicke, R. (1992) Quantitative analysis of protein far UV circular dichroism spectra by neural networks. *Protein Eng.*, 5, 191-195.
- Borges, J.C., Fischer, H., Craievich, A.F., Hansen, L.D., & Ramos, C.H.I. (2003) Free human mitochondrial GrpE is a symmetric dimer in solution. *J. Biol. Chem.*, 278, 35337-35344.
- Borges, J.C. & Ramos, C.H.I. (2004) Protein folding assisted by chaperones. *Protein Pept. Lett.* (submetido) – Anexo I
- Boston, R.S., Viitanen, P.V. & Vierling, E. (1996) Molecular chaperones and protein folding in plants. *Plant. Mol. Biol.*, 32, 191-222.

- Brehmer, D., Rüdiger, S., Gässler, C.S., Klostermeier, D., Packschies, L., Reinsten, J., Mayer, M.P. & Bukau, B. (2001) Tuning of chaperone activity of Hsp70 proteins by modulation of nucleotide exchange. *Nat. Struct. Biol.*, 8, 427-432.
- Brehmer, D., Gassler, C., Rist, W., Mayer, M.P. & Bukau, B. (2004) Influence of GrpE on DnaK-substrate interactions. *J. Biol. Chem.*, 279, 27957-27964.
- Briknarova, K., Takayama, S., Brive, L., Havert, M.L., Knee, D.A., Velasco, J., Homma, S., Cabezas, E., Stuart, J., Hoyt, D.W., Satterthwait, A.C., Llinas, M., Reed, J.C. & Ely, K.R. (2001) Structural analysis of BAG1 cochaperone and its interactions with Hsc70 heat shock protein. *Nat. Struct. Biol.*, 8, 349-352.
- Bukau, B. & Horwich, A.L. (1998) The Hsp70 and Hsp60 chaperone machines. *Cell*, 92, 351-366.
- Carrasco, B. & García de la Torre, J. (1999) Hydrodynamic Properties of Rigid Particles: Comparison of Different Modeling and Computational Procedures. *Biophys. J.*, 76, 3044-3057.
- Chai, Y., Koppenhafer, S.L., Bonini, N.M. & Paulson, H.L. (1999) Analysis of the role of heat shock protein (Hsp) molecular chaperones in polyglutamine disease. *J. NeuroSci.*, 19, 10338-10347.
- Cheetham, M.E. & Caplan, A.J. (1998) *Structure*, function and evolution of DnaJ: conservation and adaptation of chaperone function. *Cell Stress Chaperones*, 3, 28-36.
- Chesnokova, L.S., Slepnev, S.V., Protasevich, I.I., Sehorn, M.G., Brouillette, C.G. & Witt, S.N. (2003) Deletion of DnaK's lid strengthens binding to the nucleotide exchange factor, GrpE: a kinetic and thermodynamic analysis. *Biochemistry*, 42, 9028-9040.
- Choglay, A.A., Chapple, J.P., Blatch, G.L. & Cheetham, M.E. (2001) Identification and characterization of a human mitochondrial homologue of the bacterial co-chaperone GrpE. *Gene*, 267, 125-134.
- Cole, J.L. & Hansen, J.C. (1999) Analytical Ultracentrifugation as a contemporary biomolecular research tool. *J. Biomol. Tech.* 10, 163-176.
- Cyr, D.M., Langer, T. & Douglas, M.G. (1994) DnaJ-like proteins: molecular chaperones and specific regulators of Hsp70. *Trends Biochem. Sci.*, 19, 176-81.
- Doniach, S., Basile, J., Garel, T. & Orland, H. (1995) Partially folded states of proteins: characterization by X-ray scattering. *J. Mol. Biol.*, 254, 960-967.
- Easton, D.P., Kaneko, Y. & Subject, J.R. (2000) The hsp110 and Grp170 stress proteins: newly recognized relatives of the Hsp70s. *Cell Stress Chaperones*, 5, 276-290.

- Edelhock, H. (1967) Spectroscopic Determination of Tryptophan and Tyrosine in Proteins. *Biochemistry*, 6, 1948-1954.
- Feigin, L.A. & Svergun, D.I. (1987) Structure analysis by Small-angle X-ray and Neutron Scattering, pp. 83-87, Plenum Press, New York.
- Fink, A.L. (1998) Protein aggregation: folding aggregates, inclusion bodies and amyloid. *Fold Des.*, 3, R9-23.
- Fink, A.L. (1999) Chaperone-mediated protein folding. *Physiol. Rev.*, 79, 425-449.
- Flaherty, K.M., DeLuca-Flaherty, C. & McKay, D.B. (1990) Three-dimensional structure of the ATPase fragment of a 70K heat-shock cognate protein. *Nature*, 346, 623-628.
- Frydman, J., Nimmesgern, E., Ohtsuka, K. & Hartl, F.U. (1994) Folding of nascent polypeptide chains in a high molecular mass assembly with molecular chaperones. *Nature*, 370, 111-117.
- García de la Torre, J., Carrasco, B. & Harding, S.E. (1997) SOLPRO: theory and computer program for the prediction of SOLution PROPERTIES of rigid macromolecules and bioparticles. *Eur. Biophys. J.*, 25, 361-372.
- García de la Torre, J., Huertas, M.L. & Carrasco, B. (2000) Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys. J.*, 78, 719-730.
- Gelinas, A.D., Langsetmo, K., Toth, J., Bethoney, K.A., Stafford, W.F. & Harrison, C.J. (2002) A Structure-based Interpretation of E.coli GrpE Thermodynamic Properties. *J. Mol. Biol.*, 323, 131-142.
- Gelinas, A.D., Toth, J., Bethoney, K.A., Langsetmo, K., Stafford, W.F. & Harrison, C.J. (2003) Thermodynamic linkage in the GrpE nucleotide exchange factor, a molecular thermosensor. *Biochemistry*, 42, 9050-9059.
- Gething, M.J. & Sambrook, J. (1992) Protein folding in the cell. *Nature*, 355, 33-45.
- Gill, S.C. & von Hippel, P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.*, 182, 319-326.
- Glover, R. & Lindquist, S. (1998) Hsp104, Hsp70, and Hsp40: A novel chaperone system that rescues previously aggregated proteins. *Cell*, 94, 73-82.
- Goldberg, R.J. (1953) Sedimentation in the ultracentrifuge. *J. Phys. Chem.*, 57, 194-202.
- Goloubinoff, P., Mogk, A., Zvi, A.V.I., Tomoyasu, T. & Bukau, B. (1999) Sequential mechanism of solubilization and refolding of stable protein aggregates by a bichaperone network. *Proc. Natl. Acad. Sci. U.S.A.*, 96, 13732-13737.

- Greene, M.K., Maskos, K. & Landry, S.J. (1998) Role of the J-domain in the cooperation of Hsp40 with Hsp70. *Proc. Natl. Acad. Sci. U.S.A.*, 95, 6108-6113.
- Greenfield, N. J. (1996) Methods to estimate the conformation of proteins and polypeptides from circular dichroism data. *Anal. Biochem.*, 235, 1-10.
- Grimshaw, J.P., Jelesarov, I., Schönfeld, H.J. & Christen, P. (2001) Reversible thermal transition in GrpE, the nucleotide exchange factor of the DnaK heat-shock system. *J. Biol. Chem.*, 276, 6098-6104.
- Grimshaw, J.P., Jelesarov, I., Siegenthaler, R.K. & Christen, P. (2003) Thermosensor action of GrpE. The DnaK chaperone system at heat shock temperatures. *J. Biol. Chem.*, 278, 19048-19053.
- Groemping, Y. & Reinstein, J. (2001) Folding properties of the nucleotide exchange factor GrpE from *Thermus thermophilus*: GrpE is a thermosensor that mediates heat shock response. *J. Mol. Biol.*, 314, 167-178.
- Guinier, A. & Fournet, G. (1955) *Small-Angle Scattering of X-Rays* - John Wiley & Sons, New York.
- Ha, J-H. & McKay, D.B. (1994) ATPase kinetics of recombinant bovine 70 kDa heat shock cognate protein and its amino-terminal ATPase domain. *Biochemistry*, 33, 14625-14635.
- Harrison, C.J., Hayer-Hartl, M., Di Liberto, M., Hartl, F.U. & Kuriyan J. (1997) Crystal Structure of the Nucleotide Exchange Factor GrpE Bound to the ATPase Domain of the Molecular Chaperone DnaK. *Science*, 276, 431-435.
- Harrison, C.J. (2003) GrpE, a nucleotide exchange factor for DnaK. *Cell Stress Chaperones*, 8, 218-224.
- Hartl, F.U. & Hayer-Hartl, M. (2002) Molecular chaperones in the cytosol, from nascent chain to folded protein. *Science*, 295, 1852-1858.
- Hendrick, J.P. & Hartl, F.U. (1993) Molecular chaperone function of heat-shock proteins. *Annu. Rev. Biochem.*, 62, 349-84.
- Hendrick, J.P., Langer, T., Davis, T.A., Hartl, F.U. & Wiedmann, M. (1993) Control of Folding and Membrane Translocation by Binding of the Chaperone DnaJ to Nascent Polypeptides. *Proc. Natl. Acad. Sci. U.S.A.*, 90, 10216-10220.
- Jaenicke, R. (1987) Folding and association of proteins. *Prog. Biophys. Mol. Biol.*, 49, 117-237.
- Jaenicke, R. (1999) Stability and folding of domain proteins. *Prog. Biophys. Mol. Biol.*, 71, 155-241.

- Jelesarov, I. & Bosshard, H. (1999) Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition. *J. Mol. Recognit.*, 12, 3-18.
- Johnson, M.L., Correia, J.J., Yphantis, D.A. & Halvorson, H.R. (1981) Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques. *Biophys. J.*, 36, 575-588.
- Karzai, A.W. & MaMacken, R. (1996) A bipartite signaling mechanism involved in DnaJ-mediated activation of the Escherichia coli DnaK protein. *J. Biol. Chem.*, 271, 11236-11246.
- Kelly, S. M. & Price, N. C. (1997) The application of circular dichroism to studies of protein folding and unfolding. *Biochim. Biophys. Acta*, 1338, 161-185
- Kellermann, G., Vicentin, F., Tamura E., Rocha, M., Tolentino, H., Barbosa, A., Craievich, A. & Torriani, I. (1997) The small-angle X-ray scattering beamline at the Brazilian synchrotron light laboratory. *J. Appl. Crystallog.* 30, 880-883.
- Koch, M.H, Vachette, P. & Svergun, D.I. (2003) Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. *Q. Rev. Biophys.*, 36, 147-227
- Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head bacteriophage T4. *Nature*, 227, 680-5.
- Lakowicz, J. R. (1983) In: Principles of fluorescence spectroscopy. Plenum Press, New York, 341 – 379.
- Langer, T., Lu, C., Echols, H., Flanagan, J., Hayer, M.K. & Hartl, F.U. (1992) Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding. *Nature*, 356, 683-689.
- Laue, T.M., Shal, B.D., Ridgeway, T.M. & Pelletier, S.M. (1992) Analytical ultracentrifugation in biochemistry and polymer science. Pp. 90-125, Eds. S. Harding & A. Rowe, Royal Society of Chemistry, London.
- Laue, T.M. & Stafford, W.F. (1999) Modern applications of analytical ultracentrifugation. *Annu. Rev. Biophys. Biomol. Struct.*, 28, 75-100.
- Laue, T.M. (2001) Biophysical studies by ultracentrifugation. *Curr. Opin. Struct. Biol.*, 11, 579-583.

- Laufen, T., Mayer, M.P., Beisel, C., Klostermeier, D., Mogk, A., Reistein, J. & Bukau, B. (1999) Mechanism of regulation of hsp70 chaperones by DnaJ cochaperones. *Proc. Natl. Acad. Sci. U.S.A.*, 96, 5452-5457.
- Lebowitz, J., Lewis, M.S. & Schuck, P. (2002) Modern analytical ultracentrifugation in protein science: a tutorial review. *Prot. Sci.*, 11, 2067-2079.
- Lee, S., Fan, C.Y., Younger, J.M., Ren, H. & Cyr, D.M. (2002) Identification of essential residues in the type II Hsp40 Sis1 that function in polypeptide binding. *J. Biol. Chem.*, 277, 21675-21682.
- Levitt, M. & Chothia, C. (1976) Structural Patterns in Globular Proteins. *Nature*, 261, 552-558.
- Li, J., Qian, X. & Sha, B. (2003) The crystal structure of the yeast Hsp40 Ydj1 complexed with its peptide substrate. *Structure*, 11, 1475-1483.
- Liberek, K., Marszalek, J., Ang, D., Georgopoulos, C. & Zylicz, M. (1991) Escherichia coli DnaJ and GrpE heat shock proteins jointly stimulate ATPase activity of DnaK. *Proc. Natl. Acad. Sci. U.S.A.*, 88, 2874-2878.
- Liberek, K., Wall, D. & Georgopoulos, C. (1996) The DnaJ chaperone catalitically activates the DnaK chaperone to preferentially bind the s32 heat shock transcriptional regulator. *Proc. Natl. Acad. Sci. U.S.A.*, 92, 6224-6228.
- Linke, K., Wolfram, T., Bussemer, J., & Jakob, U. (2003) The roles of the two zinc binding sites in DnaJ. *J. Biol. Chem.* 278, 44457-44466.
- Mally, A. & Witt, S.T. (2001) GrpE accelerates peptide binding and release from the high affinity state of DnaK. *Nat. Struct. Biol.*, 8, 254-257.
- Martin, J. & Hartl, F.U. (1997) Chaperone-assisted protein folding. *Curr. Opin. Struct. Biol.*, 7, 41-52.
- Martinez-Yamout, M., Legge, G.B., Zhang, O., Wright, P.E. & Dyson, H.J. (2000) Solution Structure of the Cysteine-Rich Domain Of The *Escherichia Coli* Chaperone Protein DnaJ. *J. Mol. Biol.*, 300, 805-818.
- Matouschek, A. (2003) Protein unfolding – an important process in vivo? *Curr. Opin. Struct. Biol.*, 13, 98-109.
- Mayer, M.P., Schröder, H., Rüdiger, S., P aminoácidos I, K., Laufen, T. & Bukau, B. (2000) Multispet mechanism of substrate binding determines chaperone activity of Hsp70. *Nature Struct. Biol.*, 7, 586-593.
- Mayer, M.P., Brehmer, D., Gässler, C.S. & Bukau, B. (2001) Hsp70 Chaperone Machines. *Adv. Protein Chem.*, 59, 1-45.

- Mehl, A.F., Heskett, L.D., Jain, S.S. & Demeler, B. (2003) Insights into dimerization and four-helix bundle formation found by dissection of the dimer interface of the GrpE protein from *Escherichia coli*. *Prot. Sci.*, 12, 1205-1215.
- Miernyk, J.A. (1999) Protein folding in the plant cell. *Plant Physiol.*, 121, 695-703.
- Mogk, A., Tomoyasu, T., Goloubonoff, P., Rüdiger, S., Röder, D., Langen, H. & Bukau, B. (1999) Identification of thermolabile *Escherichia coli* proteins: prevention and reversion of aggregation by DnaK and ClpB. *EMBO J.*, 18, 6934-6949.
- Montgomery, D.L.; Morimoto, R.I. & Gierasch, L.M. (1999) Mutations in the Substrate Binding Domain of the *Escherichia coli* 70 kDa Molecular Chaperone, DnaK, which Alter Substrate Affinity or Interdomain coupling. *J. Mol. Biol.*, 286, 915-32.
- Naylor, D.J., Stines, A.P., Hoogenraad, N.J. & Hoj. P.B. (1998) Evidence for the Existence of Distinct Mammalian Cytosolic, Microsomal, and Two Mitochondrial GrpE-like Proteins, the Co-chaperones of Specific Hsp70 Members. *J. Biol. Chem.*, 273, 21169-21177.
- O'Brien, M.C. & McKay, D.B. (1995) How potassium affects the activity of the molecular chaperone Hsc70. I. Potassium is required for optimal ATPase activity. *J. Biol. Chem.*, 270, 2247-2250.
- Oliveira, C.L.P. (2003) "TRAT1D – Computer Program for SAXS Data Treatment", LNLS technical Manual MT 01/2003.
- Osipiuk, J., Walsh, M. A., Freeman, B. C., Morimoto, R. I. & Joachimiak, A. (1999) Structure of a new crystal form of human Hsp70 ATPase domain. *Acta Crystallogr. D. Biol. Crystallogr.*, 55, 1105-1107.
- Palleros, D.R., Welch, W.J. & Fink, A.L. (1991) Interaction of Hsp70 with unfolded proteins: Effects of temperature and nucleotides on the kinetics of binding. *Proc. Natl. Acad. Sci. U.S.A.*, 88, 5719-5723.
- Palleros, D.R., Reid, K.L., McCarty, J.S., Walker, G.C. & Fink, A.L. (1992) DnaK, hsp73, and their molten globules. Two different ways heat shock proteins respond to heat. *J. Biol. Chem.*, 267, 5279-5285.
- Palleros, D.R., Reid, K.L., Shi, L., Welch, W.J. & Fink, A.L. (1993) ATP-induced protein-Hsp70 complex dissociation requires K⁺ but not ATP hydrolysis. *Nature*, 365, 664-666.
- Pellecchia, M., Montgomery, D.L., Stevens, S.Y., Vander Kooi, C.W., Feng, H.P., Gierasch, L.M. & Zuiderweg, E.R. (2000) Structural insights into substrate binding by the molecular chaperone DnaK. *Nat. Struct. Biol.*, 4, 298-303.

- Pierce, M.M., Raman, C.S. & Nall, B.T. (1999) Isothermal titration calorimetry of protein-protein interactions. *Methods.*, 19, 213-221.
- Pierpaoli, E.V., Sandmeier, E., Baici, A., Schönfeld, H.J., Gisler, S. & Christen, P. (1997) The power stroke of the DnaK/DnaJ/GrpE molecular chaperone system. *J. Mol. Biol.*, 269, 757-768.
- Pierpaoli, E.V., Sandmeier, E., Schönfeld, H.J. & Christen, P. (1998) Control of the DnaK chaperone cycle by substoichiometric concentrations of the co-chaperones DnaJ and GrpE. *J. Biol. Chem.*, 273, 6643-6649.
- Porod, G. (1982) in: Small-angle X-ray Scattering (Glatter, O. & Kratky, O., Eds) pp. 17-51, Academic Press, London.
- Privalov, P.L. (1996) Intermediate states in protein folding. *J. Mol. Biol.*, 258, 707-725.
- Ptitsyn, O.B. (1991) How does protein synthesis give rise to the 3D-structure? *FEBS Lett.*, 285, 176-181.
- Qian, X., Hou, W., Zhengang, L., & Sha, B. (2002) Direct interactions between molecular chaperones heat-shock protein (Hsp) 70 and Hsp40: yeast Hsp70 Ssa1 binds the extreme C-terminal region of yeast Hsp40 Sis1. *Biochem. J.*, 361, 27-34.
- Ralston, G. (1993) Introduction to Analytical Ultracentrifugation. Fullerton, CA: Beckman Instruments.
- Ruddon, R.W. & Bedows, E. (1997) Assisted protein folding. *J. Biol. Chem.*, 272, 3125-128.
- Rüdiger, S., Germeroth, L., Schneider-Mergener, J. & Bukau, B. (1997) Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *EMBO J.*, 16, 1501-1507.
- Rüdiger, S., Schneider-Mergener, J. & Bukau, B. (2001) Its substrate specificity characterizes the DnaJ co-chaperone as a scanning factor for the DnaK chaperone. *EMBO J.*, 20, 1042-1050.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) In: Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory Press (Ed.) New York, United State of America.
- Santoro, M.M. & Bolen, D.W. (1988) Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry*, 27, 8063-8068.
- Schlicher, T. & Soll, J. (1997) Chloroplastic isoforms of DnaJ and GrpE in pea. *Plant. Mol. Biol.*, 33, 181-185.

- Schlossman, D.M., Schmid, S.L., Braell, W.A. & Rothman, J.E. (1984) An enzyme that removes clathrin coats: purification of an uncoating ATPase. *J. Cell. Biol.*, 99, 723-733.
- Schönfeld, H.J., Schmidt, D., Schröder, H. & Bukau, B. (1995) The DnaK chaperone system of *Escherichia coli*: quaternary structures and interactions of the DnaK and GrpE components. *J. Biol. Chem.*, 270, 2183-2189.
- Schönfeld, H.-J. & Behlke, J. (1998) Molecular chaperones and their interactions investigated by analytical ultracentrifugation and other methodologies. *Methods Enzymol.*, 290, 269-96.
- Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* 78, 1606-1619.
- Semenyuk, A.V. & Svergun, D.I. (1991) GNOM – a program package for small-angle scattering data processing. *J. Appl. Cryst.*, 24, 537-540.
- Sha, B., Lee, S. & Cyr, D.M. (2000) The crystal structure of the peptide binding fragment from the yeast Hsp40 protein Sis1. *Structure*, 8, 799-807.
- Shi, L., Kataoba, M. & Fink, A.L. (1996) Conformational Characterization of DnaK and Its Complexes by Small-Angle X-ray Scattering. *Biochemistry*, 96, 3297-308.
- Shnyrov, V.L., Zhadan, G.G., Cobaleda, C., Sagrera, A., Munoz-Barroso, I. & Villar, E. (1997) A differential scanning calorimetric study of Newcastle disease virus: identification of proteins involved in thermal transitions. *Arch. Biochem. Biophys.*, 341, 89-97.
- Stafford, W.F. (1994) Boundary analysis in sedimentation velocity experiments. *Methods Enzymol.*, 240, 478-501.
- Stafford, W.F. (1997) Sedimentation velocity spins a new weave for an old fabric. *Curr. Opin. Biotechnol.*, 8, 14-24.
- Stuart, J.K., Myszka, D.G., Joss, L., Mitchell, R.S., McDonald, S.M., Xie, Z., Takayama, S., Reed, J.C. & Ely, K.R. (1998) Characterization of Interactions between the Anti-apoptotic Protein BAG-1 and Hsc70 Molecular Chaperones. *J. Biol. Chem.*, 273, 22506-22514.
- Studier, W.F., Rosenberg, A.H., Dunn, J.J. & Dubendorff, J.W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.*, 185, 60-89.
- Suh, W-C., Lu, C.Z. & Gross, C.A. (1999) Structural Features Required for the Interaction of the Hsp70 Molecular Chaperone DnaK with Its Cochaperone DnaJ. *J. Biol. Chem.*, 274, 30534-30539.
- Svergun, D.I. & Stuhrmann, H.B. (1991) New developments in direct shape determination from small-angle scattering. 1. Theory and model-calculations. *Acta Cryst.*, A47, 736-744.

- Svergun, D.I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.*, 76, 2879-2886.
- Svergun, D.I., Petoukhov, M.V. & Koch, M.H.J. (2001) Determination of domain structure of proteins from X-ray solution scattering. *Biophys. J.*, 80, 2946-2953.
- Svergun, D.I. & Koch, M.H.J. (2003) Small-angle scattering studies of biological macromolecules in solution. *Rep. Prog. Phys.*, 66, 1735-1782.
- Szabo, A., Langer, T., Schroder, H., Flanagan, J., Bukau, B. & Hartl, F.U. (1994) The ATP hydrolysis-dependent reaction cycle of the Escherichia coli Hsp70 system DnaK, DnaJ, and GrpE. *Proc. Natl. Acad. Sci. U.S.A.*, 91, 10345-10349.
- Szabo, A., Korszun, R., Hartl, F.U. & Flanagan, J. (1996) A zinc finger-like domain of the molecular chaperone DnaJ is involved in binding to denatured protein substrates. *EMBO J.*, 15, 408-417.
- Szyperski, T., Pellechia, M., Wall, D., Georgopoulos, C. & Wuthrich, K. (1994) NMR Structure Determination of the Escherichia coli DnaJ Molecular Chaperone: Secondary Structure and Backbone Fold of the N-Terminal Region (Residues 2-108) Containing the Highly Conserved J Domain. *Proc. Natl. Acad. Sci. U.S.A.*, 91, 11343-11347.
- Tama, F., Miyashita, O. & Brooks, C.L.3rd. (2004) Flexible Multi-scale Fitting of Atomic Structures into Low-resolution Electron Density Maps with Elastic Network Normal Mode Analysis. *J. Mol. Biol.*, 337, 985-999.
- Takayama, S. & Reed, J.C. (2001) Molecular chaperone targeting and regulation by BAG family proteins. *Nat. Cell Biol.*, 3, E237-241.
- Theysen, H., Schuster, H-P., Packschies, L., Bukau, B. & Reinstein, J. (1996) The Second Step of ATP Binding Induces Peptide Release. *J. Mol. Biol.*, 263, 657-670.
- Van Holde, K.E. & Weischet, W.O. (1978) Boundary analysis of sedimentation-velocity experiments with monodisperse and paucidisperse solutes. *Biopolymers*, 17, 1397-1403.
- Venter, J.C., Adams, M.D., Myers, E.W., *et al.* (2001) The sequence of the human genome. *Science*, 291, 1304-1351.
- Volkov, V.V. & Svergun, D.I. (2003) Uniqueness of *ab initio* shape determination in small-angle scattering. *J. Appl. Cryst.*, 36, 860-864.
- Waxman, E., Laws, W.R., Laue, T.M. & Ross, J.B.A. (1994) Refining hydrodynamics shape of proteins: The combination of data from analytical ultracentrifugation and time-resolved fluorescence anisotropy decay. pp. 189-208. in *Modern Analytical Ultracentrifugation*. (Eds Todd M. Schuster & Thomas M. Laue) Birkhäuser, Boston.

- Wilbanks, S.M. & McKay, D.B. (1995) How potassium affects the activity of the molecular chaperone Hsc70. II. Potassium binds specifically in the ATPase active site. *J. Biol. Chem.*, 270, 2251-2257
- Wilson, W.W. (2003) Light scattering as a diagnostic for protein crystal growth – A practical approach. *J. Struct. Biol.*, 42, 56-65.
- Woody, R.W. (1995) Circular dichroism. *Methods Enzymol.*, 246, 34-71.
- Yan, W. & Craig, E.A. (1999) The Glycine-Phenylalanine-Rich Region Determines the Specificity of the Yeast Hsp40 Sis1. *Mol. Cell. Biol.*, 19, 7751-7758.
- Zhu, X., Zhao, X., Burkholder, W.F., Gragerov, A., Ogata, C.M., Gottesman, M.E. & Hendrickson, W.A. (1996) Structural analysis of substrate binding by the molecular chaperone DnaK. *Science*, 272, 1606-1614.

7.1 Lista de endereços na rede mundial de computadores – Internet

BLAST: www.ncbi.nlm.nih.gov/BLAST/

CDNN Deconvolution (Versão 2): Bioinformatik.biochemtech.uni-halle.dee/cdnn

ClustalW: clustalw.genome.ad.jp/

Deep-View/Swiss-PDB Viewer: www.expasy.ch/swiss-mod/SWISS-MODEL.html

LALIGN: www.ch.embnet.org/software/LALIGN_form.html

MITOPROT II 1.0a4: www.mips.biochem.mpg.de/cgi-bin/proj/medgen/mitofilter

NCBI: www.ncbi.nlm.nih.gov

Oligo Analysis & Plotting Tool: oligos.qiagen.com/oligos/toolkit.php

PDB: pd-beta.rcsb.org/pdb/Welcome.do (Berman *et al.*, 2000)

Portal Periódicos: www.periodicos.capes.gov.br/

PredictProtein: cubic.bioc.columbia.edu/predictprotein/

ProtParam Tool: ca.expasy.org/tools/protparam.html

Sednterp: www.jphilo.mailway.com/download.htm

Swiss-Model: www.expasy.ch/swiss-mod/SWISS-MODEL.html

WebLab ViewerLite: www.accelrys.com

8 APÊNDICES

8.1 Anexo I

Borges, J.C. & Ramos, C.H.I. (2004) Protein folding assisted by chaperones. *Protein Pept. Lett.* (submetido).

Protein folding assisted by chaperones

Júlio C. Borges and Carlos H.I. Ramos*

Centro de Biologia Molecular Estrutural, Laboratório Nacional de Luz Síncrotron, CP 6192, 13084-971, Campinas, SP, Brazil. Departamento de Bioquímica, Instituto de Biologia, Universidade Estadual de Campinas, Campinas SP, Brazil. E-mail: cramos@lnls.br

Abstract

Molecular chaperones are one of the most important cell defense mechanisms against protein aggregation and misfolding. These specialized proteins bind non-native states of other proteins and assist them in reaching a correctly folded and functional conformation. Chaperones also participate in protein translocation by membranes, in the stabilization of unstable protein conformers and regulatory factors, in the delivery of substrates for proteolysis and in the recovery of proteins from aggregates.

Key Words: Molecular chaperones, Heat shock proteins, Hsp70, chaperonin.

Molecular chaperones and heat-shock proteins

The term “molecular chaperone” was first used to describe the function of nucleoplasmin, a protein that promotes the correct oligomerization of histone by preventing its aggregation during assembly in chromatin [1]. This concept was further developed by Hendrick and Hartl [2] who stated: “we define a molecular chaperone as a protein that binds to and stabilizes an otherwise unstable conformer of another protein, facilitates its correct fate in vivo: be it in folding, oligomeric assembly, transport to a particular subcellular compartment, or controlled switching between active/inactive conformations”. Chaperone proteins were initially referred to as heat shock proteins (Hsp) because of their increased abundance in cells exposed to thermal stress, although many of these proteins are constitutively expressed [3]. The importance of chaperone proteins is indicated by the high number of publications in this area and by the considerable number of reviews covering this subject [2,4-8].

Chaperones can assist the folding of an unfolded protein within cells by recognizing exposed hydrophobic surfaces that will be buried in the native state, but without modifying the final native structure of the protein. This characteristic does not contradict Anfinsen's classic theory [9] that the acquisition of the specific three-dimensional structure of a protein depends exclusively on the amino acid sequence of the individual protein.

There are at least five main molecular chaperone families that recognize and bind to hydrophobic surfaces on unfolded or partially folded proteins, and are classified by their molecular mass (Table 1). We present a brief description of these main families. The lack of space precludes a detailed discussion of the other groups.

Table 1: The main chaperone families and their properties and major co-chaperones (see Figure 1).

Chaperone	Co-chaperone	Mass (kDa)	n-mers	General comments
Hsp100	?	~100	6-7	ATPase activity; recovers proteins from aggregates; form proteolytic complexes.
Hsp90	HIP, HOP, p23	~90	2	Prevents protein aggregation; binds ATP; interacts with large range of proteins; constitutive.
Hsp70	Hsp40, GrpE	~70 ~40*, ~21*	1 2*, 2*	ATPase activity; binds nascent peptide chains; involved in folding, transport through membranes, degradation, etc.
Hsp60 (GroEL)	Hsp10 (GroES)	~60 ~10*	14 7*	ATPase activity; assists protein folding.
smHsp	?	~15-30	9-24	ATP-independent; involved in thermotolerance and escape from aggregation by keeping proteins in a refoldable state

*co-chaperone

Hsp70 – a pivot chaperone

One of the most important families of chaperones is the Hsp70 system formed by Hsp70 (or DnaK in *E. coli*) and its co-chaperones Hsp40 (DnaJ) and GrpE. This system assists many cellular processes involving proteins, including folding, transport through membranes, degradation and escape from aggregation [8,10]. Hsp70 functions as a pivot that receives and distributes unfolded proteins from and to other molecular chaperones (Fig. 1). Hsp70 transports proteins from ribosomes to the Hsp60 system for folding [11-13], participates in a system with Hsp100 that recovers protein aggregates [14-16], and cooperates with smHsp to reactivate heat-denatured proteins [17-19]. Several reports have shown that Hsp70 also cooperates with the Hsp90 system to form a foldosome, or cytoplasmic chaperone complex, that can interact with a high number of other chaperones, co-chaperones (p23, Hip and Hop), foldases (PDIase and PPIase), steroid hormone receptors, transcription factors, and others [20-22].

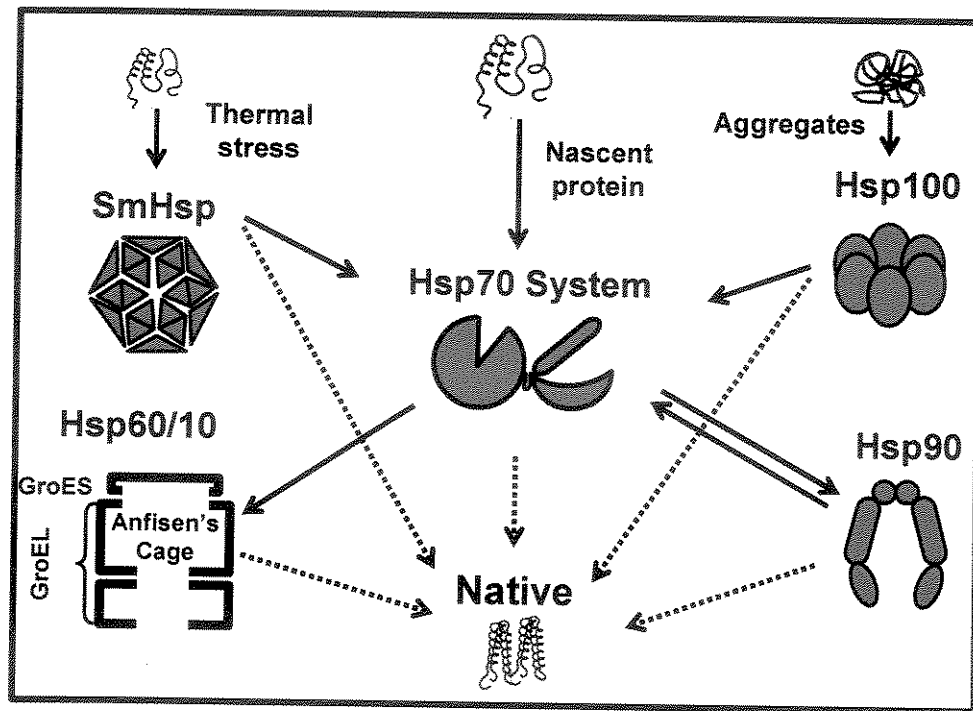


Figure 1: The pivotal role of the Hsp70 system. Hsp70 receives substrates directly from ribosomes and from the Hsp90 systems, and can receive substrates from Hsp100 and SmHsp to recover unfolded proteins. In addition, Hsp70 distributes substrates to the Hsp60 and Hsp90 systems. Dashed arrows: protein folding; Solid arrows: delivery of unfolded or partially folded proteins.

Hsp70 consists of a nucleotide-binding domain (NBD, ~45 kDa) that binds adenine nucleotide and has ATPase activity, and a substrate-binding domain (SBD, ~25 kDa) that binds to short, extended hydrophobic segments (Fig. 1) [23,24]. The affinity of SBD for hydrophobic residues is regulated by nucleotide binding in the NBD; the binding of ADP increases the affinity whereas the binding of ATP decreases the affinity for hydrophobic residues [8,10]. ADP binding causes a sequential conformational change, first in the NBD and then in the SBD, that shifts the equilibrium of the latter to the closed conformation but is reversed when ATP is bound [25,26]. The NBD and SBD domains can be subdivided in two sub-domains each: NBD1 and NBD2 and SBD1 and SBD2. SBD1 is formed mainly by β -sheets and functions as a “bed” for the substrate, whereas SBD2 consists of a long α -helix and functions as a “lid” [23,26].

Although the Hsp70 family is structurally and functionally highly conserved, these proteins are classified in three subfamilies based on their affinity for ADP: DnaK>Hsc70>HscA [27]. The DnaK subfamily has an exposed loop in the NBD (A276-R302 in *E. coli*) that is four and 10 amino acids longer than the Hsc70 and HscA loops, respectively. The DnaK subfamily is the only one with a hydrophobic patch at the top of its cleft and two putative salt bridges responsible for interface polarity.

The different properties of the subfamilies may explain the need for GrpE and Bag co-chaperones during nucleotide exchange by DnaK and Hsc70, respectively [27].

The cycle of ATP binding, hydrolysis and replacement of ADP for ATP is assisted by Hsp70 co-chaperones, the most studied of which are: Hsp40, GrpE, Hip (Hsp-interacting protein) and Hop (Hsp70 organization protein that participates in the Hsp70-Hsp90 heterocomplex) [26]. Hsp40 proteins are ubiquitous, diverse, and differ from each other in the architecture of their domains and in their cellular location. Hsp40 binds to unfolded peptides and interacts with ATP-ligated Hsp70 to promote ATP hydrolysis and consequent folding of the peptide [28-30]. Hsp40 can also function as a chaperone by hydrophobic interaction with substrates [11,28,30-33], and may act as a substrate scanning factor for Hsp70 proteins [24]. Hsp40 consists of four regions [34], from the N- to the C-terminus: 1) An α -helix J_domain in the N-terminus that is characteristic of proteins from this family and binds to Hsp70 [28,35], 2) A disordered region rich in glycine/phenylalanine that may be responsible for flexibility as a spacer between adjacent regions and is also essential for Hsp40 to function since the J_domain alone is unable to bind to Hsp70 proteins [29,25], 3) A central region with a cysteine-rich domain that enclose four repeats of the motif CXXCXGXG (where X is any amino acid) and folds in a zinc-dependent fashion with two repeats binding to one zinc [36,37], and 4) A C-terminal domain formed by a β -sheet structure and that is involved in the protein dimerization [38]. The cysteine-rich and the C-terminal domains are involved in substrate binding and presentation [32,33,36]. Based on their architecture, Hsp40 proteins can be classified into three main subfamilies: subfamily A that consists of the four domains described above, subfamily B that lacks the cysteine-rich domain, and subfamily C that has only the J-domain that is not necessarily located at the N-terminus [10, 34].

GrpE is essential for bacterial viability at all temperatures [11] because it acts as an exchange factor that releases nucleotides bound to DnaK [28] and is important for DnaK recycling [39]. GrpE is a symmetric dimer in solution [40] and becomes asymmetric when bound to the DnaK-NBD [41]. The GrpE has an elongated shape and consists of four regions: 1) A β -structured domain at the C-terminus (EcGrpE138-197) that binds to the DnaK-NBD causing a mechanical opening movement that releases ADP from the NBD [27,41], 2) A middle region formed by a four-helix bundle (EcGrpE89-137) that acts as a stabilization center for dimerization [42], 3) An elongated helix connected to a α -structured region (EcGrpE40-88) that works as a thermosensor [42,43], and 4) A region composed of the first 34 residues that enhances binding in the SBD of DnaK by reducing its affinity for substrate [41,44]. In contrast to prokaryotes, eukaryotes have at least two isoforms of GrpE in the mitochondria [45,46]. These isoforms can stimulate the ATPase activity of mitochondrial Hsp70 [45].

Chaperonins – the Anfinsen’s cage

GroEL (Hsp60) and its co-chaperone GroES (Hsp10) are one of the best-studied complexes involved in protein folding, as has been extensively reviewed elsewhere [5,6,8,47,48]. Hsp60 and Hsp10 are ubiquitous proteins that assist the folding of a large variety of newly synthesized proteins and newly translocated proteins [5,8,47]. The chaperonin mechanism of action consists in the capture of a substrate within its central channel – the Anfinsen’s cage – that, at the cost of ATP, provides an adequate environment for folding. The chaperonin crystal structure and mechanism of action have been analyzed in detail by Xu *et al.* [49] and are summarized in Figure 2. The chaperonin receives substrates from the Hsp70 system to mediate their folding (Fig. 1) [11-13]. There are two types of chaperonins and they show little similarity with each other. One type includes the classic GroEL and GroES (chaperonins) from bacteria, mitochondria and chloroplasts (initially referred to as RUBISCO binding protein), and the cytoplasmic homologues TCP (Tric complex proteins) [8,50].

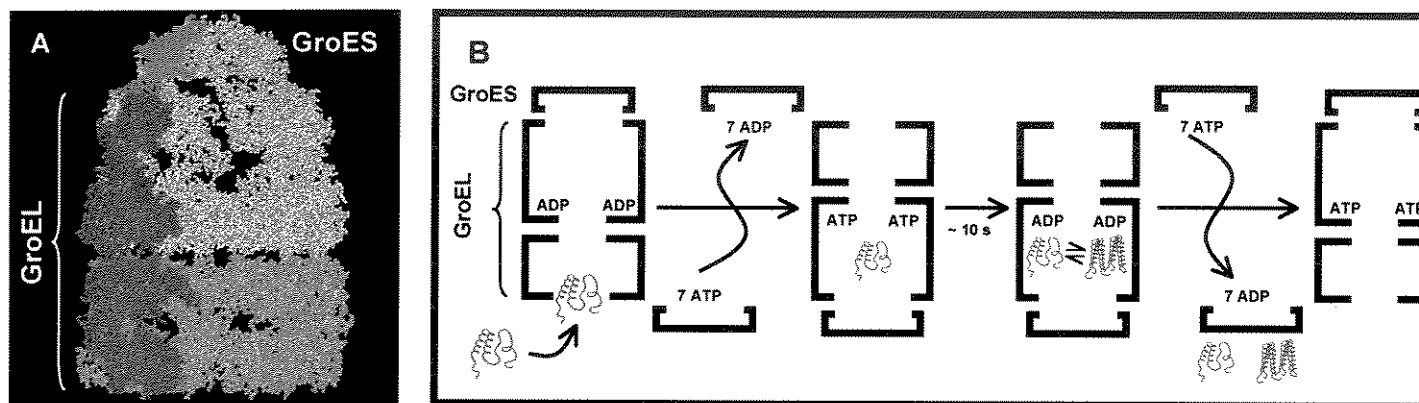


Figure 2. The chaperonin system. A) The chaperonin complex is active as a double ring oligomer structure with a cylindrical shape and a mass of almost 1 million Daltons. The chamber is composed of two heptamers of GroEL (GroES-bound monomer is pink and unliganded monomer is orange) arranged as two stacked rings with a GroES heptamer (monomer in green) at the top acting as a lid. B) A substrate protein with a hydrophobic surface can bind to the hydrophobic surface in the apical domain of GroEL (trans complex, shown on the left). The binding of ATP causes a large conformational change in GroEL, creating a cage and duplicating its initial internal volume. GroES functions as a lid and closes the cavity of the ring (cis complex, shown in the middle). The hydrolysis of ATP induces new conformational changes that switch the internal surface from hydrophobic to hydrophilic, thus allowing the substrate to undergo proper folding. The binding of ATP releases GroES, and the substrate, if not yet folded, can rebind to GroEL and undergo a new cycle.

The TCP, a double-ring hetero-oligomer complex formed by different units that share about 50% identity, differs from GroEL because it is co-chaperone independent. However, thermosome, an archeal chaperonin, has a built-in lid that functions as a co-chaperone [51,52]. The main function of TCP is to assist the folding of multi-domain proteins in the eukaryote cytoplasm in a co-translational

way [53]. TCP can fold the N-terminus of a protein while its C-terminus is being built by the ribosome, which may explain the need for an attached “lid” to keep the cavity open to accommodate the growing polypeptide [50,54]. The mechanism of action of the TCP is not well understood but its structural shape and function are similar to GroEL, although there are important differences [8,50].

Hsp90 – Holding for folding

Hsp90 is essential for the growth of eukaryotic cells and is a specialized protein folding tool. Hsp90 can interact with proteins that are not in their native state, thereby keeping them in a refoldable state, and are important for maintaining the conformation, stability and function of a wide variety of proteins involved in signal transduction, the cell cycle and apoptosis [20-22,55]. The function of Hsp90 is driven by a dynamic association with its client proteins and their co-chaperones to form protein complexes. More than 40 Hsp90-client proteins are known and are classified in three groups: transcription factors, protein kinases and unrelated proteins [20,22]. Hsp90 does not bind nascent protein but receives its substrates from the Hsp70 and returns them to the Hsp70 system in a refoldable state (Fig. 1) [20-22].

Hsp90 consists of an N-terminal domain (25 kDa) that contains an ATP-binding site followed by a charged region (36 amino acids) of unknown function, a flexible domain in the middle (35 kDa) and a C-terminal domain (12 kDa) that is responsible for dimerization (Fig. 1) [56,57]. Both the N- and C-terminal domains have been implicated in substrate binding that is influenced by nucleotide binding [58,59]. Hsp90 proteins are assisted by several co-chaperones that, together with other chaperones and cellular proteins, form larger complexes known as foldosomes or cytoplasm chaperone complexes [7]. Among these proteins, Hsp90 proteins interact with Hsp70, Hsp40 homologues, foldases (prolyl and disulphide isomerases), and Hip (Hsp70 interacting protein). Several proteins that interact with Hsp90 contain the tetratricopeptide repeat motif that forms a helix-turn-helix structure [60], including Hop (Hsp70-Hsp90 organizing protein) that recognizes the C-terminus of Hsp90 and Hsp70 and acts as a mediator between them for substrate presentation [22]. The Hsp90 ATPase activity regulates the substrate binding cycle and Hop inhibits this activity by preventing nucleotide binding. The p23 co-chaperone recognizes Hsp90 bound to ATP and stimulates dissociation of the complex [22]. Hsp90 has become an interesting target for cancer therapies because it interacts with geldanamycin in tumour cells in a manner that is regulated by Hsp90 co-chaperones and the Hsp70 system [61,62]. These findings emphasize the importance of the cooperation between Hsp70 and Hsp90.

Hsp100 – Recovering aggregates

The proteins of the Hsp100/Clp family have a molecular mass ranging from 80 to 105 kDa, have ATPase activity, form ring oligomers and are involved in several biological phenomena, including stress tolerance, proteolysis, DNA transposition and gene regulation [63,64]. The Hsp100/Clp family is divided into two classes: Class I consists of subfamilies ClpA, B, C and D, and Class II contains subfamilies ClpM, N, X and Y [63]. Proteins belonging to Class I are hexameric in the presence of ATP and their structure is similar to a ring with two highly conserved ATP binding domains (GX2GXGKT) in each linear extremity [63,64]. Proteins belonging to Class II are smaller than Class I proteins and characteristically have only one ATP binding domain [63]. The ClpA and ClpB chaperones from *E. coli* have ATPase activity but different physico-chemical properties and functions [64]. ClpA and ClpX have intrinsic chaperone activity and participate in proteolytic complexes whereas ClpB functions alone [63]. The subfamily Hsp104/ClpB is involved in the recovery of proteins from aggregates and in thermal tolerance in *S. cerevisiae* [14,63,65]. Hsp100/ClpB binds to aggregates that have their hydrophobic sites exposed to the Hsp70 system, resulting in their disaggregation and refolding (Fig. 1) [15,66]. Mogk *et al.* [67] suggested the participation of smHsp in this system. The crystallographic structure of ClpB from *Thermus thermophilus* has recently been determined [68] and will help to elucidate how Hsp100 recovers proteins from aggregates.

SmHsp –Preventing aggregation

The small heat shock protein family (smHsp) contains proteins with a molecular mass of 15-40 kDa and forms large-spherical assemblies of 200-800 kDa (9 to 24-mers) [19,69,70]. SmHsps are involved in thermotolerance, are present in all cellular compartments, are ATP-independent, and are highly diverse (whereas two proteins have been described in bacteria, more than 20 have been described in plant cells [46]). The C-terminus (~90 amino acids) of smHsp is conserved and is also found in α -crystallin, an abundant constituent of the eye lens of most vertebrate species that is able to avoid protein aggregation [71,72]. Like other molecular chaperone families, smHsp proteins interact with their substrate through hydrophobic amino acids residues, thereby hiding the substrate surface and preventing its aggregation [73,74]. SmHsp can bind to thermally denatured proteins by trapping them in a non-native but refoldable state [73,75] that can be released after stress or transferred to the Hsp70 system for refolding (Fig. 1) [17-19,69,].

The smHsps have a high content of β -sheet conformation that is located mainly in the highly conserved C-terminal region and forms a β -sandwich structure [19,70,72,76]. SmHsps can dimerize via a highly stable interaction involving the β -sandwich region, and these dimers can oligomerize through

interactions between the N- and C-termini [19]. The oligomers have a dynamic subunit exchange that suggests a rapid, temperature-dependent equilibrium between the oligomer subunits [19]. An increase in temperature induces conformational changes in the smHsp oligomers that results in the exposure of a hydrophobic surface and in dissociation of the oligomers into their subunits [19], although the latter phenomenon is still controversial [77].

Perspectives

Molecular chaperones have one of the most important functions in intracellular protein metabolism because they assist protein folding. In particular, Hsp70 has a central role in assisting protein folding because it can receive and distribute protein substrates in cooperation with other molecular chaperone families. However, many questions still remain unanswered for these and other chaperones. The ability of Hsp100 to recover proteins from aggregates and the temperature-dependent mechanism of action of smHsps have yet to be understood in detail. Hsp90 is involved in numerous signal transduction pathways and its role, especially in cancer, requires clarification. Several other molecular chaperones and/or co-chaperones are likely to be discovered using Genomic and Proteomic approaches and should help to address intriguing questions such as how many intracellular pathways there are for avoiding protein aggregation and how this knowledge can be used to prevent or treat degenerative conditions like Alzheimer's disease. However, much work remains to be done in elucidating the structural and functional properties of these proteins in order to reach this stage.

Acknowledgments: The authors thank Prof. S. Hyslop for critical reading of this manuscript, and Fundação de Amparo à Pesquisa do Estado de São Paulo and Conselho Nacional de Pesquisa e Desenvolvimento for fellowships and financial support.

References

- [1] Laskey, R.A., Honda, B.M., Mills, A.D. and Finch, J.T. (1978) *Nature*, 275, 416-420.
- [2] Hendrick, J.P. and Hartl, F.U. (1993) *Annu. Rev. Biochem.*, 62, 349-384.
- [3] Lindquist, S. and Craig, E.A. (1988) *Annu. Rev. Genet.*, 22, 631-677.
- [4] Martin, J. and Hartl, F.U. (1997) *Curr. Opin. Struct. Biol.*, 7, 41-52.
- [5] Bukau, B. and Horwich, A.L. (1998) *Cell*, 92, 351-366.
- [6] Fink, A.L. (1999) *Physiol. Rev.*, 79, 425-449.
- [7] Miernyk, J.A. (1999) *Plant Physiol.*, 121, 695-703.
- [8] Hartl, F.U. and Hayer-Hartl, M. (2002) *Science*, 295, 1852-1858.
- [9] Anfisen, C.B. (1973) *Science*, 181, 223-230.
- [10] Mayer, M.P., Brehmer, D., Gässler, C.S. and Bukau, B. (2001) *Adv. Protein Chem.*, 59, 1-45.
- [11] Langer, T., Lu, C., Echols, H., Flanagan, J., Hayer, M.K. and Hartl, F.U. (1992) *Nature*, 356, 683-689.
- [12] Frydman, J., Nimmesgern, E., Ohtsuka, K. and Hartl, F.U. (1994) *Nature*, 370, 111-117.
- [13] Siegers, K., Waldmann, T., Leroux, M.R., Grein, K., Shevchenko, A., Schiebel, E. and Hartl, F.U. (1999) *EMBO J.*, 18, 75-84.

- [14] Glover, R. and Lindquist, S. (1998) *Cell*, 94, 73-82.
- [15] Goloubinoff, P., Mogk, A., Zvi, A.P., Tomoyasu, T. and Bukau, B. (1999) *Proc. Nat. Acad. Sci. U.S.A.*, 96, 13732-13737.
- [16] Mogk, A., Tomoyasu, T., Goloubinoff, P., Rüdiger, S., Röder, D., Langen, H. and Bukau, B. (1999) *EMBO J.*, 18, 6934-6949.
- [17] Veinger, L., Diamant, S., Buchner, J. and Goloubinoff, P. (1998) *J. Biol. Chem.*, 273, 11032-11037.
- [18] Lee, G.J. and Vierling, E. (2000) *Plant Physiol.*, 122, 189-197.
- [19] van Montfort, R., Slingsby, C. and Vierling, E. (2001) *Adv. Protein Chem.*, 59, 105-156.
- [20] Buchner, J. (1999) *Trends Biochem. Sci.*, 24, 136-141.
- [21] Young, J.C., Moarefi, I. and Hartl, F.U. (2001) *J. Cell Biol.*, 154, 267-273.
- [22] Pearl, L.H. and Prodromou, C. (2001) *Adv. Protein Chem.*, 59, 157-186.
- [23] Zhu, X., Zhao, X., Burkholder, W.F., Gragerov, A., Ogata, C.M., Gottesman, M.E. and Hendrickson, W.A. (1996) *Science*, 272, 1606-1614.
- [24] Rüdiger, S., Germeroth, L., Schneider-Mergener, J. and Bukau, B. (1997) *EMBO J.*, 16, 1501-1507.
- [25] Theysen, H.; Schuster, H-P.; Packschies, L.; Bukau, B. and Reinstein, J. (1996) *J. Mol. Biol.*, 263, 657-670.
- [26] Mayer, M.P., Schröder, H., Rüdiger, S., P aminoácidos I, K., Laufen, T. and Bukau, B. (2000) *Nature Struct. Biol.*, 7, 586-593.
- [27] Brehmer, D., Rüdiger, S., Gässler, C.S., Klostermeier, D., Packschies, L., Reinsten, J., Mayer, M.P. and Bukau, B. (2001) *Nat. Struct. Biol.*, 8, 427-432.
- [28] Szabo, A., Langer, T., Schroder, H., Flanagan, J., Bukau, B. and Hartl, F.U. (1994) *Proc. Natl. Acad. Sci. U.S.A.*, 91, 10345-10349.
- [29] Karzai, A.W. and MaMacken, R. (1996) *J. Biol. Chem.*, 271, 11236-11246.
- [30] Laufen, T., Mayer, M.P., Beisel, C., Klostermeier, D., Mogk, A., Reistein, J. and Bukau, B. (1999) *Proc. Natl. Acad. Sci. U.S.A.*, 96, 5452-5457.
- [31] Rüdiger, S., Scheneider-Mergener, J. and Bukau, B. (2001) *EMBO J.*, 20, 1042-1050.
- [32] Lee, S., Fan, C.Y., Younger, J.M., Ren, H. and Cyr, D.M. (2002) *J. Biol. Chem.*, 277, 21675-21682.
- [33] Li, J., Qian, X. and Sha, B. (2003) *Structure*, 11, 1475-1483.
- [34] Cheetham, M.E. and Caplan, A.J. (1998) *Cell Stress Chaperones*, 3, 28-36
- [35] Suh, W-C., Lu, C.Z. and Gross, C.A. (1999) *J. Biol. Chem.*, 274, 30534-30539.
- [36] Banecki, B., Liberek, K., Wall, D., Wawrzynow, A., Georgopoulos, C., Bertoli, E., Tanfani, F. and Zylicz, M. (1996) *J. Biol. Chem.*, 271, 14840-14848.
- [37] Martinez-Yamout, M., Legge, G.B., Zhang, O., Wright, P.E. and Dyson, H.J. (2000) *J. Mol. Biol.*, 300, 805-818.
- [38] Sha, B., Lee, S. and Cyr, D.M. (2000) *Structure*, 8, 799-807.
- [39] Liberek, K., Marszalek, J., Ang, D., Georgopoulos, C. and Zylicz, M. (1991) *Proc. Natl. Acad. Sci. U.S.A.*, 88, 2874-2878.
- [40] Borges, J.C., Hannes, F., Craievich, A.F., Hansen, L.D. and Ramos, C.H.I. (2003) *J. Biol. Chem.*, 278, 35337-35344.
- [41] Harrison, C.J., Hayer-Hartl, M., Di Liberto, M., Hartl, F.U. and Kuriyan J. (1997) *Science*, 276, 431-435.
- [42] Gelinas, A.D., Langsetmo, K., Toth, J., Bethoney, K.A, Stafford, W.F. and Harrison, C.J. (2002) *J. Mol. Biol.*, 323, 131-142.
- [43] Grimshaw, J.P., Jelesarov, I., Schönfeld, H.J. and Christen, P. (2001) *J. Biol. Chem.*, 276, 6098-6104.

- [44] Chesnokova, L.S., Slepenev, S.V., Protasevich, I.I., Sehorn, M.G., Brouillette, C.G. and Witt, S.N. (2003) *Biochemistry*, 42, 9028-9040.
- [45] Naylor, D.J., Stines, A.P., Hoogenr aminoácidos d, N.J. and Hoj. P.B. (1998) *J. Biol. Chem.*, 273, 21169-21177.
- [46] Borges, J.C., Peroto, M.C. and Ramos, C.H.I. (2001) *Genet. Mol. Biol.*, 24, 85-92.
- [47] Ellis, R.J. and Hartl, F.U. (1996) *FASEB J.*, 10, 20-26.
- [48] Rason, N.A., White, H.E. and Saibil, R.R. (1998) *Biochem. J.*, 333, 233-242.
- [49] Xu, Z., Horwich, A.L. and Sigler, P.B. (1997) *Nature*, 388, 741-750.
- [50] Kusmierczyk, A.R. and Martin, J. (2001) *FEBS Lett.*, 505, 343-347.
- [51] Klumpp, M., Baumeister, W. and Essen, L.O. (1997) *Cell*, 91, 263-270.
- [52] Ditzel, L., Lowe, J., Stock, D., Stetter, K.O., Huber, H., Huber, R. and Steinbacher, S. (1998) *Cell*, 93, 125-138.
- [53] Netzer, W.J. and Hartl, F.U. (1997) *Nature*, 388, 343-349.
- [54] Netzer, W.J. and Hartl, F.U. (1998) *Trends Biochem. Sci.*, 23, 68-73.
- [55] Freeman, B.C. and Morimoto, R.I. (1996) *EMBO J.*, 15, 2969-2979.
- [56] Stebbins, C.E., Russo, A.A., Schneider, C., Rosen, N., Hartl, F.U. and Pavletich, N.P. (1997) *Cell*, 89, 239-250.
- [57] Maruya, M., Sameshima, M., Nemoto, T. and Yahara, I. (1999) *J. Mol. Biol.*, 285, 903-907.
- [58] Young, J.C., Schneider, C. and Hartl, F.U. (1997) *FEBS Lett.*, 418, 139-143.
- [59] Scheibel, T., Siegmund, H.I., Jaenick, R., Ganz, P., Lilie, H. and Buchner, J. (1999) *Proc. Natl. Acad. Sci. U.S.A.*, 96, 1297-1302.
- [60] Scheufler, C., Brinker, A., Bourenkov, G., Pegorato, S., Moroder, L., Bartunik, H., Hartl, F.U. and Moarefi, I. (2000) *Cell*, 101, 199-210.
- [61] Maloney, A. and Workman, P. (2002) *Expert Opin. Biol. Ther.*, 2, 3-24.
- [62] Kamal, A., Thao, L., Sensintaffar, J., Zhang, L., Boehm, M.F., Fritz, L.C. and Burrows, F.J. (2003) *Nature*, 425, 407-410.
- [63] Schirmer, E.C., Glover, J.R., Singer, M.A. and Lindquist, S. (1996) *Trends Biochem. Sci.*, 21, 289-296.
- [64] Schirmer, E.C., Queitsh, C., Kowall, A.S., Parsell, D.A. and Lindquist, S. (1998) *J. Biol. Chem.*, 273, 15546-15552.
- [65] Sanchez, Y. and Lindquist, S.L. (1990) *Science*, 248, 1112-1115.
- [66] Schlee, S., Beinker, P., Akhrymuk, A. and Reinstein, J. (2004) *J. Mol. Biol.*, 336, 275-285.
- [67] Mogk, A., Deuerling, E., Vorderwulbecke, S., Vierling, E. and Bukau, B. (2003) *Mol. Microbiol.*, 50, 585-595.
- [68] Lee, S., Sowa, M.E., Watanabe, Y.H., Sigler, P.B., Chiu, W., Yoshida, M. and Tsai, F.T. (2003) *Cell*, 175, 229-240.
- [69] Waters, E.R., Lee, G.J. and Vierling, E. (1996) *J. Exp. Bot.*, 47, 325-338.
- [70] Kim, K.K., Kim, R. and Kim, S.H. (1998) *Nature*, 394, 595-599.
- [71] Plesofsky-Vig, N., Vig, J. and Brambl, R. (1992) *J. Mol. Evol.*, 35, 537-545.
- [72] de Jong, W.W., Caspers, G.J. and Leunissen, J.A. (1998) *Int. J. Biol. Macromol.*, 22, 151-162.
- [73] Lee, G.J., Roseman, A.M., Saibil, H.R. and Vierling, E. (1997) *EMBO J.*, 16, 659-671.
- [74] Haslbeck, M., Walke, S., Stromer, T., Ehrnsperger, M., White, H.E., Chen, S., Saibil, H.R. and Buchner, J. (1999) *EMBO J.*, 18, 6744-6751.
- [75] Ehrnsperger, M., Graber, S., Gaestel, M. and Buchner, J. (1997) *EMBO J.*, 16, 221-229.
- [76] van Montfort, R.L., Basha, E., Friedrich, K.L., Slingsby, C. and Vierling, E. (2001) *Nat. Struct. Biol.*, 8, 1025-1030.
- [77] Kim, R., Lai, L., Lee, H.H., Cheong, G.W., Kim, K.K., Wu, Z., Yokota, H., Marqusee, S. and Kim, S.H. (2003) *Proc. Natl. Acad. Sci. U.S.A.*, 100, 8151-8155.

8.3 Anexo II

Borges, J.C., Fischer, H., Craievich, A.F., & Ramos, C.H.I. (2004) Low-resolution structural study of two human Hsp40 chaperones in solution. DjA1 from subfamily A and DjB4 from subfamily B, have different quaternary structures. *J. Biol. Chem.*, (Submetido).

Low-resolution structural study of two human Hsp40 chaperones in solution. DjA1 from subfamily A and DjB4 from subfamily B, have different quaternary structures.

Júlio C. Borges^{1,2}, Hannes Fischer³, Aldo F. Craievich³ and Carlos H.I. Ramos^{1*}

¹Centro de Biologia Molecular Estrutural. Laboratório Nacional de Luz Síncrotron; PO Box 6192 Zip code: 13084-971 – Campinas SP, Brazil. ²Departamento de Bioquímica, Instituto de Biologia. Universidade Estadual de Campinas; Campinas SP, Brazil. ³Departamento de Física Aplicada, Instituto de Física, Universidade de São Paulo; São Paulo SP, Brazil.

*Corresponding author: Phone: 55-19-3287-4520/Fax: 55-19-3287-7110. E-mail: cramos@lnls.br

Running title: Structural insights on human Hsp40.

Summary

The proteins belonging to the Hsp40 family assist the Hsp70 in many cellular functions, helping to form an essential system for cell viability. Thus, the understanding of the structural and functional characteristics of the Hsp40 family is relevant to the general knowledge of the role of the Hsp70 chaperone system in the cell. We have investigated two human Hsp40 representatives, DjA1 (Hdj2/dj2/HSDJ/Rdj1) from subfamily A and DjB4 (Hlj1/DnaJW) from subfamily B, by small angle X-ray scattering and analytical ultracentrifugation methods, and determined their quaternary structure. We have also determined low-resolution models for the structure of a C-terminal deleted mutant of DjA1, DjA1₁₋₃₃₂, in which the dimer formation was interrupted. Our results were used in combination with the current structural information of Hsp40 C-terminal and J-domains, to generate models of the internal structural organization of DjA1 and DjB4. The characteristics of these models indicate that DjA1 and DjB4 were both dimers but with substantial differences in their quaternary structures: while DjA1 consisted of a compact dimer in which the N- and C-termini of the two monomers are facing each other, DjB4 formed a dimer where only the C-termini of the two monomers are in contact. The two proteins also differed in their capability to bind unfolded luciferase. Together, our results

indicate that the representatives of subfamilies A and B of human Hsp40 have different quaternary structures and chaperone functions.

Keywords: heat shock protein, molecular chaperone, Hsp40, small angle X-ray scattering, analytical ultracentrifugation, low-resolution structure.

The Hsp70¹ chaperone system is formed by Hsp70 (DnaK-related) and its co-chaperones Hsp40 (DnaJ-related) and GrpE. It assists many cellular processes involving proteins: including folding, transport through membranes, degradation and escape from aggregation (1-5). The affinity of Hsp70 for unfolded proteins is regulated by the binding of either ADP (high affinity) or ATP (low affinity) into its nucleotide-binding domain (NDB), and by interaction with its co-chaperones (5, 6). Hsp40 assists the folding of nascent proteins, aggregation prevention and refolding of aggregates by presenting nascent proteins to Hsp70 and stimulating its hydrolysis of ATP (7-11).

Hsp40 can also work as a chaperone by itself (12-14) and can be divided, as determined by genetic and mutational studies *in vivo* (15-18) and by biophysical methods (10, 19-21), in four conserved functional regions (Fig. 1A). The highly conserved α -helical N-terminal domain is called J-domain, is characteristic of proteins from this family and binds to Hsp70 stimulating its ATPase activity (10, 11, 13, 20, 22). Following the J-domain is found the glycine/phenylalanine rich region (G/F-rich; Fig. 1A) that is disordered and likely to be responsible for flexibility (10, 23). Some authors have shown that the G/F-rich region is essential for the function of Hsp40 (10, 15) since the J-domain alone is unable to bind to Hsp70 proteins (10, 19, 20). The central region is composed of a cysteine-rich domain (Cys-rich; Fig. 1A) that encloses four repeats of the motif CXXCXGXG (where X is any amino acid) and folds in a zinc-dependent fashion with two repeats bound to one zinc ion (19, 24, 25). The C-terminus domain (Fig. 1A) is formed by a β -sheet structure and is involved in the dimerization of Hsp40 (26). The Cys-rich and the C-terminal domains are involved in substrate binding and presentation (21, 24, 27).

¹Abbreviations: Hsp: heat shock protein; LUC, luciferase; CD: circular dichroism; SAXS: small angle X-ray scattering; AUC: analytical ultracentrifugation; SDS-PAGE: sodium dodecyl sulfate polyacrilamide gel electrophoresis; NMR: nuclear magnetic resonance; G/F-rich: glycine/phenylalanine rich region.

Hsp40 are ubiquitous proteins, they are found in all cellular locations and are highly diverse in the eukaryotic genomes (28, 29), with at least 44 genes present in the human genome (30). Based on their architecture and cellular location (4, 31, 32), Hsp40 proteins are classified as three main subfamilies named A to C (referred also as types I to III (4, 32, 33); Fig. 1A): subfamily A which consists of the four domains described above, subfamily B that lacks the Cys-rich domain, and subfamily C that has only the J-domain that is not necessarily located at the N-terminus (4, 32). Hsp40 proteins from subfamily A have autonomous chaperone activity and therefore may work in either an Hsp70-independent or dependent way. Otherwise, Hsp40 proteins from subfamily B do not present autonomous chaperone activity and work in an Hsp70-dependent way (17, 32, 34, 35).

Although high resolution structures of the C-terminal, Cys-rich, and J-domains are available, there is still a lack of information about how these domains interact with each other, with other domains and also with Hsp70. In order to obtain further information about relevant structural and functional characteristics of human Hsp40 proteins, we have decided to study DjA1 (also known as Hdj2/dj2/HSDJ/Rdj1), a representative of subfamily A, its C-terminal deletion mutant (DjA1₁₋₃₃₂), and DjB4 (also known as Hlj1/DnaJW, and 65% identical to DjB1/Hdj1), a representative of family B, by a gathered number of biophysical methods, such as small-angle X-ray scattering (SAXS) and analytical ultracentrifugation (AUC).

EXPERIMENTAL PROCEDURES

Cloning, expression and purification – DjA1 (also known as Hdj2, dj2, HSDJ, and Rdj1) was cloned from the human gene DNAJA1 (cDNA clone, GenBank accession number AW247277) and DjB4 (also known as Hlj1 and DnaJW) was cloned from the human gene DNAJB4 (cDNA clone, GenBank accession number AA081471). The Hsp40 nomenclatures DjA1 and DjB4 are as described by Ohtsuka and Hata (33). Two primers were used to amplify the DjA1 cDNA by PCR and to create the restriction enzyme sites for *Nde*I and *Xho*I: the DjA1 5'-primer (5' CCGGCAGGCTAGCATGGTGAAAGAAACAAC 3') containing an *Nde*I restriction site and a DjA1 3'-primer (5' TGAGTGTTATTCTCGAGTCATTAAGAGGTCTG 3') containing a *Xho*I restriction site. Two different primers were used to amplify the DjB4 cDNA by PCR and to create the restriction enzyme sites *Nde*I and *Bam*HI: the DjB4 5'-primer (5'

TCAAGGCATTCCATATGGGGAAAGACTATTA 3') containing an *NdeI* restriction site and the DjB4 3'-primer (5' GTGTAACAAAGTGGATCCTACTATGAGGCAGG 3') containing a *Bam*HI restriction site. The C-terminal deletion of DjA1 was constructed by site-directed mutagenesis using a primer that created a stop codon at the residue Phe333 followed by a restriction site for *XhoI* (5' TTATCAGGCTCGAGTTAGCCATTCTC 3').

The PCR products were cloned into the pET28a expression vector (Novagen) for His-tag purification methods. The correct cloning was confirmed by DNA sequencing using an ABI 377 Prism system (Perkin Elmer). These procedures created the vectors pET28aDjA1, pET28aDjB4 and pET28aDjA1₁₋₃₃₂, which were transformed into the *E. coli* strain BL21(DE3) for heterologous protein expression by adding 0.4 mmol/liter of isopropyl thio- β -D-galactoside (IPTG) at OD₆₀₀ = 0.6. The induced cells were grown for 5 hours and harvested by centrifugation for 10 minutes at 2,600 x g. The bacterial pellet was resuspended in lysis buffer (50 mmol/liter Tris-HCl, pH 8.0, 500 mmol/liter KCl and 10 mmol/liter EDTA – 15 mliter per liter of medium), in the presence of 5 U of DNase (GIBCO BRL) and 30 μ g/mliter of lysozyme (Sigma) for 30 min. The pellet was then disrupted by sonication in an ice bath, and centrifuged as described above. The supernatant was submitted to metal affinity chromatography in a HiTrap Chelating column (Pharmacia Biotech) using an ÄKTA FPLC (Pharmacia Biotech). The proteins were eluted with 500 mmol/liter imidazol and loaded into a HiLoad Superdex 200 pg 26/60 molecular exclusion column (Pharmacia Biotech) using an ÄKTA FPLC (Pharmacia Biotech). The degree of purification was estimated by SDS-PAGE (36) and the concentration was determined spectrophotometrically using the calculated extinction coefficient for denatured proteins (37, 38).

Analytical molecular exclusion chromatography – The experiments of analytical molecular exclusion chromatography were carried out with a Superose 12 HR 10/30 (Pharmacia Biotech) column coupled to an ÄKTA Purifier (Pharmacia Biotech) equilibrated with buffer Tris-HCl 25 mmol/liter buffer (pH 7.5), NaCl 500 mmol/liter, glycerol 1% containing β -mercaptoethanol (1 to 10 mmol/liter). The column was washed with 2 column volumes with a flow rate of 0.5 mL/min and aliquots of 100 μ liters of proteins were loaded onto the column with absorbance monitoring at 280 nm. The apparent molecular mass was calculated using a plot of the *ln* of the molecular mass (kDa) of standard proteins (thyroglobulin, 669 kDa; gamma-globulin, 160 kDa; bovine serum albumin, 69 kDa; chicken ovoalbumin, 45 kDa; and cytochrome c, 12 kDa) as a function of the time retention.

Circular dichroism spectroscopy – Circular dichroism (CD) measurements were performed using a Jasco J-810 spectropolarimeter with temperature controlled by Peltier Type Control System PFD 425S. Hsp40 proteins were resuspended in Tris-HCl 25 mmol/liter buffer (pH 7.5), NaCl 500 mmol/liter and β -mercaptoethanol 1 mmol/liter at final concentrations of 20 to 50 μ mol/liter. Data were collected at a scanning rate of 50 nm/min with spectral band width of 1 nm using a 0.1 mm path length cell. CDNN Deconvolution software (Version 2 – <http://Bioinformatik.biochemtech.uni-halle.dee/cdnn>) was employed for secondary structure prediction. All buffers used were of chemical grade and were filtered before their use to avoid scattering from small particles.

Chaperone activity measurement – The method described by Lu and Cyr (27, 34) was used to verify the ability of the Hsp40 proteins to interact with unfolded proteins. Briefly, luciferase (Promega) was chemically denatured by dilution into guanidinium-HCl 6 mol/liter, for 40 min at room temperature and then diluted 25 x into 50 μ liter of Tris-HCl 25 mmol/liter buffer (pH 7.5), NaCl 500 mmol/liter and β -mercaptoethanol 5 mmol/liter, in the presence and in the absence of His-tagged Hsp40 proteins (see figure captions for concentration details). Otherwise, luciferase (Promega) at 2 μ mol/liter was thermically denatured by incubation at 42 °C for 10 min, in the presence and in the absence of His-tagged Hsp40 proteins (see figure captions for concentration details), and then cooled back down to room temperature. The solutions were centrifuged for 10 min at 21,000 x g and 50 μ liter of the supernatant was incubated with 50 μ liter of a solution of 50% of Talon metal chelate (Clontech) in Tris-HCl 25 mmol/liter buffer (pH 7.5), NaCl 500 mmol/liter and β -mercaptoethanol 5 mmol/liter for 1 hour at room temperature. The mixture was then centrifuged for 1 min at 10,000 x g and 4 °C, and the pellet, in which the resin was contained, was washed twice with 75 μ liter of Tris-HCl 25 mmol/liter buffer (pH 7.5), NaCl 500 mmol/liter and β -mercaptoethanol 5 mmol/liter containing 15 mmol/liter of imidazol. The complex proteins bound to the resin were eluted with the same buffer described above but with 150 mmol/liter imidazol, concentrated by precipitation with acetone 80% (Merck), and visualized by SDS-PAGE (36).

AUC – Sedimentation velocity and sedimentation equilibrium experiments were performed with a Beckman Optima XL-A analytical ultracentrifuge. The proteins DjA1, DjA1₁₋₃₃₂, and DjB4 were tested in concentrations from 50 to 1000 μ g/mliter in 25 mmol/liter Tris-HCl buffer (pH 7.5), 500 mmol/liter NaCl, glycerol 1% (not for DjA1₁₋₃₃₂) and 1 mmol/liter β -

mercaptoethanol, with no apparent aggregation. The sedimentation velocity experiments were carried out at 20 °C, 25,000 rpm for DjA1 and DjB4, and 30,000 and 40,000 rpm for DjA1₁₋₃₃₂ (AN-60Ti rotor), and the scan data acquisition at 230 and 238 nm for low and high protein concentration, respectively. The sedimentation equilibrium experiments were made at 20 °C at speeds of 6,000, 8,000, and 10,000 rpm with the AN-60Ti rotor and scan data acquisition at 238 nm. The analysis involved the fitting of a model of absorbance versus cell radius data using nonlinear regression. All fittings were performed using the Origin software package (Microcal Software) supplied with the instrument. The van Holde-Weischet (39) (sediment coefficient plot) and the Sedimentation Time Derivative ($g(s^*)$ integral distribution) (40) methods were applied to analyze the sedimentation velocity experiments. The methods used for analyzing both velocity and equilibrium experiments allow the calculation of the apparent sedimentation coefficient s , the diffusion coefficient D , and the molecular mass M . The ratio of the sedimentation to diffusion coefficient gives the molecular mass:

$$M = \frac{sRT}{D(1 - \bar{V} \rho)} \quad \text{Equation 1}$$

where R is the gas constant and T is the absolute temperature. The Sednterp software (www.jphilo.mailway.com/download.htm) was used to estimate each protein partial specific volume at 20 °C ($\bar{V} \text{DjA1} = 0.7275$ mliter/g, $\bar{V} \text{DjA1}_{1-332} = 0.7309$ mliter/g and $\bar{V} \text{DjB4} = 0.7302$ mliter /g), the buffer density ($\rho = 1.02163$ g/mliter) and the buffer viscosity ($\eta = 1.0851 \times 10^{-2}$ poise).

The Self-Association method was used to analyze the sedimentation equilibrium experiments using several models of association for DjA1 and DjB4 to fit the data. The distribution of the protein along the cell, obtained in the equilibrium sedimentation experiments, was fitted with the following equation (41):

$$C = C_0 \exp \left[\frac{M(1 - \bar{V} \rho) \omega^2 (r^2 - r_0^2)}{2RT} \right] \quad \text{Equation 2}$$

where C is the protein concentration at radial position r , C_0 is the protein concentration at radial position r_0 , and ω is the centrifugal angular velocity. The Sednterp software was applied to estimate the standard sedimentation coefficients ($s_{20,w}$) at each protein concentration to calculate $s_{20,w}^0$ at 0 mg/mliter protein concentration by extrapolation. This procedure minimizes interferences caused by temperature, viscosity solution and molecular crowd (42).

SAXS – SAXS experiments were performed at the SAS beamline of the LNLS synchrotron radiation facility, Campinas, Brazil (43) under the same conditions as those described by Borges et al. (44). Measurements were carried out using a monochromatic X-ray beam with a wavelength $\lambda = 1.488 \text{ \AA}$. For a sample-to-detector distance of 840 mm selected for the experiments, the modulus of the photon momentum transfer ($q = 4\pi\sin\theta/\lambda$, 2θ being the scattering angle) covered a range starting from $q = 0.01 \text{ \AA}^{-1}$ up to $q = 0.44 \text{ \AA}^{-1}$. The scattering intensity curves, $I(q)$, were recorded using a gas 1 dimensional position sensitive X-ray detector. In order to monitor eventual effects from radiation damage and synchrotron beam instabilities, SAXS curves were determined in many short frames (90 seconds each). The obtained SAXS data were normalized to account for the natural decay in intensity of the synchrotron incident beam and corrected for inhomogeneous detector response. Finally, the scattering intensity produced by the buffer was subtracted and the difference curves were scaled to equivalent protein concentration. Three types of samples with different concentrations were studied by SAXS: 1) 3.3 mg/mliter of DjA1 in Tris-HCl 25 mmol/liter (pH 7.5), NaCl 500 mmol/liter, glycerol 1% and dithiothreitol 1 mmol/liter; 2) 2.9 and 6.5 mg/mliter of DjA1₁₋₃₃₂ in Tris-HCl 25 mmol/liter (pH 7.5), NaCl 500 mmol/liter, and β -mercaptoethanol 1 mmol/liter; 3) 2.5, 4.4, 6.2 and 8.3 mg/mliter of DjB4 in Tris-HCl 25 mmol/liter (pH 7.5), NaCl 500 mmol/liter and β -mercaptoethanol 1 mmol/liter.

Reliable structural information about the low resolution structure of proteins in solution can be derived from SAXS data provided all proteins are in the same (monomeric or oligomeric) state and the solution is “diluted”, *i.e.*, interference effects in scattering amplitudes produced by different proteins are negligible. Under these assumptions, the total scattering amplitude is proportional to the form factor of an isolated (monomeric or oligomeric) protein, averaged for all orientations. When high concentration samples were available, a data merging was performed using low and high concentration samples for the high and low q ranges of the scattering intensity, respectively. This procedure reduced the overall statistical errors in the scattering curves without introducing unwanted interferences or spatial correlation effects in the small q range.

In order to establish the molecular masses of DjA1, DjA1₁₋₃₃₂ and DjB4, the SAXS intensity produced by a standard bovine serum albumin (69 kDa), in solutions of 5 and 10 mg/mliter, were also determined. For dilute solutions, the normalized intensity extrapolated to $q = 0$, $I(0)$, is proportional to the molecular mass, so as the molecular mass M of the proteins

investigated here were determined from the ratios between the $I(0)$ values corresponding to the different samples and that of the standard bovine serum albumin.

Computer programs – By applying the indirect Fourier transform program named GNOM (45) to the normalized SAXS curves, the distance distribution function $p(r)$ and the radius of gyration R_g of the studied proteins were evaluated (46). Prior to GNOM analysis, to ensure that the intensity at higher angle decays as q^{-4} , following Porod's law for a two-electron density model (46), a constant intensity background was subtracted from the experimental SAXS data. The low-resolution shapes (or molecular envelopes) of the proteins were restored from the experimental SAXS curves using an *ab initio* method named DAMMIN (47) and recently described in Borges et al. (44). Briefly, in this method the low-resolution shape of the protein is simulated by a set of small spheres filling, at the beginning, another sphere with a diameter equal to the maximum diameter D_{max} of the protein (previously determined by using GNOM). DAMMIN yields the structure model containing a fraction of the initial number of dummy atoms whose associated scattering intensity gives the best fit to the experimental SAXS data. The configuration of the different protein domains can still be refined by manually rotating or displacing them until the discrepancy between calculated and experimental SAXS curves is minimized. The protein models were displayed using the program WebLab ViewerLite software (www.accelrys.com).

The HydroPro software (48) was applied to estimate the translational diffusion coefficient D_t , the radii of gyration R_g , the sedimentation coefficient s , and the maxima distance (D_{max}) starting from the *ab initio* models generated by DAMMIN using SAXS data at 20 °C. The HydroPro software was setup with the radius of the atomic elements from the *ab initio* development, with sigma factors from 5 to 8 (as indicated by supplier) and minibeads radius (SIGMIN and SIGMAX) from 6 to 2 Å after initial evaluation of the two extremes. The parameters $Vbar$, ρ , and η were estimated using the software Sednterp as described above. The translational friction ratio or Perrin factor P , that indicates the relation of the frictional coefficient of the Hsp40 particles to a sphere of the same molecular mass (f/f_0), was estimated by Solpro software (49).

RESULTS AND DISCUSSION

DjA1 and DjB4 were purified as folded dimers and DjA1₁₋₃₃₂ was purified as a folded monomer – The correct cloning and sequencing of the DjA1, DjB4, and DjA1₁₋₃₃₂ were confirmed by DNA sequencing. All proteins were expressed in large quantities and were more than 95% pure as confirmed by SDS-PAGE (Fig. 2A). The proteins were unstable at low ionic strength but were soluble in the presence of 500 mmol/liter NaCl. DjA1 and DjB4 were purified as dimers and the mutant DjA1₁₋₃₃₂ was purified as a monomer, as confirmed by analytical molecular exclusion chromatography (Fig. 2B). The folding state of the proteins was investigated by CD and the resultant spectra corresponded to folded proteins with minima at 208 nm (Fig. 2C). The analysis using the CDNN Deconvolution software indicated that the proteins have similar amount of secondary structure: 35% of β -sheet structure and 10% of α -helices. The high content of β -sheet structure agrees with the data originated from X-ray crystallography and nuclear magnetic resonance experiments, which indicate the presence of β -sheet structures in the C-terminus (21, 26) and in the Cys-rich domain (21, 25) of Hsp40. The deletion of the last 54 residues of DjA1, although interrupted its dimerization, did not have a strong effect in its global conformation (Fig. 2B). However, the thermal stability of the deleted mutant was lower than the whole protein. DjA1 and DjB4 were heated up to 50 °C whereas the amount of secondary structure of DjA1₁₋₃₃₂ decreased when the protein was heated above to 38 °C (data not shown). These results indicate that the dimeric state is important for the stability of the protein at temperatures well above the body temperature, a characteristic relevant for proper function in the cell.

DjA1 was more efficient in binding unfolded luciferase than DjB4 – The chaperone action of Hsp40 occurs by hydrophobic interaction with unfolded or partially folded proteins (14, 17, 18, 21), thus the ability of the Hsp40 proteins to bind unfolded luciferase was investigated. The His-tagged chaperones have high affinity for the Talon metal chelate resin and bind to it in the absence of imidazol. And then, any unfolded luciferase that binds to a His-tagged chaperone co-purifies with it when the resin is washed with 150 mmol/liter imidazol. Fig. 3 shows SDS-PAGE profiles in which washed samples of chemical (Fig. 3A) or thermal (Fig. 3B) denaturated luciferase, in the presence or in the absence of increasing concentrations of DjA1, DjA1₁₋₃₃₂ or DjB4 (see figure caption for details), were loaded. Folded luciferase, at 22

°C, did not bind to DjA1 or DjB4, and unfolded luciferase, at 42 °C, bound to DjA1 but not to DjB4 (Fig. 3B). Similar results were obtained with chemically unfolded luciferase (Fig. 3A), in which this protein bounded to DjA1 but had a very low affinity for DjB4. The deletion of the C-terminus, mutant DjA1₁₋₃₃₂, had a major effect in the efficiency of binding (Fig. 3A).

These results indicated that DjA1 and DjB4 have no equivalent chaperone activity. A similar result was obtained when the ability to bind unfolded luciferase of the Hsp40 yeast proteins was measured: Ydj1 (subfamily A) was shown to be 2.5 times more active than Sis1 (subfamily B) (18, 27). Other authors showed that DjA1 is more efficient to refold luciferase than DjB1 (subfamily B, homologous to DjB4) (50). DjA1₁₋₃₃₂ has lower chaperone activity than intact DjA1 and was not used in the luciferase thermal unfolding experiments because it unfolds at temperatures above 38 °C. The lower chaperone activity of DjA1₁₋₃₃₂ indicated that the quaternary structure of Hsp40 is connected with its chaperone activity, as also shown for yeast Hsp40 Sis1, in which a C-terminal deletion abolishes the ability to assist Hsp70 (26).

Envelope models derived from SAXS – The experimental (corrected and normalized) SAXS curves from DjA1, DjA1₁₋₃₃₂, and DjB4 are displayed in Fig. 4A. The molecular masses of the studied proteins were obtained from the quotient between their intensity extrapolated to $q = 0$ and the $I(0)$ value corresponding to bovine serum albumin, as described before. The molecular masses of DjA1 and DjB4 were determined as 95 kDa and 90 kDa, respectively, in good agreement, inside the error, with the values calculated from their amino acid sequence (Table I) assuming that they are dimers. On the other hand, the SAXS data from DjA1₁₋₃₃₂ indicated the existence of a monodisperse set of molecules with a molecular mass of 40 kDa, in good agreement with the value calculated from the amino acid sequence (Table I) of the monomer. The second step of data analysis was to apply the GNOM program in order to determine the distance distribution function and related structure parameters, such as the radius of gyration R_g and the maximum diameter D_{max} of the proteins. The distance distribution functions $p(r)$ of the three studied proteins are presented in Fig. 4B. Since these functions do not exhibit any negative range, we concluded that all proteins were in “dilute” state as required for the further analysis of the molecular shape. The values of D_{max} and R_g yielded by GNOM are reported in Table II.

The DjA1 and its C-terminal deletion mutant, DjA1₁₋₃₃₂, have similar maximum diameters $D_{max} = 140$ Å (Fig. 4B), thus implying that they had a similar asymmetry. The DjB4 chaperone had a D_{max} of 200 Å (Fig. 4B), although it has a molecular mass smaller than DjA1

(Table I), which suggested a more elongated conformation than DjA1. According to the overall trend of the $p(r)$ functions (Fig. 4B), we concluded that DjB4 had a rather elongated shape (51), and that the shapes of DjA1 and DjB4 were clearly different. As a matter of fact, while the DjA1 had a $p(r)$ function which only exhibited a smooth and well-defined peak, the $p(r)$ function corresponding to DjB4 had two clear shoulders, one above and another below the main peak (Fig. 4B). This complex profile of the $p(r)$ function suggested that DjB4 was composed of well separated subunits.

The third step of data analysis was to apply the DAMMIN program to the SAXS data, which yields the low resolution molecular shape of the proteins. The dummy atom model (DAM) of the structure was derived from the experimental data assuming a 2-fold symmetry for the dimers. This symmetry restriction significantly reduces the number of free parameters of the models. The starting volume for the DjA1 dimer corresponded to a filled sphere with a diameter $D_{\max} = 140 \text{ \AA}$ (yielded by GNOM) containing 4,196 dummy atoms with a radius $r_a = 4.5 \text{ \AA}$. As a result of the application of the DAMMIN program, 600 ± 10 dummy atoms were assigned to the resulting final model of the DjA1 dimer. On the other hand, the starting volume for DjA1₁₋₃₃₂ was filled with DAM = 5,594 dummy atoms with a packing radius $r_a = 3.7 \text{ \AA}$ within a sphere with the diameter $D_{\max} = 140 \text{ \AA}$. Of those, 395 ± 5 dummy atoms were assigned by DAMMIN to the final model of the DjA1₁₋₃₃₂ monomer. The starting volume for DjB4 was filled with DAM = 6,699 dummy atoms with a packing radius $r_a = 4.8 \text{ \AA}$ within a sphere with the diameter $D_{\max} = 200 \text{ \AA}$. Of those, 280 ± 10 dummy atoms remained in the final model of the DjB4 dimer. Twenty independent *ab initio* DAMMIN simulations were performed for each protein. In order to verify the uniqueness of the resulting shapes, the restorations were performed with different starting conditions. All of them yielded similar results. The obtained DAM models were superimposed and averaged by the DAMAVER package software (52) and are presented in Fig. 5. Relevant parameters associated to DAMMIN application and the values of $D_{\max}$ and R_g derived from it are reported in Table II. We can notice that these values agreed well with the same parameters determined by GNOM for the three studied molecules. DAM derived hydrodynamic parameters (estimated by HydroPro) are shown in Table III. Maximum concentrations of the protein used for SAXS solution studies of the DjA1 and DjB4 dimers allowed $30 \text{ \AA}$ as the maximum resolution of the final models (Table II), even when the data was measured up to $15 \text{ \AA}$. This resolution did not permit an unambiguous determination of the spatial positions of the secondary structure

elements of DjA1, but allowed us to determine the overall shape of the molecule and the relative position of their individual domains. Table II reports different resolutions for GNOM and DAMMIN calculations. For instance, for DjA1 protein, the whole SAXS curve determined up to $q_{\max} \sim 0.4 \text{ \AA}^{-1}$ was used for GNOM, thus the resolution $2\pi/q_{\max}$ were equal to $\sim 15 \text{ \AA}$. Within the high q range from $q = 0.2 \text{ \AA}^{-1}$ up to $0.4 \text{ \AA}^{-1}$ (Fig. 4A, curve 2), the statistical errors are very high and therefore no significant structure information comes from this high q range. In order to speed up the rather long calculations of DAMMIN program without significantly losing structural information, we have used SAXS data only up to $q_{\max} = 0.2 \text{ \AA}^{-1}$ for this analysis, yielding a resolution equal to $30 \text{ \AA}$. Inherent limitations of shape determination methods such as DAMMIN avoid higher resolution protein envelopes from statistically good SAXS data taken up to higher $q_{\max}$ values, because these methods assume a uniform electron density, thus limiting the resolution to 20-30 $\text{\AA}$.

GNOM analysis yielded the same $D_{\max}$ for both DjA1₁₋₃₃₂ and DjA1, despite the dimer arrangement of the latter. This was a clear indication of the highly asymmetric shape of the monomer DjA1₁₋₃₃₂, a characteristic also noticed from the observed profile of the $p(r)$ function. Fig. 5A shows the envelope models for DjA1 (left) and for DjA1₁₋₃₃₂ (two monomers, center). The two monomers of DjA1₁₋₃₃₂ were positioned facing each other to illustrate how a dimer may be formed from two monomers and the monomer envelope placed into the dimer envelope fits very well in one half of it (Fig. 5A – right). Also, the hydrodynamic properties of the two combined monomers of DjA1₁₋₃₃₂, calculated by the HydroPro software, were similar to those determined for the DjA1 dimer model (data not shown). These results confirmed that the DAM model determined for DjA1 was a good description of its low resolution structure. Two perpendicular views of the envelope model of DjB4 are shown in Fig. 5D. The comparison of the shapes of the DjA1 and DjB4 envelopes strongly indicates that these proteins clearly have different quaternary structure.

The hydrodynamic parameters calculated by AUC corroborated the models generated from the SAXS data – The hydrodynamic properties of the DjA1 and DjB4 chaperones were established by both sedimentation equilibrium (Fig. 6A-B) and sedimentation velocity (Fig. 7A). In these experiments, the proteins behaved as a monodisperse system and did not exhibit any aggregation. The sedimentation equilibrium experimental profiles of DjA1 (Fig. 6A) and DjB4 (Fig. 6B) were well fitted with a single exponential as indicated by the analysis of their residuals. The molecular masses M of DjA1 and DjB4 derived from the sedimentation

equilibrium measurements were 93.0 ± 1.0 and 80.0 ± 1.0 kDa, respectively (Table I). For DjA1₁₋₃₃₂, the M was 42.0 ± 2.0 kDa, which was estimated from Equation 1 using the sedimentation coefficient ($s_{20,w}^0$) calculated from the sedimentation velocity analysis (see below) and the diffusion coefficient ($D_{20,w}^0$) calculated from dynamic light scattering experiments (data not shown). The molecular masses determined for the human Hsp40 proteins are in good agreement, inside the error, with the molecular masses calculated from their amino acid sequence and from SAXS results (Table I). For the determination of the sedimentation coefficient s of the studied proteins, the sedimentation velocity experiments were performed at different concentrations and velocities (Fig. 7). The obtained standard sedimentation coefficients $s_{20,w}$ were plotted as functions of protein concentration to calculate the standard sedimentation coefficient at 0 mg/mL, $s_{20,w}^0$, by extrapolation, in order to minimize interferences caused by temperature, viscosity solution and molecular crowd (42). For DjA1 we obtained $s_{20,w}^0 = 4.63$ S, and for DjB4 $s_{20,w}^0 = 3.78$ S (Fig. 7B). The sedimentation coefficient of DjA1₁₋₃₃₂ was found to be velocity-dependent, with $s_{20,w}^0 = 2.85$ S at 30,000 rpm, and of 3.40 S at 40,000, both results being larger than that predicted for the DjA1₁₋₃₃₂ DAM model ($s_{20,w}^0 = 2.46$ S; see Table III). This effect may be caused by aggregation or, more likely in this case, by an effect of high asymmetric conformation (53). Some subdomains of the Hsp40 proteins might become more flexible in the monomer form, which reduces the frictional ratio when compared with a rigid body and induces a distortion in the measured s values (54). Since the sedimentation velocity experiment measures a time average s of the solution particles, it is not possible to distinguish a molecule with a high frictional ratio from another with flexible segments (55). In addition, the Perrin factor and the elongated shape of DjA1₁₋₃₃₂ determined from SAXS data and its domain structure suggest that this protein is flexible, reinforcing the hypothesis that the apparent high $s_{20,w}^0$ is due to this molecular feature.

The hydrodynamic parameters determined from AUC experiments for DjA1, DjA1₁₋₃₃₂ and DjB4 were similar to those derived from the DAM models of DjA1 and DjB4 by the HydroPro software (Table III), which confirms the reliability of the structure models obtained from SAXS results. Together, the data generated by SAXS and AUC data indicated that DjA1 and DjB4 were dimers with different structural features, and that DjA1₁₋₃₃₂ was a monomer with flexible segments (Tables I and III). The structural parameters obtained from AUC measurements strongly suggest that the DAM models derived from SAXS data represent a good low resolution structure of the real conformation of the studied proteins.

The results described and discussed in this and in the preceding sections indicated that DjA1, DjA1₁₋₃₃₂, and DjB4 were produced as folded proteins with conformational properties of members from the Hsp40 chaperone family. Then, the next step was to use the low-resolution models of the studied proteins determined from SAXS data and the known information about the high-resolution structures of Hsp40 domain available in the Protein Data Bank to obtain a more detailed insight into the structural conformation of both DjA1 (dimer and monomer) and DjB4.

Quaternary structure of DjA1 – The DjA1₁₋₃₃₂ mutant was a monomer with an elongated shape, as determined by analytical molecular exclusion chromatography, AUC and SAXS experiments. The NMR structure of the J-domain (56) displayed in red in Fig. 5B, and the crystallographic structure of the Cys-rich and C-terminal domains (21), showed in green in Fig. 5B, were fitted into the DjA1₁₋₃₃₂ envelope. The final arrangement is displayed in Fig. 5B, where some important features are noticed, namely the C-terminus being located inside the thinnest part of the protein envelope, the J-domain being located inside the thickest part of the envelope (Fig. 5B), and some portions of the domains in this model being highly flexible (arrows in Fig. 5B) due to the presence of the G/F-rich region (see below). As argued before, this may explain the fact that the hydrodynamic properties predicted by the HydroPro software for the DjA1₁₋₃₃₂ envelope were not in perfect agreement with the experimental AUC data. The apparent high sedimentation coefficient of DjA1₁₋₃₃₂ may be justified by the presence of the G/F-rich region, which is believed to work as a flexible spacer between the J-domain and the remaining structure of the protein (10, 23). In the Cys-rich domain determined by crystallography studies (21), the sub-domains I and II (Fig. 5B) form a L-shaped structure (90° angle), whereas the same domain determined by NMR, forms an angle of about 45° resulting in a V-shaped structure (25). This indicates that the subdomain II may have some flexibility, which yields the apparent high sedimentation coefficient of DjA1₁₋₃₃₂ calculated by AUC experiments. Also, the Hsp40 proteins studied here exhibited high degradation susceptibility (data not shown) and the high flexibility of the G/F-rich region may be an explanation for this behavior.

The results of SAXS experiments indicated that both the DjA1 and the DjA1₁₋₃₃₂ have an elongated shape and similar D_{max} , showing that they probably share the same axial orientation. The superposition of the two models (Fig. 5A, right) illustrated that a dimer composed of two DjA1₁₋₃₃₂ monomers (Fig. 5A, center) had nearly the same structure as the

DjA1 dimer. This conclusion indicated that the region involved in the dimer formation is mainly located at the C-terminus of DjA1. The deletion of this region blocked the formation of the dimer and decreased the protein stability, but had no effect in the secondary and tertiary (low-resolution) structures of the DjA1₁₋₃₃₂. Considering the observed orientation of the monomers in the DjA1 dimer model, we have fitted the high-resolution structures of J-, Cys-rich, and C-terminal domains into the monomer envelope (Fig. 5B), and then built a model of the Hsp40 domains arrangement into the DjA1 dimer (Fig. 5C). The DAM model generated from SAXS data indicated that DjA1 was a dimer with a bent horseshoe shape, as the structure determined for the C-terminal and Cys-rich domains of Ydj1 (subfamily A) (21). Fig. 5C shows that the DjA1 protein was bullet-shaped and the monomers C-termini were located at the thinnest side of the model and were responsible for the dimerization. The monomers N-termini (J-domains) were located in the thickest part of the model where they can act as a clamp for binding to the Hsp70 (Fig. 5C). The C-terminal and Cys-rich domains of Ydj1 (green in Fig. 5C, 1NLT; ref. 21) were placed on the axial axis in which they were well fitted. The L-shape of this complex was a consequence of the presence of the Cys-rich domain, which was placed on the center to face its counterpart in the other monomer (Fig. 5B).

Quaternary structure of DjB4 – The data generated by SAXS was used to create an envelope model of the quaternary structure of the human DjB4 that agreed with the measured structural parameters (Fig. 5D). The envelope model of DjB4 was compatible with an elongated structure, with a maximum diameter $D_{\max}$ larger than the one for DjA1 (200 Å *versus* 140 Å). This was an unexpected finding because DjA1 has a molecular mass higher than DjB4 as shown by AUC experiments (Table I). However, the similarities between the predicted and the calculated hydrodynamics parameters of the envelope model were a strong indication of the accuracy of the model (Table III). The DjB4 *ab initio* model generated from SAXS data by using DAMMIN and DAMAVER is displayed in Fig. 5D. The high-resolution structures of the Hsp40 domains available: the NMR structure of the J-domain (56), which is shown in red in Fig. 5E, and the crystallographic structure of the C-terminus of Sis1 (subfamily B) (21), which is shown in green in Fig. 5E; were fitted in the best possible position into the DjB4 envelope. Studies on C-terminus deletions of Hsp40 (21, 26, this work) indicate that this region is responsible for dimerization, which implies that the monomers C-termini must be located near to each other. The localization of the two C-terminal domains of DjB4 in the center of the model as shown in Fig. 5E fulfilled the requirement stated above. Any other arrangement for

these domains would difficult the formation of the dimer, which is against of the strong tendency to form dimers observed for Hsp40 proteins (21, 26). Each J-domain was then placed at the extremities in the model (Fig. 5E) because they are connected to the C-terminus by a long and flexible region formed by the G/F-rich region (residues 68-139) plus a sequence of 60 amino acids with unknown function (residues 140-199). High resolution structure of this region is not available, making therefore difficult to predict the correct disposition of these regions (residues 68-199) in the model, and this is represented by a dotted line (Fig. 5E).

DjA1 and DjB4: similarities and differences – Several genetic and biochemical studies have pointed out that the Hsp40 family is functionally diverse and the different domain composition among them point that they are also structurally diversified (32). Such diversity means that the interaction of Hsp70 with different Hsp40 generates specialized combinations that facilitate specific actions of the Hsp70 chaperone machinery in the cell (18, 32, 34). It is reasonable that Hsp40 from subfamilies A and B have evolved to have differences in chaperone activity, however, the reason why they exhibit such differences are not completely understood. The determination of the tertiary structure of some of the domains of Hsp40 proteins, from subfamilies A and B (21, 25, 26, 56), are helping to understand the difference in their chaperone activity. The data on the quaternary structure of these proteins presented here add important contributions to the task of explaining the differences (exhibit by subfamilies A and B Hsp40s) in binding to substrates and to Hsp70 proteins.

The presence of a J-domain defines a protein as a member of the Hsp40 family. Since the conserved residues in this domain are responsible for the stabilization of the anti-parallel coiled coil (56), we suggest that this domain has the same conformation on both DjA1 and DjB4 proteins. However, the J-domains are located differently in their quaternary structure (because the dimerization sites, at the C-terminus, must be in the center of the model): the J-domains from each monomer of DjA1 are displaced facing each other, whereas, in the DjB4 model, these domains are placed in the extremities. There is no high-resolution structure of the G/F-rich region available, an important site for the Hsp40 binding to Hsp70 (19, 20), therefore we represented it as a dotted line in our models (Fig. 5B-C and 5E). According to the primary sequence homology, the G/F-rich region is bigger in the DjB4 than in DjA1 (Fig. 1B), and in combination with a region of about 60 residues which also has no available structure, allowed the spreader conformation presented for DJB4 (Fig. 5E), in which the J-domains are localized in the extremities. The C-terminal regions of both subfamilies A and B are

responsible for substrate binding and are structurally similar (21), but present differences in their ability to refold luciferase (27, 34).

Lu & Cyr (34) studied the interaction of Ydj1 (A) and Sis1 (B) with two different Hsp70s and concluded that the two Hsp40 proteins differ in their chaperone functions, a characteristic that may be caused by the presence of the zinc-finger region in the first, which may play a structural role in stabilizing the polypeptide binding groove of Hsp40. Fan et al. (18) engineered two Hsp40 chimeras in which the central part of subfamily A Ydj1 protein (residues 101-255) was exchanged by the central part of subfamily B Sis1 protein (residues 108-257) and *vice versa*. They neatly found that the exchange switch the specificity of the proteins in binding substrates and in stimulating the luciferase refolding activity by Hsp70. These results suggest that selectively substrate-binding sites are needed for substrate selection in Hsp40 and that the central portion of this protein is important for chaperone activity specification (18). It is worthy to note that the central part of the DjA1 and DjB4 Hsp40s seem to be the main factor responsible for the differences in the low resolution models showed here: the central part of DjA1 (G/F-rich region plus Cys-rich region – residues 68-207) seems to keep the J-domains close to each other whereas the central part of DjB4 (residues 68-199) keeps the J-domains far apart from each other. A different global structure conformation rather than specific amino acid interactions may be the responsible for the difference in substrate specificity of the chaperones as suggested by the findings of Lopez et al. (57) who compared the Sis1/Hdj1 (subfamily B) and Ydj1/Hdj2 (subfamily A) Hsp40s ability to maintain the Rnq1 prion. Thus, the differences on the DjA1 and DjB4 quaternary structure above mentioned may be a factor affecting the differences in chaperone activity. However, the answer that justifies why DjA1 and DjB4 present distinct quaternary structure needs further investigation.

The highly diversity of Hsp40 proteins may be a responsible factor for several kinds of interactions between Hsp70 and Hsp40 and, thus, the necessity of different quaternary structure for proteins from subfamilies A and B. Several lines of evidence suggest multiple binding sites among Hsp70 and Hsp40 partners: Suh et al. (20) showed the existence of at least two binding sites for DnaK in DnaJ, both located at the J-domain. Linke et al. (35) suggest that a possible contact between the Cys-rich domain may be important to fulfill the transference of substrate from DnaJ to DnaK. Demand et al. (58) and Qian et al. (59) presented evidences that eukaryotic Hsp70 C-terminus interacts with Hsp40 from subfamily B. Laufen et al. (11) have proposed that the interaction cycle among the Hsp40 and Hsp70

may work in multiple steps, in which the Hsp40 protein might “target” unfolded substrate to Hsp70, as follows: 1) interaction between the J-domain with the Hsp70 NBD-ATP; 2) substrate shift from Hsp40 to the Hsp70 substrate-binding domain (SBD); 3) transference of the information regarding substrate binding, from the Hsp70 SBD to the Hsp70 NBD, resulting in the full ATP hydrolysis and closing of the Hsp70 SBD ‘lid’ and, finally, 4) Hsp40 dissociation from the Hsp70-ADP-substrate complex. Due to the high diversity of Hsp40 and the existence of multiple contacts among different Hsp40-Hsp70 partners (20, 35, 58, 59), the functional cycle described above may be constituted of further steps that depend on the Hsp40-Hsp70 partners.

The fact that DnaJ (subfamily A) is unable to bind to the isolated N- and C-domains of DnaK (20), but that Hsp40 from subfamily B does (58, 59), is in good agreement with the quaternary models proposed here: the J-domains of DjA1 are arranged in a position that allows binding to both domains of Hsp70 at once, while the J-domains of DjB4 can act independently of each other when binding to Hsp70. The quaternary structure of the DjA1 presented here is compatible with the model of interaction between DnaJ and DnaK proposed by Suh et al. (20; Fig. 5 in their paper) and the quaternary structure of the DjB4 presented here is compatible with the model of interaction between Sis1 and Hsp70 proposed by Sha et al. (26; Fig. 6 in their paper). Several authors (21, 26, 59, 60) have suggested that the peptide binding domain of Hsp70 nestles into the groove between the monomers of Hsp40. The size and the shape of the models presented here by SAXS, although they are at low-resolution, are compatible with these views. We show in Fig. 8 a diagram suggesting the possible interactions between DjA1 and DjB4 with Hsp70, considering the quaternary structure determined here for DjA1 and DjB4 and the current view regarding interaction of Hsp40 from subfamilies A and B with Hsp70 (20, 35, 58, 59). In this model, the interaction between DjA1-Hsp70 might occur by two binding sites located in each J-domain, one interacting with the Hsp70 NBD and the other with a region near to the Hsp70 substrate binding site (20). In the interacting model DjA1-Hsp70, the Cys-rich domain is in a proper position to interact with the Hsp70 and to transfer substrate to it, as proposed by Linke et al. (35). In the interaction model DjB4-Hsp70, the binding site located at one J-domain interacts with the Hsp70 NBD and the C-terminus of DjB4 is able to interact with the C-terminus of Hsp70, as previously suggested (58, 59). The size of DjB4 allows the suggested double interaction, in one cycle, between the J-domain and the Hsp70 NBD and between the Hsp40 C-terminus and the Hsp70 C-terminus.

The data presented here is in good agreement with the current view of the interaction between Hsp40 and Hsp70 described above. The conformational arrangement of the J-domain and G/F-rich region found for DjA1 and DjB4 may provide flexibility to these proteins, allowing them to accommodate and interact with both the nucleotide and substrate binding domains of Hsp70 in the same cycle of interaction (see discussion above). Low or high resolution structure studies from others Hsp40, from subfamilies A and B, would add more information about the interactions between Hsp70-Hsp40 partners. Finally, further dissections about those interactions in different Hsp40 proteins must be done to fully understand them and our results are relevant to clarify the features, from a structural point of view, that DjA1 and DjB4 are crucial to the chaperone function inside cells.

CONCLUSION

Small angle X-ray scattering experiments yielded low resolution structure models (molecular envelopes) of DjA1 and DjB4 proteins. A low-resolution model for the structure of a C-terminal deleted mutant of DjA1, DjA1₁₋₃₃₂, in which the dimer formation was interrupted, was also determined. The hydrodynamic properties calculated from these protein models agreed well with those directly determined from AUC experiments. This agreement provided support and confidence to the structure models derived from SAXS results. The characteristics of the low resolution models and the values of their molecular masses indicated that DjA1₁₋₃₃₂ was a monomer, and DjA1 and DjB4 were both dimers. An additional analysis combining the envelope functions derived from SAXS data and the known information about the high-resolution structure of Hsp40 C-terminal and J-domains, showed that DjA1 had a rather compact structure in which two highly asymmetric monomers bound in a bullet-shape, while in the DjB4 protein the N-termini were located in a completely different arrangement, pointing outward the compact C-termini core responsible for dimerization. The observed differences may explain why these two Hsp40 subfamilies have different chaperone activity and help to understand the activities of Hsp40 in substrate binding, transport to and interaction with Hsp70 proteins.

Acknowledgments: We thank Fundação de Amparo à Pesquisa do Estado de São Paulo and Conselho Nacional de Desenvolvimento Científico e Tecnológico for fellowships and financial support. We thank H.L.M. Guedes for discussion and the technical staff of LNLS for assistance.

References

- 1 Gething, M.J., and Sambrook, J. (1992) *Nature* **355**, 33-45.
- 2 Bukau, B., and Horwich, A.L. (1998) *Cell* **92**, 351-366.
- 3 Fink, A.L. (1999) *Physiol. Rev.* **79**, 425-449.
- 4 Mayer, M.P., Brehmer, D., Gässler C.S., and Bukau, B. (2001) *Adv. Protein Chem.* **50**, 1-45.
- 5 Hartl, F.U., and Hayer-Hartl, M. (2002) *Science* **295**, 1852-1858.
- 6 Mayer, M.P., Schröder H., Rüdiger, S., Pall, K., Laufen, T., and Bukau, B. (2000) *Nat. Struct. Biol.* **7**, 586-593.
- 7 Martin, J., and Hartl, F.U. (1997) *Curr. Opin. Struct. Biol.* **7**, 41-52.
- 8 Hendrick, J.P., Langer, T., Davis, T.A., Hartl, F.U., and Wiedmann, M. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 10216-10220.
- 9 Liberek, K., Wall, D., and Georgopoulos, C. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6224-6228.
- 10 Karzai, A.W., and MaMacken, R. (1996) *J. Biol. Chem.* **271**, 11236-11246.
- 11 Laufen, T., Mayer, M.P., Beisel, C., Klostermeier, D., Mogk, A., Reistein, J., and Bukau, B. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 5452-5457.
- 12 Langer, T., Lu, C., Echols, H., Flanagan, J., Hayer, M.K., and Hartl, F.U. (1992) *Nature* **356**, 683-689.
- 13 Szabo, A., Langer, T., Schroder, H., Flanagan, J., Bukau, B., and Hartl, F.U. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 10345-10349.
- 14 Rüdiger, S., Scheneider-Mergener, J., and Bukau, B. (2001) *EMBO J.* **20**, 1042-1050.
- 15 Yan, W., and Craig, E.A. (1999) *Mol. Cell. Biol.* **19**, 7751-7758.
- 16 Johnson, J.L., and Craig, E.A. (2001) *J. Cell. Biol.* **152**, 851-856.
- 17 Lee, S., Fan, C.Y., Younger, J.M., Ren, H., and Cyr, D.M. (2002) *J. Biol. Chem.* **277**, 21675-21682.
- 18 Fan, C.Y., Lee, S., Ren, H.Y., and Cyr, D.M. (2004) *Mol. Biol. Cell* **15**, 761-773.

- 19 Szabo, A., Korszun, R., Hartl, F.U., and Flanagan, J. (1996) *EMBO J.* **15**, 408-417.
- 20 Suh, W-C., Lu, C.Z., and Gross, C.A. (1999) *J. Biol. Chem.* **274**, 30534-30539.
- 21 Li, J., Qian, X., and Sha, B. (2003) *Structure* **11**, 1475-1483.
- 22 Greene, M.K., Maskos, K., and Landry, S.J. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6108-6113.
- 23 Szyperski, T., Pellechia, M., Wall, D., Georgopoulos, C., and Wuthrich, K. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 11343-11347.
- 24 Banecki, B., Liberek, K., Wall, D., Wawrzynow, A., Georgopoulos, C., Bertoli, E., Tanfani, F., and Zylicz, M. (1996) *J. Biol. Chem.* **271**, 14840-14848.
- 25 Martinez-Yamout, M., Legge, G.B., Zhang, O., Wright, P.E., and Dyson, H.J. (2000) *J. Mol. Biol.* **300**, 805-818.
- 26 Sha, B., Lee, S., and Cyr, D.M. (2000) *Structure* **8**, 799-807.
- 27 Lu, Z., and Cyr, D.M. (1998) *J. Biol. Chem.* **273**, 5970-5978.
- 28 Miernyk, J.A. (2001) *Cell Stress Chaperones* **6**, 209-218.
- 29 Borges, J.C., Peroto, M.C., and Ramos, C.H.I. (2001) *Genet. Mol. Biol.* **24**, 85-92.
- 30 Venter, J.C., Adams, M.D., Myers, E.W., *et al.* (2001) *Science* **291**, 1304-1351.
- 31 Cyr, D.M., Langer, T., and Douglas, M.G. (1994) *Trends Biochem. Sci.* **19**, 176-181.
- 32 Cheetham, M.E., and Caplan, A.J. (1998) *Cell Stress Chaperones* **3**, 28-36.
- 33 Ohtsuka, K., and Hata, M. (2000) *Cell Stress Chaperones* **5**, 98-112.
- 34 Lu, Z., and Cyr, D.M. (1998) *J. Biol. Chem.* **273**, 27824-27830.
- 35 Linke, K., Wolfram, T., Bussemer, J., and Jakob, U. (2003) *J. Biol. Chem.* **278**, 44457-44466.
- 36 Laemmli, U.K. (1970) *Nature* **227**, 680-685.
- 37 Edelhock, H. (1967) *Biochemistry* **6**, 1948-1954.
- 38 Gill, S.C., and von Hippel, P.H. (1989) *Anal. Biochem.* **182**, 319-326.
- 39 Van Holde, K.E., and Weischet, W.O. (1978) *Biopolymers* **17**, 1397-1403.
- 40 Stafford, W.F. (1994) *Methods Enzymol.* **240**, 478-501.
- 41 Johnson, M.L., Correia, J.J., Yphantis, D.A., and Halvorson, H.R. (1981) *Biophys. J.* **36**, 575-588.
- 42 Laue, T.M. (2001) *Curr. Opin. Struct. Biol.* **11**, 579-583.
- 43 Kellermann, G., Vicentin, F., Tamura E., Rocha, M., Tolentino, H., Barbosa, A., Craievich, A., and Torriani, I. (1997) *J. Appl. Crystallog.* **30**, 880-883.

- 44 Borges, J.C., Fischer, H., Craievich, A.F., Hansen, L.D., and Ramos, C.H.I. (2003) *J. Biol. Chem.* **278**, 35337-35344.
- 45 Svergun, D.I. (1992) *J. Appl. Crystallog.* **25**, 495-503.
- 46 Porod, G. (1982) in *Small-angle X-ray Scattering* (Glatter, O. and Kratky, O., Eds) pp. 17-51, Academic Press, London.
- 47 Svergun, D.I. (1999) *Biophys. J.* **76**, 2879-2886.
- 48 Garcia de la Torre, J., Huertas, M.L., and Carrasco, B. (2000) *Biophys. J.* **78**, 719-730.
- 49 Garcia de la Torre, J., Carrasco, B., and Harding, S.E. (1997) *Eur. Biophys. J.* **25**, 361-372.
- 50 Terada, K., and Mori, M. (2000) *J. Biol. Chem.* **275**, 24728-24734.
- 51 Feigin, L.A., and Svergun, D.I. (1987) *Structure analysis by Small-angle X-ray and Neutron Scattering*, pp. 83-87, Plenum Press, New York.
- 52 Volkov, V.V. and Svergun, D.I. (2003) *J. Appl. Cryst.* **36**, 860-864.
- 53 Zimm, B.H. (1974) *Biophys. Chem.* **1**, 279-291.
- 54 Laue, T.M., Shal, B.D., Ridgeway, T.M., and Pelletier, S.M. (1992) pp. 90-125, Eds. S. Harding and A. Rowe, Royal Society of Chemistry, London.
- 55 Waxman, E., Laws, W.R., Laue, T.M., and Ross, J.B.A. (1994) pp. 189-208, Eds Todd M. Schuster and Thomas M. Laue, Birkhäuser, Boston.
- 56 Qian, Y.Q., Patel, D., Hartl, F.U., and McColl, D.J. (1996) *J. Mol. Biol.* **260**, 224-235.
- 57 Lopez, N., Aron, R., and Craig, E.A. (2003) *Mol. Cell Biol.* **14**, 1172-1181.
- 58 Demand, J., Lüders, L., and Höhfeld, J. (1998) *Mol. Cell Biol.* **18**, 2023-2028.
- 59 Qian, X., Hou, W., Zhengang, L., and Sha, B. (2002) *Biochem. J.* **361**, 27-34.
- 60 Landry, S.J. (2003) *Structure* **11**, 1465-1466.

Tables

Table I: The molecular mass M from DjA1, DjA1₁₋₃₃₂ and DjB4 predicted from their amino acids sequence and determinate by SAXS and AUC. Data from SAXS were determined from the values of the extrapolated scattering intensity $I(0)$ by comparison with that one of bovine serum albumin (69 kDa). Sedimentation equilibrium technique was used to calculate M for DjA1 and DjB4. For DjA1₁₋₃₃₂, the M was calculated by the relation s/D (Equation 1) with the sedimentation coefficient ($s_{20,w}^0$) estimated from sedimentation velocity analysis and the diffusion coefficient ($D_{20,w}^0$) estimated from dynamic light scattering experiment (data not shown).

Method	Molecular Mass (kDa)		
	DjA1 dimer	DjA1 ₁₋₃₃₂	DjB4 dimer
Amino acid sequence	94.0	39.5	80.0
SAXS	95 ± 10	40 ± 5	90 ± 10
AUC	93.0 ± 1.0	42.0 ± 2.0	80.0 ± 1.0

Table II: Structural parameters corresponding to DjA1, DjA1₁₋₃₃₂ and DjB4 proteins in solution, studied by SAXS, and mathematical parameters associated to GNOM and DAMMIN programs. Errors are less than 10%.

Mathematical and structural parameters	DjA1 dimer		DjA1 ₁₋₃₃₂		DjB4 dimer	
	GNOM	DAMMIN	GNOM	DAMMIN	GNOM	DAMMIN
Free parameter	19	2098	19	5594	26	3350
Discrepancy χ	-	0.85	-	1.26	-	1.32
Resolution (Å)	15	30	15	21	15	30
$D_{\max}$ (Å)	140	145	140	140	200	195
R_g (Å)	46	46	41	41	56	55

Table III: Hydrodynamics parameters of DjA1, DjA1₁₋₃₃₂ and DjB4 in solution calculated from both AUC data and structural parameters derived from DAM model generated by SAXS experiments using the HydroPro software. The partial specific volume, solvent viscosity, and density were determined by Sednterp software. Errors are less than 2%.

Hydrodynamic parameters	Method					
	DjA1 dimer		DjA1 ₁₋₃₃₂		DjB4 dimer	
	AUC	DAM by HydroPro	AUC	DAM by HydroPro	AUC	DAM by HydroPro
$s_{20,w}^0$ (S)	4.63	4.44	2.85	2.46	3.78	3.63
$D_{20,w}^0$ (10^7 cm ² /s)	4.5	4.2	6.5	5.6	4.4	4.1
P (f/f ₀)	-	1.33	-	1.40	-	1.36

Figure captions

Figure 1: Hsp40 chaperones. **A)** Schematic representation of the Hsp40 subfamilies (A, B and C) domain structure. **B)** Homology comparison of the amino acid sequences of DjA1 (gene DNAJA1 – GenBank accession no. NM_001539) and DjB4 (gene DNAJB4 – GenBank accession no. NP_008965). The nomenclature is from Ohtsuka and Hata (33). Gray box: J-domain; underlined gray line: G/F-rich region; underlined black line: Cys-rich domain; white box: C-terminal domain. The arrow indicates the last residue in the DjA1₁₋₃₃₂ deletion mutant. DjA1 has 47% of similarity with DjB4 (81% at the J-domain).

Figure 2: Expression, purification, secondary and quaternary structure. **A)** SDS-PAGE 12%. Lane 1: Molecular mass markers in kDa (left); lanes 2, 6 and 10: bacterial pellet before induction; lanes 3, 7 and 11: bacterial pellet after induction with IPTG; lane 4, 8 and 12: supernatant of lysed cells; lanes 5, 9 and 13: purified proteins. **B)** Analytical molecular exclusion chromatography. Bottom panel: The measured retention times of standard proteins (see Material and Methods). Top panel: The measured retention times for DjA1, DjA1₁₋₃₃₂, and DjB4. The estimated apparent molecular masses were 120 kDa, 100 kDa and 40 kDa, respectively. **C)** Circular dichroism. Residual molar ellipticity of DjA1, DjA1₁₋₃₃₂ and DjB4 were measured from 195 to 260 nm in 25 mmol/liter Tris-HCl buffer (pH 7.5) containing 500 mmol/liter of NaCl and 1 mmol/liter β -mercaptoethanol at 20 °C. The

amount of secondary structure was estimated by CDNN Deconvolution software and they were similar among the two proteins: 35% β -sheet, 10% α -helix, 20% of turn, and 35% random coil (errors are less than 10%).

Figure 3: Chaperone activity measurements. SDS-PAGE 10% showing the samples of chemical - (A) and thermal (B) unfolded luciferase (LUC) in the presence and in the absence of increasing concentrations of Hsp40 proteins (5 to 15 μ mol/liter) (see Material and Methods). Samples of free luciferase were loaded for comparison purpose.

Figure 4: Experimental SAXS curves of DjA1₁₋₃₃₂, DjA1 and DjB4 in solution and results of the fitting procedures. A) log I versus q experimental curve. The open squares (DjA1₁₋₃₃₂), circles (DjA1), and triangles (DjB4) represent the experimental curve and the continuous line of the scattering intensity from the DAM (obtained from DAMMIN). B) Distance distribution functions of DjA1, DjB4 and DjA1₁₋₃₃₂. The arrows point out the maximum diameters D_{max} .

Figure 5: Low-resolution DAM models generated from SAXS data from human Hsp40 and molecular fitting. A) The DjA1 dimer is shown in white (left, dimensions: 140 x 92 x 52 Å), the and DjA1₁₋₃₃₂ monomers are shown in red and yellow (center, dimensions each: 140 x 57 x 42 Å). DjA1₁₋₃₃₂ is presented as a dimer model constructed from monomers for comparison with DjA1. The dimer DjA1₁₋₃₃₂ has hydrodynamic properties, predicted by HydroPro software (data not shown), very similar to those from DjA1. One view of the superposition of DjA1 and DjA1₁₋₃₃₂ is represented in right. B) The DjA1₁₋₃₃₂ *ab initio* model showed to be a good envelope for the available structure of Hsp40 domains. Red: the NMR structure of the J-domain (residues 2-77, 1HDJ – ref. 56), and Green: the crystallographic structure of the Cys-rich and C-terminal domains (residues 103-350, 1NLT – ref. 21), connected by a dot line (G/F-rich region, structure not available). The subdomains (I to III, ref. 21) of the Cys-rich and C-terminal domains are shown. C) The domains disposition profile obtained in C) was used to construct the domains disposition profile for the DjA1 SAXS model. Red: the NMR structure of the J-domain (residues 2-77, 1HDJ – ref. 56), and Green: the crystallographic structure of the Cys-rich and C-terminal domains (residues 103-350, 1NLT – ref. 21), connected by a dot line (G/F-rich region, structure not available). D) *Ab initio* model developed to DjB4 (dimensions: 200 x 93 x 48 Å). E) The DjB4 *ab initio*

model showed to be a good envelope for the available structure of Hsp40 domains. Red: the NMR structure of the J-domain (residues 2-77, 1HDJ – ref. 56), and Green: the crystallographic structure of the C-terminal domains (residues 180-349, 1C3G – ref. 26), connected by a dot line (G/F-rich region, structure not available). The relative position between DjB4 and DjA1 J-domains is a strong indication of different quaternary structure between these proteins. The superposition was performed using the WebLab ViewerLite software.

Figure 6: Sedimentation equilibrium experiment. Sedimentation equilibrium experiments were performed at 6,000, 8,000, and 10,000 rpm (AN-60Ti rotor) with scan data acquisition at 238 nm and in 25 mmol/liter Tris-HCl buffer (pH 7.5), 500 mmol/liter NaCl, glycerol 1% and 0.5 mmol/liter β -mercaptoethanol. The best fits of experimental data for 250 μ g/mL of either DjA1 or DjB4 at 6,000, 8,000 and 10,000 rpm at 20 °C are shown (see Material and Methods for details). The random distribution of the residuals (bottom panel) indicates that the fit is satisfactory. **A)** Sedimentation equilibrium data of DjA1 agrees with a dimer structure with 93.0 ± 1.0 KDa. **B)** Sedimentation equilibrium data of DjB4 agrees with a dimer structure with 80.0 ± 1.0 KDa.

Figure 7: Sedimentation velocity experiments. **A)** Sedimentation velocity experiments carried out at 20 °C, with scan data acquisition at 230 and 238 nm at low and high protein concentration, respectively. Rotor (AN-60Ti) velocities were 25,000 rpm for DjA1 and DjB4 and 30,000 rpm for DjA1₁₋₃₃₂. The figure displays experiments with 0.6 mg/mliter of DjA1, 0.8 mg/mliter of DjA1₁₋₃₃₂, and 0.48 mg/mL of DjB4 in 25 mmol/liter Tris-HCl buffer (pH 7.5), 500 mmol/liter NaCl, glycerol 1% and 0.5 mmol/liter β -mercaptoethanol (except for DjA1₁₋₃₃₂ that was carried out in the absence of glycerol. The $g(s^*)$ distributions were fitted using the Origin (Microcal Software) with a Gaussian giving apparent sedimentation coefficients s^* of about 3.8, 3.2 and 2.65 Svedberg for DjA1, DjB4 and DjA1₁₋₃₃₂, respectively. **B)** Plots of $s_{20,w}$ versus protein concentrations fitted by linear regression to calculate the $s_{20,w}^0$: 4.63 ± 0.05 , 3.78 ± 0.01 and 2.85 ± 0.03 , for DjA1, DjB4 and DjA1₁₋₃₃₂, respectively.

Figure 8: Proposed models for the interaction of DjA1 and DjB4 with Hsp70. See text for discussion.

Figure 1A

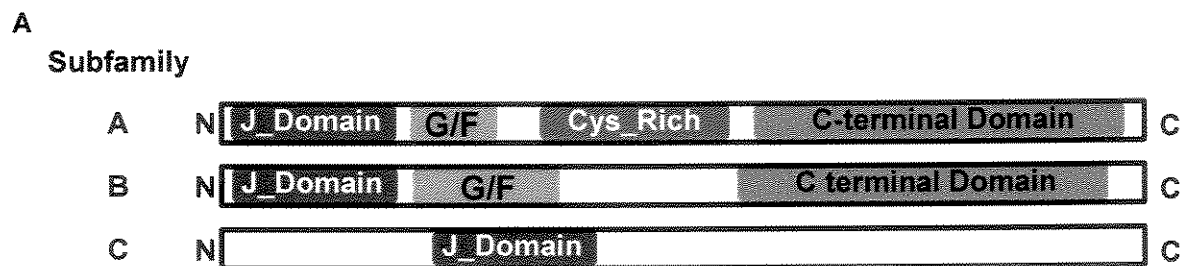


Figure 1B

B

J_Domain

DjA1	1	MVKETTTYDVLGVKPNATQEELEKKAYRKLALKYHPDKN--PNEGEKFKQISOAYEVLSDAKKRELYDKGGEOAIKEGGAG
DjB4	1	--MGKDYCYCILGIEKGASDEDIKKAYRKQALKFHPDKNKSPQAEKFKVAEAYEVLSDPKKREIYDQFGEEGLK-GGAG
DjA1	79	GGFGSPMDIFDMFFGGGGRMORERRGKNNVHOLSVTLEDLYNGATRKALOKNVICDKCEGRGKKGAVECCPNCRGTCM
DjB4	78	GTDTGGGTFRYTFHGDPHATFAAFFGGSNPFEIFFGRRMGGRDSEEMEIDGDPFSAFGFSMNG-----YPRDRNSVG
DjA1	159	QIRIHOIGPGMVQOIOSVCMEOGHERISPKDRCKSCNGRKIVREKKILEVHIDKGMKDQKITFHGEGDQEPGLEPGD
DjB4	150	PSRLKQDPPVIHELRSLEEIYSGCTKRMKISRKLNADGRSYRSEDKILTIEIKKGWKEGTKITFPREGDETFNSIPAD
DjA1	238	IIIVLDQKDHAVFTRRGEDLFMCMDIQLVEALCGFQKPISTLDNRTIVITSHPGQIVKHGDIKCVLNEGMPYRYPYKKG
DjB4	230	IVFIIKDKDHPKFKRDGSNIITYTAKISLREALCGCSINVPTLDGRNIPMSVN--DIVKPGMRRIIGYGLPFPKNPDQRG
DjA1	318	RLIEFKVNFPENGFLSPDKLSLLEKLLPERKEVEETDEMDOVELVDFDPNQERRRHYNGEAYEDDEHHPRGGVQC
DjB4	308	DLLEFEVSFPDT--ISSSSKEVLRKHLPAS-----

Figure 2A

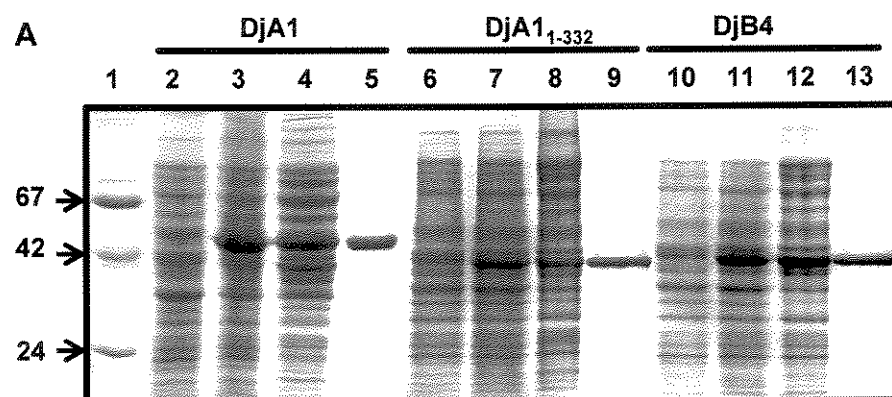


Figure 2B

B

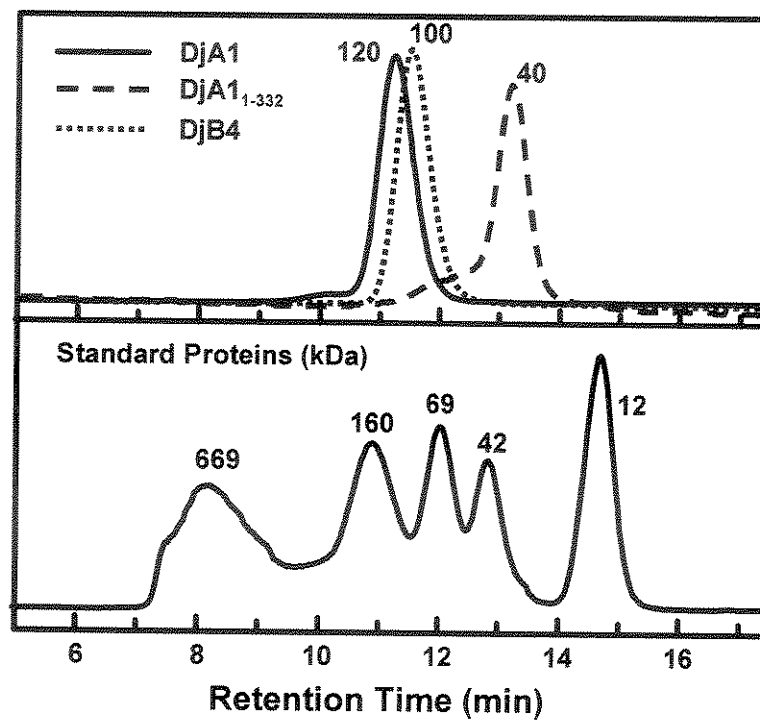


Figure 2C

C

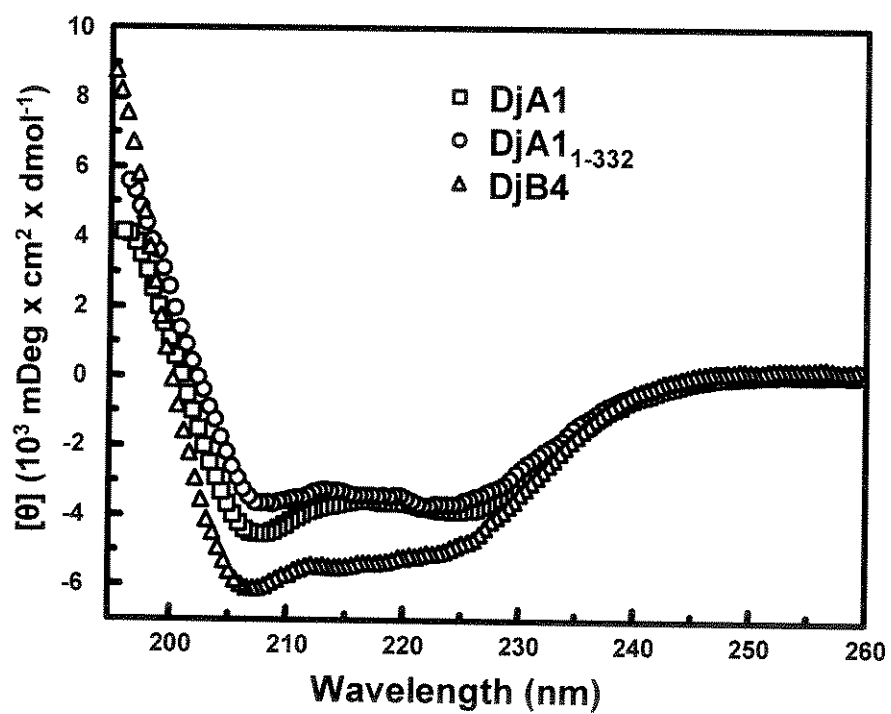


Figure 3A

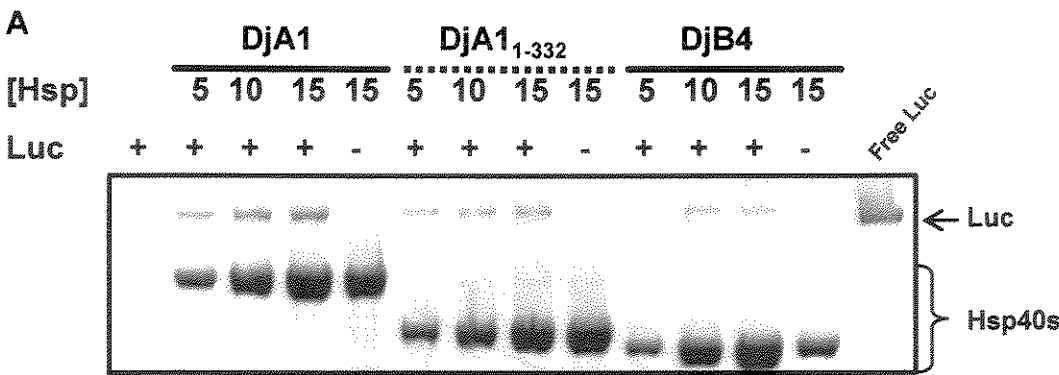


Figure 3B

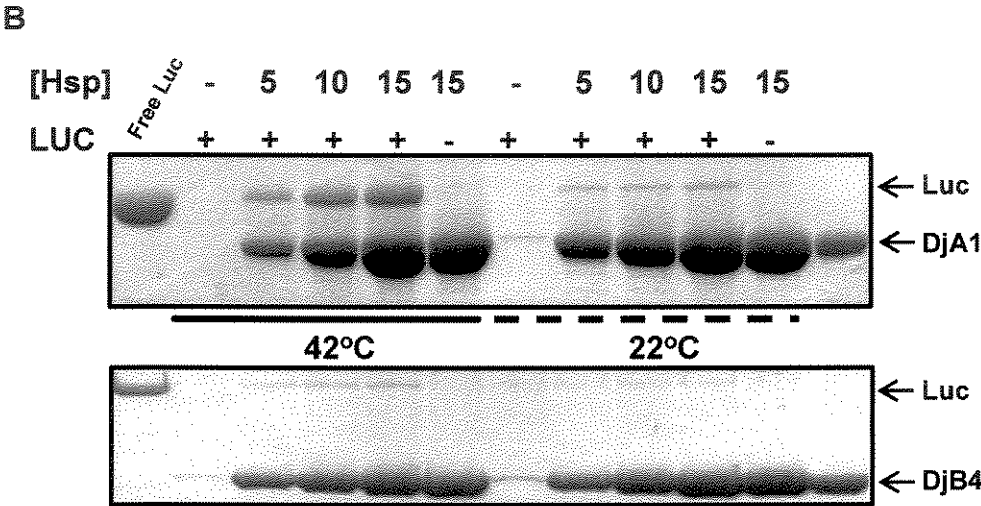


Figure 4A

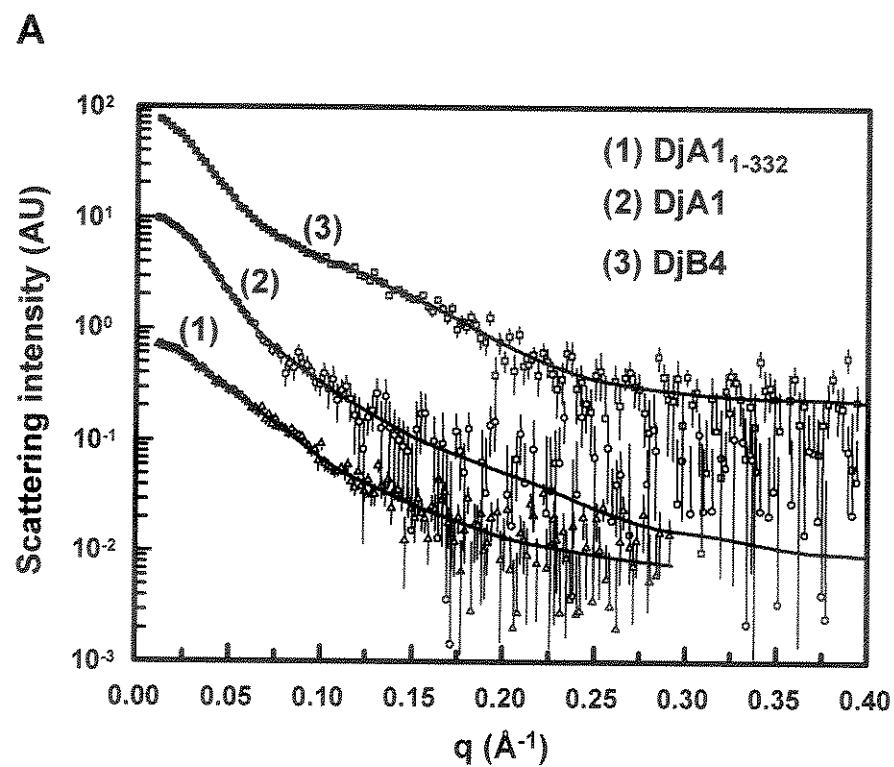


Figure 4B

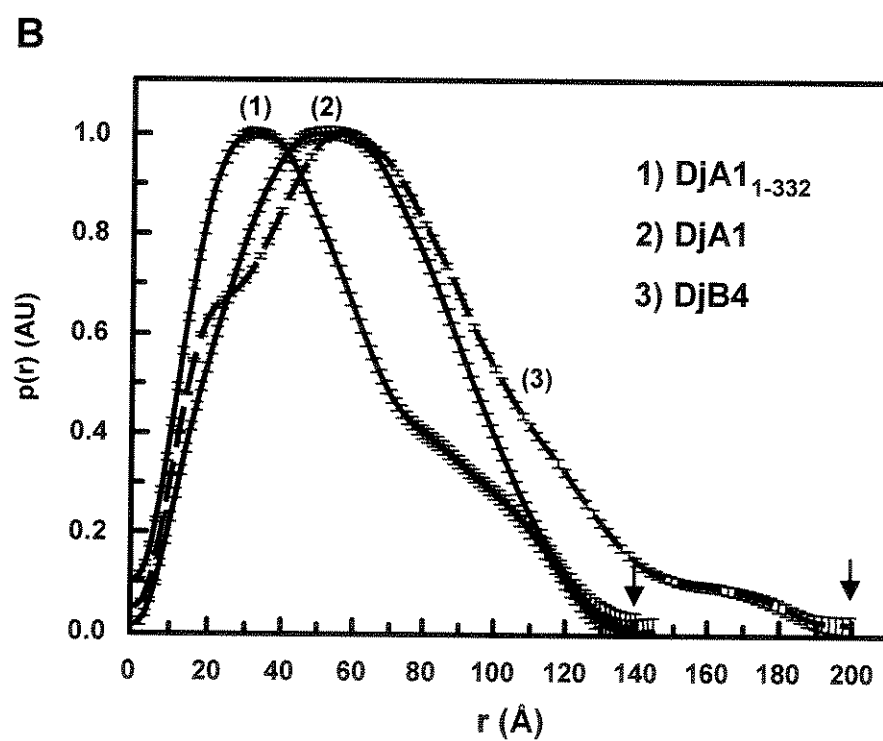


Figure 5

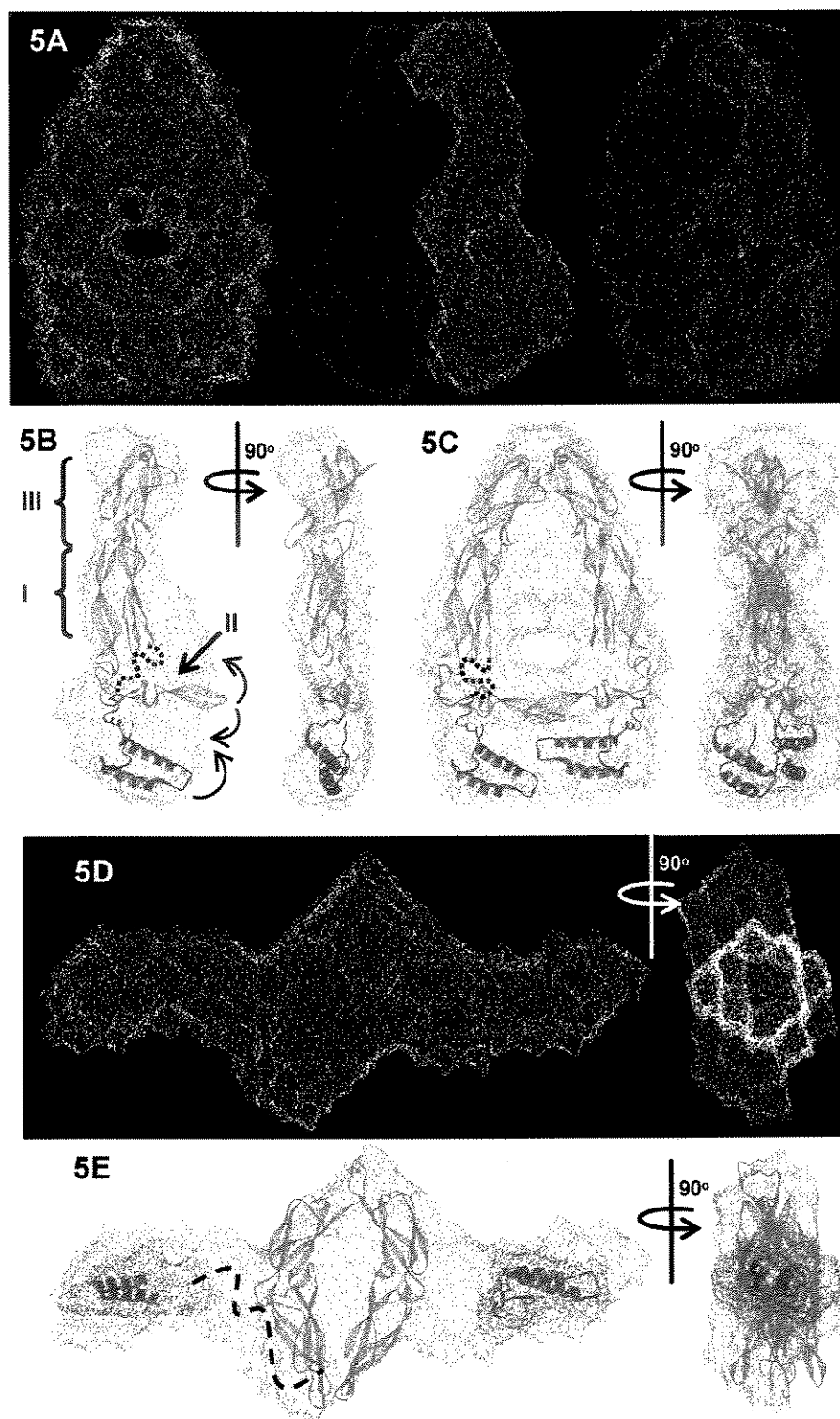


Figure 6A

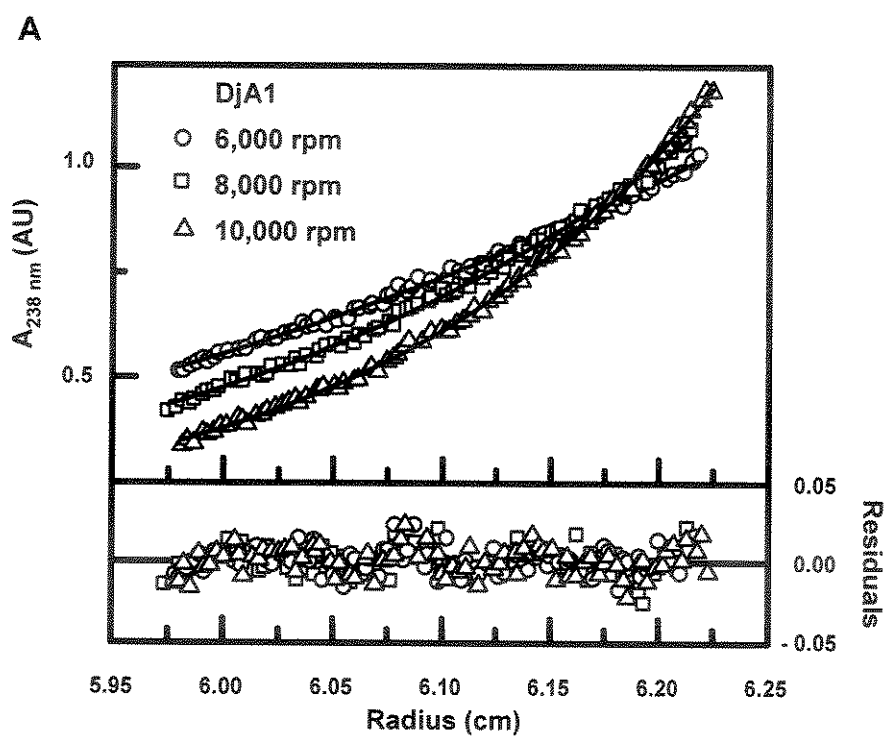


Figure 6B

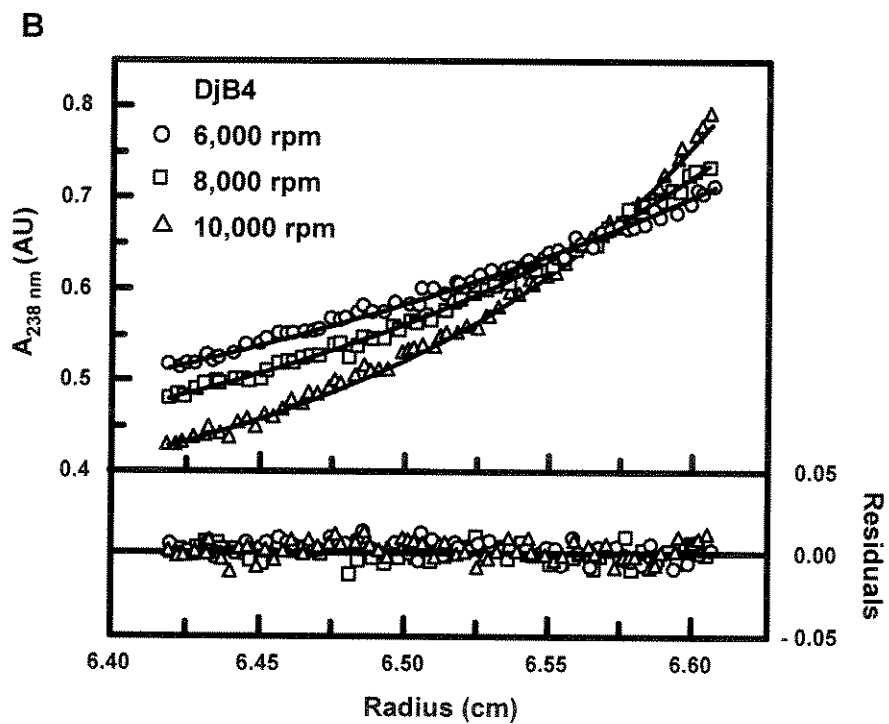


Figure 7A

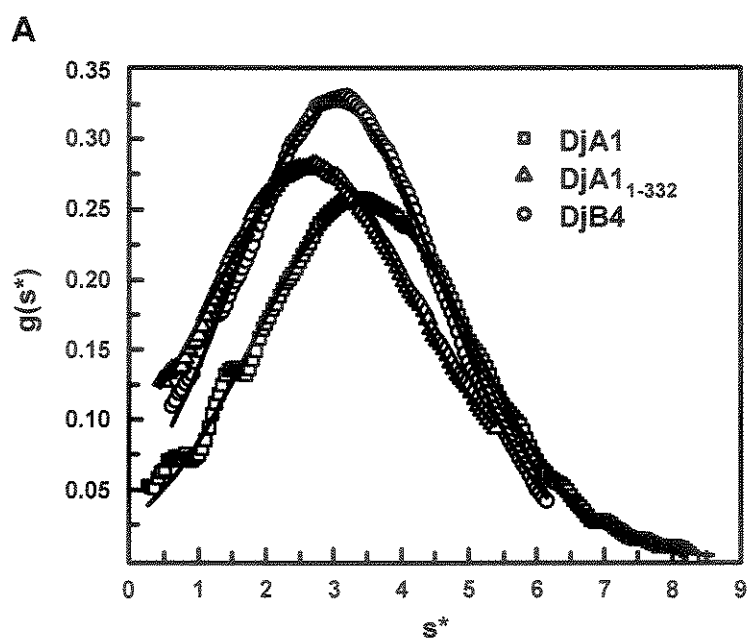


Figure 7B

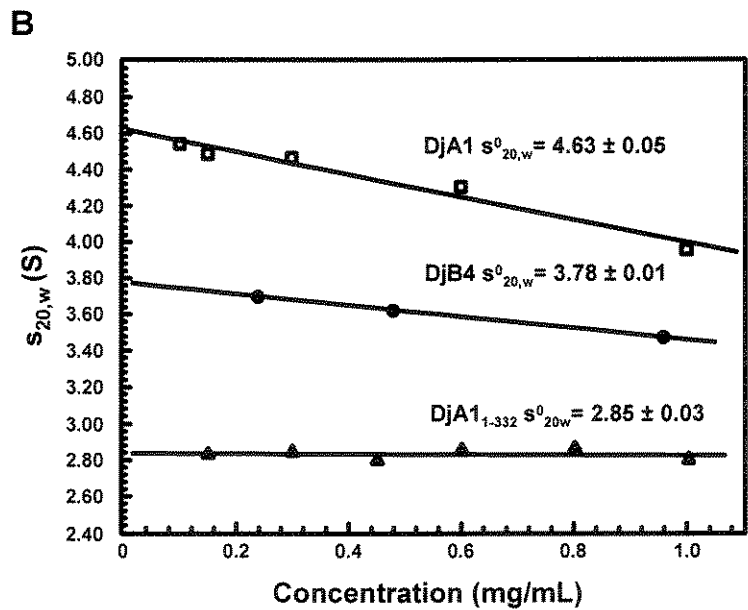
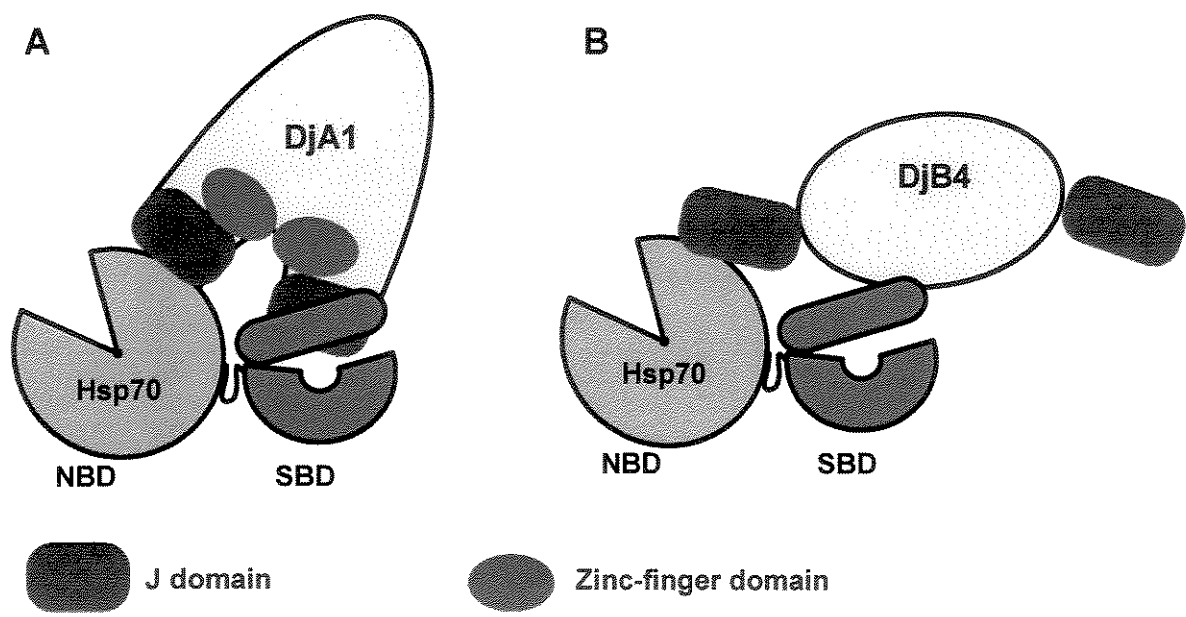


Figure 8



Abaixo é apresentada uma cópia da mensagem recebida do editor da revista *J. Biol. Chem.*, certificando a submissão de tal manuscrito a esta revista.

-----Original Message-----

From: jbc@asbmb.faseb.org [mailto:jbc@asbmb.faseb.org]

Sent: sábado, 24 de julho de 2004 09:00

To: cramos@lnls.br

Subject: From the JBC re: Manuscript M4:08349

M4:08349

Dear Dr. Ramos:

We have received your manuscript entitled: "Low-resolution structural study of two human Hsp40 chaperones in solution. DjA1 from subfamily A and DjB4 from subfamily B, have different quaternary structures".

It has been given the number M4:08349 and has been assigned to:

Dr. Norma M. Allewell
Dean, College of Life Sciences
University of Maryland
2300 Symons Hall
College Park, MD 20742
Tel: 301-854-3413
Fax: 301-314-9949
E-mail: nallewell@asbmb.faseb.org

Manuscripts submitted under multiple authorship are reviewed on the assumption that all listed authors concur in the submission and that the final version of the manuscript has been seen and approved by all authors. Papers are published with the implicit understanding that you will pay charges at a rate of \$65.00 per page for all typeset pages.

A detailed description of JBC's Editorial policies and practices can be found on the Journal's website at <http://www.jbc.org/misc/edpolicy.shtml>. These can be summarized as follows:

The primary criteria for judging the acceptability of a manuscript are its originality and scientific importance. Manuscripts judged lacking in these respects will be declined, even if the experimental work appears technically sound. This policy permits declination of a manuscript solely on the Editorial Board's judgment that the studies reported are not sufficiently original or important to merit publication in the JOURNAL. After receipt of a manuscript by the Editor, it is sent to an Associate Editor who usually assigns it to a member of the Editorial Board. The Board member then makes a definitive recommendation for acceptance, revision, or declination based on the scientific merit and technical quality of the studies reported. The Board member may consult outside referees when additional expertise is required. If a manuscript is not recommended by the Board member for publication, the Associate Editor can decline it without further review.

8.2 Anexo III

Coeficiente de absorvidade molar ($\text{mol/L}^{-1} \cdot \text{cm}^{-1}$) para as proteínas recombinantes utilizados para a quantificação das proteínas. Os dados foram calculados pelo programa ProtParam Tool (<http://ca.expasy.org/tools/protparam.html>) de acordo com *Gill et al.* (1989). As proteínas Hsp70-1A, mortalina, DjA1, DjA1₁₋₃₃₂ e DjB4 e foram consideradas em fusão com a cauda amino-terminal de poli-histidina. Dados para as proteínas GrpE#1 e GrpE#2 humanas foram calculados sem o peptídeo precursor.

Proteína Recombinante	$\epsilon - \text{mol/L}^{-1} \cdot \text{cm}^{-1}$ (por λ em nm)				
	276	278	279	280	282
Hsp70-1A	32.840	32.454	31.735	30.820	29.440
Mortalina	20.190	19.854	19.350	18.730	17.840
DjA1	16.820	16.162	15.515	14.800	13.920
DjA1 ₁₋₃₃₂	13.920	13.362	12.825	12.240	11.520
DjB4	21.640	24.254	20.695	20.010	19.040
GrpE#1	7.395	7.127	6.845	6.520	6.120
HisGrpE#2	4.350	4.200	4.035	3.840	3.600
GrpE#2	4.785	4.581	4.395	4.200	3.960

8.3 Anexo IV

Borges, J.C., Fischer, H., Craievich, A.F., Hansen, L.D. & Ramos, C.H.I. (2003) Free human mitochondrial GrpE is a symmetric dimer in solution. *J. Biol. Chem.*, 278, 35337-35344.

Free Human Mitochondrial GrpE Is a Symmetric Dimer in Solution*

Received for publication, May 14, 2003, and in revised form, June 24, 2003
Published, JBC Papers in Press, July 2, 2003, DOI 10.1074/jbc.M305083200

Júlio C. Borges^{‡§¶}, Hannes Fischer^{¶||}, Aldo F. Craievich^{||}, Lee D. Hansen^{**},
and Carlos H. I. Ramos^{‡§¶¶}

From the [‡]Centro de Biologia Molecular Estrutural, Laboratório Nacional de Luz Sincrotron, Campinas, São Paulo 13084-971, Brazil, the [§]Departamento de Bioquímica, Instituto de Biologia, UNICAMP, Campinas, São Paulo 13084-971, Brazil, the ^{||}Departamento de Física Aplicada, Instituto de Física, Universidade de São Paulo, São Paulo 05389-970, Brazil, and the ^{**}Department of Chemistry and Biochemistry, Brigham Young University, Provo, Utah 84602

The co-chaperone GrpE is essential for the activities of the Hsp70 system, which assists protein folding. GrpE is present in several organisms, and characterization of homologous GrpEs is important for developing structure-function relationships. Cloning, producing, and conformational studies of the recombinant human mitochondrial GrpE are reported here. Circular dichroism measurements demonstrate that the purified protein is folded. Thermal unfolding of human GrpE measured both by circular dichroism and differential scanning calorimetry differs from that of prokaryotic GrpE. Analytical ultracentrifugation data indicate that human GrpE is a dimer, and the sedimentation coefficient agrees with an elongated shape model. Small angle x-ray scattering analysis shows that the protein possesses an elongated shape in solution and demonstrates that its envelope, determined by an *ab initio* method, is similar to the high resolution envelope of *Escherichia coli* GrpE bound to DnaK obtained from single crystal x-ray diffraction. However, in these conditions, the *E. coli* GrpE dimer is asymmetric because the monomer that binds DnaK adopts an open conformation. It is of considerable importance for structural GrpE research to answer the question of whether the GrpE dimer is only asymmetric while bound to DnaK or also as a free dimer in solution. The low resolution structure of human GrpE presented here suggests that GrpE is a symmetric dimer when not bound to DnaK. This information is important for understanding the conformational changes GrpE undergoes on binding to DnaK.

Nascent proteins in the cell sometimes require the assistance of one or more protein complexes named molecular chaperones to fold correctly (1, 2). An important chaperone complex is composed of the molecular chaperones Hsp70,¹ Hsp40 (or

DnaK and DnaJ, respectively) and GrpE, which are highly expressed and important for several cell processes (3–7). The Hsp70 affinity for unfolded proteins is regulated by nucleotide binding to its nucleotide binding domain (NBD), which has a molecular mass of about 45 kDa (5, 8). The Hsp70 C terminus forms the substrate binding domain (SBD), which is capable of binding hydrophobic amino acid residues and has a molecular mass of about 20 kDa (8, 9). Co-chaperones Hsp40 and GrpE interact both *in vivo* and *in vitro* with DnaK (6, 10, 11), stimulating its ATPase activity (12) and regulating the ability of DnaK to bind and stabilize unfolded proteins (13). The importance of GrpE in the Hsp70 chaperone machinery is shown by the following: it is essential for bacterial viability at all temperatures (14), GrpE acts as an exchange factor that releases nucleotides bound to DnaK (12), and it is important for DnaK recycling (11). The first indication that GrpE was a homodimer came from cross-linking studies with glutaraldehyde (15). Subsequently, analytical ultracentrifugation experiments (16) showed that GrpE has a dimeric structure with an elongated shape that binds DnaK with 2:1 stoichiometry.

The crystallographic high resolution structure of residues 34–197 of *Escherichia coli* GrpE (EcGrpE_{34–197}) complexed with the *E. coli* DnaK-NBD (EcDnaK_{3–383}) corroborates that GrpE forms a dimer and shows that only one of the subunits, known as the proximal monomer, binds to Hsp70 (17). In this structure, the GrpE dimer is asymmetric because the proximal monomer adopts a more open conformation than the distal monomer. Knowing whether this asymmetric conformation remains while GrpE is free in solution is necessary for understanding this protein structure-function relationship in the cell. The GrpE C-terminal domain, EcGrpE_{141–197}, is β -structured and binds to the DnaK-NBD causing the release of ADP from the NBD (17). The function of the GrpE N terminus remains to be understood fully. The GrpE_{40–86} domain forms a long coil-coiled α -helical structure (17) and may function as a thermosensor because it appears to be responsible for the GrpE thermal transition at physiological temperatures (18). GrpE appears to be the only component of the Hsp70 chaperone machinery which undergoes a thermal transition at a physiologically relevant temperature (19). The GrpE_{89–137} domain forms a four-helix bundle (17) and likely acts as the stabilization center for dimerization (18). GrpE is present in eukaryotes (20–22), and it is generally assumed that they share high structural similarity with their prokaryote homologs. Thus, it

* This work was supported in part by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), PEW Charitable Trust, and Conselho Nacional de Desenvolvimento Científico Tecnológico. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

This paper is dedicated to Prof. Rogério Meneghini on the occasion of his retirement.

¶ Supported by a fellowship from the Fundação de Amparo à Pesquisa do Estado de São Paulo.

¶¶ To whom correspondence should be addressed: Centro de Biologia Molecular Estrutural, Laboratório Nacional de Luz Sincrotron, P. O. Box 6192, Campinas, SP 13084-971, Brazil. Tel.: 55-19-3287-4520; Fax: 55-19-3287-7110; E-mail: cramos@lnls.br.

¹ The abbreviations used are: Hsp, heat shock protein; D_{max} , maxi-

mum diameter; DAM, dummy atom model; DSC, differential scanning calorimetry; DTT, dithiothreitol; NBD, nucleotide binding domain of Hsp70; R_g , radius of gyration; SAXS, small angle x-ray scattering analysis; SBD, substrate binding domain of Hsp70.

is important to characterize the structure and function of GrpE from diverse organisms to test this hypothesis. The sequence of human mitochondrial GrpE (23) is represented in Fig. 1 along with the *E. coli* GrpE sequence showing that they share about 30% identity. The mitochondrial form of human GrpE is expressed as an immature protein that matures when its mitochondrial signaling peptide is lost.

Human GrpE was cloned and expressed in *E. coli* and purified by ion exchange chromatography and preparative molecular exchange chromatography. The secondary structure characterized by circular dichroism (CD) showed that the protein is folded. Temperature-induced unfolding followed either by CD or by differential scanning calorimetry (DSC) showed that human GrpE unfolding is only partially reversible, whereas *E. coli* GrpE exhibits reversible unfolding up to 60 °C (19). Analytical ultracentrifugation data indicate that the protein has the molecular mass of a dimer, and the calculated sedimentation coefficient agrees with the value expected for a protein with an elongated shape. The elongated shape was also derived independently from small angle x-ray scattering analysis (SAXS). The hydrodynamic parameters derived from SAXS data agree with those determined by analytical ultracentrifugation. Our envelope models are similar to the envelope of the crystallographic structure of *E. coli* GrpE bound to the DnaK-NBD determined by Harrison *et al.* (17). The low resolution structure generated from our SAXS data in solution suggests that GrpE is a symmetric dimer when not bound to DnaK. The implications of our conclusions concerning human GrpE conformational changes in structure-function relationships of this co-chaperone are discussed.

EXPERIMENTAL PROCEDURES

Cloning, Expression, and Purification—The human cDNA IMAGE clone (GenBank accession number BE614754) was used for cloning the mitochondrial GrpE (Mt-GrpE#1, GenBank accession number Q9HAV7). Two primers were used to amplify the cDNA by PCR and to create the restriction enzyme sites *NdeI* and *XhoI* for cloning into pET23a vector (Novagen). The 5'-primer (5'-TCTCCCGGCATATGTGCACAG-3') containing the *NdeI* restriction site was designed to anneal downstream to the mitochondrial peptide signal, eliminating this sequence in the recombinant protein. The correct cloning was confirmed by DNA sequencing using an ABI 377 Prism system (PerkinElmer Life Sciences). These procedures created the pET23aHMGpE#1 vector, which was transformed in *E. coli* strain BL21(DE3) for protein expression by adding 0.4 mmol/liter isopropyl thio- β -D-galactoside at $A_{600} = 0.8$. The induced cells were grown for 5 h and harvested by centrifugation for 10 min at $2,600 \times g$. The bacterial pellet was resuspended in lysis buffer (50 mmol/liter Tris-HCl, pH 8.0, 50 mmol/liter KCl, 10 mmol/liter EDTA, 15 ml/liter medium), disrupted by sonication in an ice bath, and centrifuged as described above. The supernatant was dialyzed against equilibration buffer (25 mmol/liter Tris-HCl, pH 7.5, 1 mmol/liter DTT) and then submitted to ion exchange chromatography on Q-Sepharose resin using an ÄKTA FPLC (Amersham Biosciences). Human GrpE was eluted in 100 mmol/liter NaCl, dialyzed against the second equilibration buffer (25 mmol/liter Tris-HCl, pH 7.5, 150 mmol/liter NaCl, 1 mmol/liter DTT) and loaded on a HiLoad Superdex 200pg molecular exclusion column using an ÄKTA FPLC. The degree of purification of human GrpE was estimated by SDS-PAGE (24), and its concentration was determined spectrophotometrically using the calculated extinction coefficient for denatured proteins (25, 26).

Circular Dichroism (CD) Spectroscopy—CD measurements were recorded on a Jasco J-810 spectropolarimeter with temperature controlled by Peltier Type Control System PFD 425S. Human GrpE was resuspended in 10 mmol/liter phosphate buffer, pH 7.2, and 1 mmol/liter β -mercaptoethanol at a final concentration of 20 μ mol/liter. Data were collected at a scanning rate of 50 nm/min with spectral band width of 1 nm using a 1-mm path length cell at increasing temperatures (for details, see Fig. 3 legend). CDNN Deconvolution software (version 2; Bioinformatik.biochemtech.uni-halle.dee/cdnn) was employed for secondary structure prediction. All buffers used were of chemical grade and were filtered before use to avoid scattering from small particles.

DSC—The measurements of human GrpE thermal denaturation were done both in an N-DSC III differential scanning calorimeter (Cal-

orimetry Sciences Corp.) and in a VP-DSC Differential Scanning Calorimeter (MicroCal), which gave similar results. The measurements were performed with protein concentrations of 7, 45, and 130 μ mol/liter and in two sets of buffers; 25 mmol/liter Hepes pH 7.6, 50 mmol/liter KCl, 5 mmol/liter $MgCl_2$, and 25 mmol/liter Tris-HCl, 150–500 mmol/liter NaCl, 1 mmol/liter β -mercaptoethanol. The scan rate was varied from 0.5 to 1.5 °C/min, and the temperature measurement range was from 10 to 100 °C. The reversibility of unfolding was tested by performing several consecutive up and down scans and by scan rate variation. Base lines were run several times in all of the conditions mentioned above.

Analytical Ultracentrifugation—Sedimentation velocity and sedimentation equilibrium experiments were performed with a Beckman Optima XL-A analytical ultracentrifuge. The protein was tested in concentrations of 50, 100, and 200 μ g/ml in 25 mmol/liter Tris-HCl buffer at pH 7.5, with 150 mmol/liter NaCl and 0.5 mmol/liter DTT, with no apparent aggregation. The sedimentation velocity experiments were carried out at 20 °C, 40,000 rpm (AN-60Ti rotor), and the scan data acquisition was taken at 230 nm. The sedimentation equilibrium experiments were made at 20 °C at speeds of 7,000, 9,000, and 11,000 rpm using the AN-60Ti rotor. Scan data acquisition was done at 230 nm. The analysis involved fitting a model of absorbance *versus* cell radius data by nonlinear regression. All fits were done with the ORIGIN software package (MicroCal Software) supplied with the instrument. The van Holde-Weischet (27) (sediment coefficient plot), Second Moment (28), and the Sedimentation Time Derivative ($g(s^*)$ integral distribution) (29) methods were used to analyze the sedimentation velocity experiments. The methods used for analyzing both velocity and equilibrium experiments allow the calculation of the apparent sedimentation coefficient s , the diffusion coefficient D , and the molecular weight M . The ratio of the sedimentation to diffusion coefficient gives the molecular weight;

$$M = \frac{sRT}{D(1 - V\bar{v}\rho)} \quad (\text{Eq. 1})$$

R is the gas constant and T is the absolute temperature. The software Sednterp (www.jphilo.mailway.com/download.htm) was used to estimate protein partial specific volume ($V\bar{v} = 0.7432$ ml/g), buffer density ($\rho = 1.0052$ g/ml), and buffer viscosity ($\eta = 0.01$ poise). The Self-association method was used to analyze the sedimentation equilibrium experiments using several models of association for human GrpE to fit the data. The distribution of the protein along the cell, obtained in the equilibrium sedimentation experiments, was fitted with the Equation 2 (30),

$$C = C_0 \exp\left(\frac{M(1 - V\bar{v}\rho)\omega^2(r^2 - r_0^2)}{2RT}\right) \quad (\text{Eq. 2})$$

where C is the protein concentration at radial position r , C_0 is the protein concentration at radial position r_0 , and ω is the centrifugal angular velocity.

SAXS—Small angle x-ray scattering experiments were performed at the SAS beam line of the LNLS synchrotron radiation facility, Campinas, Brazil (31). Measurements were made with a monochromatic x-ray beam with a wavelength $\lambda = 1.488$ Å for a sample-detector distance of 840 mm covering the momentum transfer range $0.01 < q < 0.44$ Å⁻¹ ($q = 4\pi\sin\theta/\lambda$, where 2θ is the scattering angle). The scattering intensity was recorded using a one-dimensional position-sensitive x-ray detector. The scattering curves produced by the protein solutions of either 5.0 or 19.4 mg/ml and of the solvent (25 mM Tris-HCl, pH 7.5, and 1 mmol/liter DTT) were collected in many short (90 s) frames to monitor radiation damage and beam stability. The data were normalized to account for the natural decay in intensity of the synchrotron incident beam and corrected for inhomogeneous detector response. The scattering intensity produced by the buffer was subtracted, and the difference curves were scaled to equivalent protein concentration.

To determine the molecular mass of human GrpE, a 5 mg/ml bovine serum albumin (66 kDa) solution was used as a standard. The molecular mass of human GrpE was inferred from the ratio of the extrapolated $I(0)$ value of human GrpE to that of bovine serum albumin. Because there was no negative region in the pair distribution function, we can safely conclude that all solutions studied were in the "dilute" state, i.e. no interferences of scattering amplitudes were produced by the interaction of different isolated scattering objects.

Models and Computer Programs—The distance distribution function $p(r)$ and the radius of gyration R_g of the human GrpE protein were

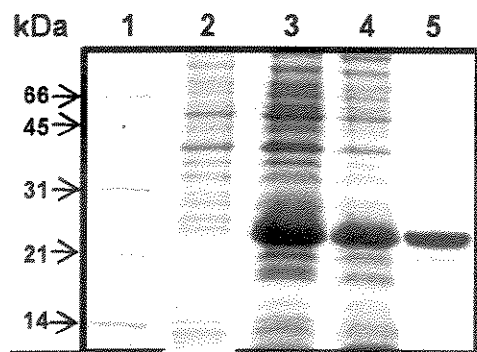


FIG. 2. SDS-PAGE showing recombinant human GrpE expression profile and the purified protein. Lane 1, molecular mass markers in kDa (right); lane 2, bacterial pellet before induction; lane 3, bacterial pellet after 5 h of induction with isopropyl thio- β -D-galactoside; lane 4, supernatant of lysed cells; lane 5, human GrpE after chromatographic purification. Human GrpE is expressed in large quantity and is highly pure.

in the CD signal at 222 nm from 4 to 50 °C and a more accentuated change above 50 °C (Fig. 3b). This temperature dependence is verified by DSC, the protein started to unfold at about 50 °C (Fig. 4). The unfolding of human GrpE is not completely reversible even if the protein is heated to only 50 °C as shown by both CD and DSC experiments (Figs. 3b and 4). The unfolding measured by DSC shows at least two unfolding events, one with a midpoint at 57 °C and the other, probably aggregation, occurs above about 62 °C (Fig. 4).

Hydrodynamic Analysis by Analytical Ultracentrifugation—The hydrodynamic properties of human GrpE were analyzed by both sedimentation velocity (Fig. 5a) and sedimentation equilibrium (Fig. 5b). In all of the experiments performed, GrpE behaved as a monodisperse protein system and did not show any sign of aggregation at concentrations from 50 up to 200 μ g/ml. Sedimentation velocity analysis shows that human GrpE has an apparent sedimentation coefficient of 2.25 ± 0.05 s and a predicted molecular mass of 42 ± 1 kDa (Table I). The results of sedimentation equilibrium experiments were consistent with a dimeric protein state and a molecular mass of 43 ± 1 kDa at all velocities and concentrations tested; the analysis of residuals evidenced the high quality of the fit. The data generated by analytical ultracentrifugation are thus consistent with a dimeric model for GrpE with an elongated shape and a molecular mass of 42.5 ± 1.5 kDa (Table I).

SAXS Measurements—The SAXS curves at 5.0 and 19.4 mg/ml exhibit a similar q dependence, and therefore the protein is in the same configuration in both solutions, and the solutions are both in the dilute state so that the basic assumption of isolated proteins producing independent scattering amplitudes is obeyed. The experimental SAXS curves after normalization and subtraction of parasitic scattering are displayed in Fig. 6, a and b, together with the computed curve from a DAM model as described under "Experimental Procedures." In Fig. 6a the SAXS intensity is plotted on a logarithmic scale to show the goodness of fit of the modeled function within the high q range. In Fig. 6b the SAXS intensity is given on a linear scale together with a Guinier plot ($\log I(q)$ versus q^2 , inset) (45), which allows determination of the R_g of the protein from the slope of the linear portion within the low q limit of the curve. The distance distribution function associated with the scattering object, $p(r)$, was evaluated with the GNOM package, and R_g was determined from the Guinier plot, also by applying the GNOM program. The molecular mass was obtained from SAXS results by comparing the intensity extrapolated to $q = 0$ with the corresponding data on bovine serum albumin and determined to be 38 ± 2 kDa. This value indicates that the GrpE

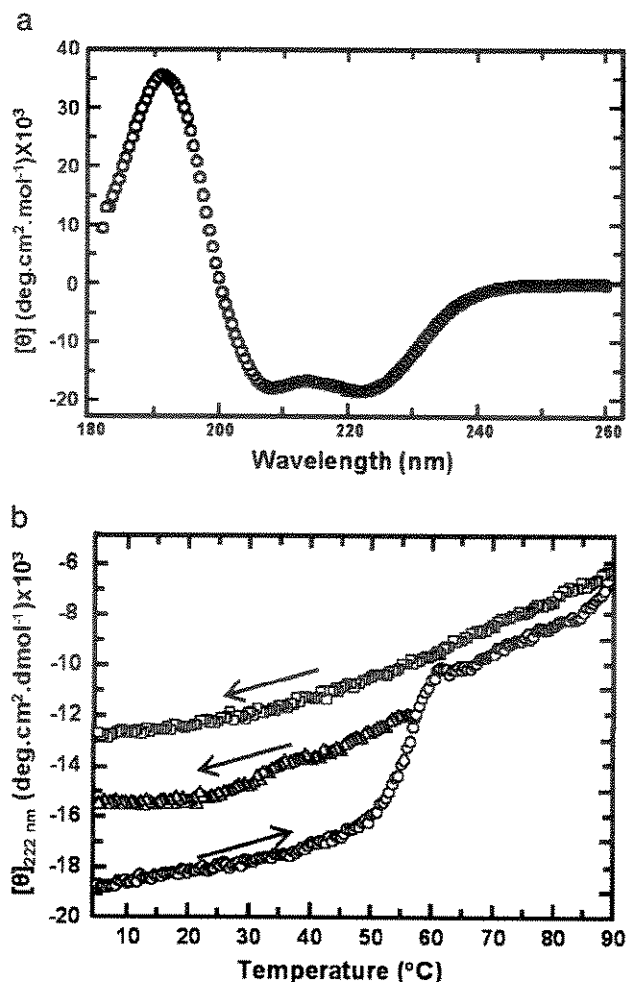


FIG. 3. CD experiments. a, residual molar ellipticity $[\theta]$ of human GrpE was measured from 180 to 260 nm in 10 mmol/liter phosphate buffer, pH 7.2, and 1 mmol/liter β -mercaptoethanol at 20 °C. The amount of secondary structure was calculated by CDNN deconvolution software: 55% α -helix, 25% β -sheet and turn, and 20% random coil (errors are less than 10%). b, temperature-induced unfolding of human GrpE as followed by CD at 222 nm at a scan rate of 1 °C/min. Unfolding from 4 to 90 °C data is shown by open circles, refolding from 57 to 4 °C data is shown by open triangles, and refolding from 90 to 4 °C data is shown by open squares. The refolding is partially reversible from 57 and 90 °C or if the protein was heated only to 50 °C (data not shown).

protein is in a dimeric state, both at low and high concentrations (5.0 and 19.4 mg/ml, respectively).

The experimental values of the maximum diameter of the protein, $D_{\max} = 155$ Å obtained from the curve plotted in Fig. 7 and the $R_g = 45.2$ Å, obtained from the Guinier plot, and calculated by GNOM program, suggest that the protein is very elongated. This is also evidenced by the shape of the distance distribution function $p(r)$ plotted in Fig. 7, which exhibits the typical asymmetry corresponding to very elongated scattering particles. This elongated shape of human GrpE caused the protein to elute from a gel filtration column earlier than expected for a typical globular protein with similar molecular mass (data not shown).

Ab Initio Shape Determination of Human GrpE by DAM-MIN—The human GrpE shape was determined by an *ab initio* procedure using the DAMMIN program (35). In this method, a sphere of diameter $D_{\max}$ is filled with densely packed small spheres (dummy atoms) with radius $r_0 \ll D_{\max}$. This method uses a simulated annealing algorithm and searches to minimize a function that accounts for the differences between ex-

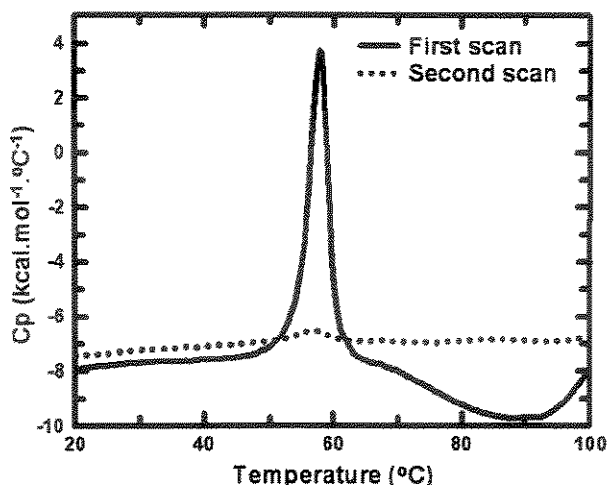


FIG. 4. DSC. Human GrpE temperature-induced unfolding was investigated by DSC with protein concentrations of 7, 45, and 130 $\mu\text{mol/liter}$ and in two sets of buffers: 25 mmol/liter Hepes pH 7.6, 50 mmol/liter KCl, and 5 mmol/liter MgCl_2 ; and 25 mmol/liter Tris-HCl, 150–500 mmol/liter NaCl, and 1 mmol/liter β -mercaptoethanol. The scan rate varied from 0.5 to 1.5 $^\circ\text{C/min}$, and the temperature measurement range was from 10 to 100 $^\circ\text{C}$. Curves are presented for the first (solid line) and second (dotted line) temperature-induced unfolding of GrpE, showing that the unfolding is irreversible (even if the first scan was up to only 57 $^\circ\text{C}$). The protein started to unfold at about 50 $^\circ\text{C}$. There are at least two unfolding events, one has a midpoint at 57 $^\circ\text{C}$, and the other is probably aggregation that occurs above about 62 $^\circ\text{C}$. The human GrpE unfolding profile shown here is independent of salt concentration from 50 to 500 mM, of concentration from 7 to 20 $\mu\text{mol/liter}$, and of scan rate.

perimentally determined scattering curves and those calculated from DAM models. A “looseness” penalty term is added to the minimization function ensuring that the procedure yields a compact and interconnected model. The low resolution structure model was derived from the experimental data assuming a 2-fold symmetry for the dimer. The symmetry restriction resulted in a significant reduction in the number of model free parameters. The starting volume for human GrpE dimer corresponds to a filled sphere with a diameter $D = 170 \text{ \AA}$ containing $n_{\text{DAM}} = 4,196$ dummy atoms with a radius $r_a = 4.75 \text{ \AA}$. Twenty independent *ab initio* simulations were performed. At the end of the DAMMIN numerical process that tries to diminish the discrepancy between the experimental and modeled curves, 182 ± 1 dummy atoms remained in the final model of the human GrpE dimer at low resolution. The obtained DAMs are presented in Fig. 8a. DAM-derived structure parameters agree with those directly determined from the SAXS curve as seen in Table I. To verify the uniqueness of the resulting shape, several independent restorations were performed with different starting conditions yielding similar results. The analysis of the SAXS curve corresponding to the solution with rather high human GrpE dimer concentration (19.4 mg/ml) gave the best resolution of the final model as 15 $\text{\AA}$. This resolution does not permit unambiguous determination of the spatial positions of secondary structure elements, but it allows the overall shape of the molecule and relative positions of individual protein domains to be obtained.

The crystal structure of *E. coli* GrpE (17) was used for modeling the human GrpE with the SWISS-MODEL, PredictProtein, and HyperChem programs (see “Experimental Procedures”). These programs yielded the estimates of the global structure of human GrpE shown in Fig. 8, b–e. Fig. 8b shows the human GrpE dimer structure modeled from the proximal monomer of the *E. coli* GrpE crystallographic structure (17) fitted into the DAM model. The fit of the β -sheet domains is not satisfactory. In the GrpE dimer, the proximal monomer is

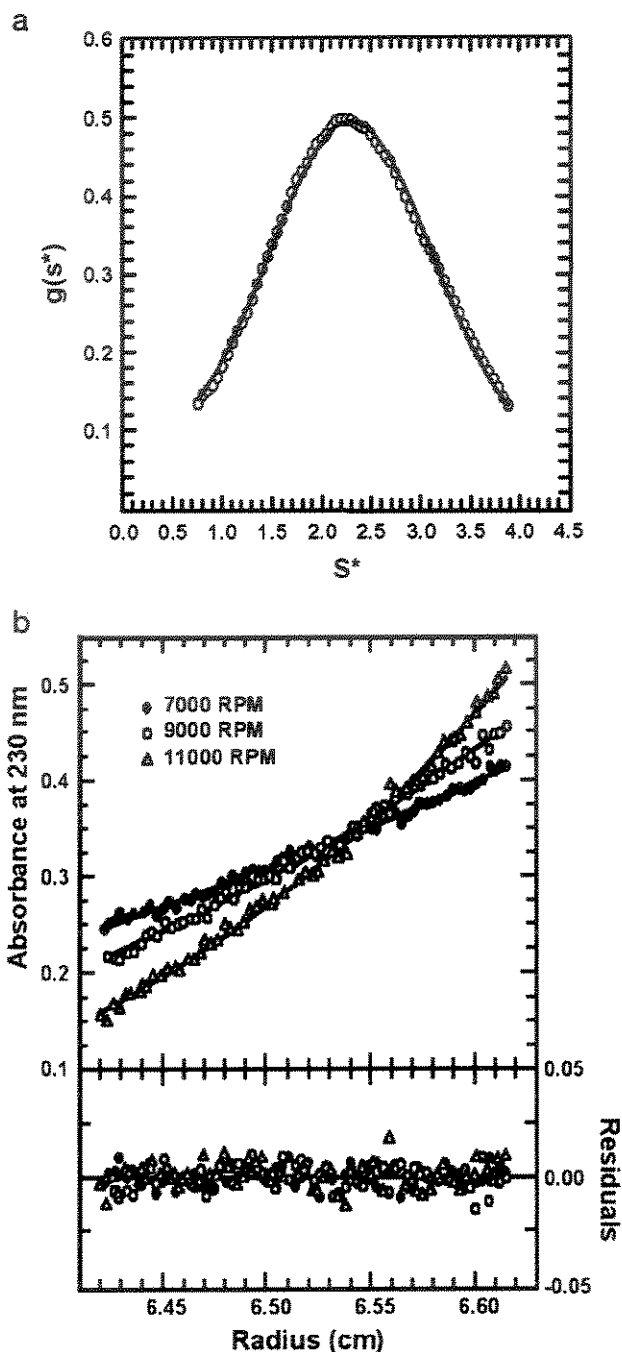


FIG. 5. Analytical ultracentrifugation experiments. *a*, sedimentation velocity. Sedimentation velocity experiments were carried out at 20 $^\circ\text{C}$, 40,000 rpm (AN-60Ti rotor), and with scan data acquisition at 230 nm. The figure shows the experiment with 200 $\mu\text{g/ml}$ GrpE in 25 mmol/liter Tris-HCl buffer, pH 7.5, 150 mmol/liter NaCl, and 0.5 mmol/liter DTT. The $g(s^*)$ distribution was fitted using the ORIGIN (MicroCal Software) with a Gaussian giving a sedimentation coefficient of about 2.25 Svedberg ($s_{20,w} = 2.34$). The plot also shows that GrpE is homogeneous. *b*, sedimentation equilibrium. Sedimentation equilibrium experiments were made at 7,000, 9,000, and 11,000 rpm (AN-60Ti rotor) with scan data acquisition at 230 nm. The best fit of experimental data for 100 $\mu\text{g/ml}$ GrpE at 7,000, 9,000 and 11,000 rpm at 20 $^\circ\text{C}$ is shown (for details, see “Experimental Procedures”). The random distribution of the residuals (bottom panel) is a good indication that the fit is correct. The fit agrees with a dimer structure of human GrpE ($43 \pm 1 \text{ kDa}$).

bound to DnaK-NBD, whereas the distal monomer is not (17). Analyses of Fig. 8, c–e, where the human GrpE dimer structure modeled from the distal monomer of *E. coli* GrpE is fitted into

TABLE I
Mathematical, structural and dynamic parameters of human GrpE in solution obtained from different experimental procedures and modeling methods

Protein parameters generated by SAXS experiments are compared with the DAM model parameters. The hydrodynamic data were from analytical ultracentrifugation experiments and also generated by HydroPro with the DAM model as template and assuming the protein is a dimer with a molecular mass of 43 kDa. The partial specific volume and solvent viscosity and density were determined by Sednterp software. For details, "Experimental Procedures." D , translational diffusion coefficient; R_g , radius of gyration of the scattering object; s^* , apparent sedimentation coefficient ($s_{20,w}$ is 2.34); P , Perrin factor or frictional ratio. Errors are all less than 2%.

Mathematical, structural, and dynamic parameters	Methods			
	SAXS ^a	DAM ^d	AUC ^c	HydroPro
Free parameters	20 ^d	2098		
Discrepancy X		1.0		
Resolution (nm)	1.4	2.8		
Molecular mass (kDa)	38.0		42.5	
R_g (Å)	45	46		45
D_{max} (Å)	155	160		156
D ($\times 10^{-7}$ cm ² /s)			4.7 ^c	4.8
s^* (Svedberg)			2.25	2.20
P (f/f_0)				1.7

^a Calculated from experimental data.

^b DAM parameters of the dummy atoms model averaged over 20 models.

^c Equilibration sedimentation and sedimentation velocity experiments.

^d Number of Shannon channels given by $D_{max} - q_{max}/\Delta T$.

^e Calculated by the second moment method.

the DAM model, indicate that the β -domains and the α -helical tail fit well. The comparison of Fig. 8, *b* and *c*, suggests that, in solution, human GrpE is in the distal monomer conformation. Fig. 8*f* shows the superposition of the *E. coli* GrpE dimer structure obtained by x-ray diffraction (17) onto the low resolution DAM model for the human GrpE. The proximal monomer (*right*), which binds to DnaK-NBD, does not exhibit a satisfactory fit, whereas the distal monomer (*left*) yields a good fit. Our model of the region corresponding to the first 33 residues of GrpE, deleted in the crystal structure (17), took into account the α -helical secondary structure prediction and the best fit in the envelope (Fig. 8, *c* and *d*). To evaluate the reliability of the structure obtained from the SAXS results, the HydroPro program was used to calculate the hydrodynamic parameters of the DAM model, yielding values very similar to those determined experimentally (Table I). Such results provide additional support for the low resolution structure proposed here for human GrpE in solution ensuring that it represents a good approximation of the real conformation of the protein.

DISCUSSION

GrpE plays an important function in protein folding. It forms a stable complex with DnaK (Hsp70) which dissociates in the presence of ATP (46), helping to fold nascent proteins (13). GrpE is essential for bacterial viability (14), lowers the amount of DnaK required for function (47), and acts as the nucleotide exchange factor for the Hsp70 DnaK subfamily (12). Human mitochondrial GrpE was cloned in its mature form, without the mitochondrial peptide signal. This recombinant protein was produced in large scale and was purified as a soluble molecule. The purified protein is folded and presents a high amount of α -helices as a component of its secondary structure. The shape of the human GrpE CD spectra and the calculated secondary structure are similar to *E. coli* GrpE (17). Although no detailed information about the tertiary structure is given by far UV CD spectra analysis, the secondary structure information itself is an indication that human and *E. coli* GrpE share at least some structural homology. However, the thermal unfolding profile of

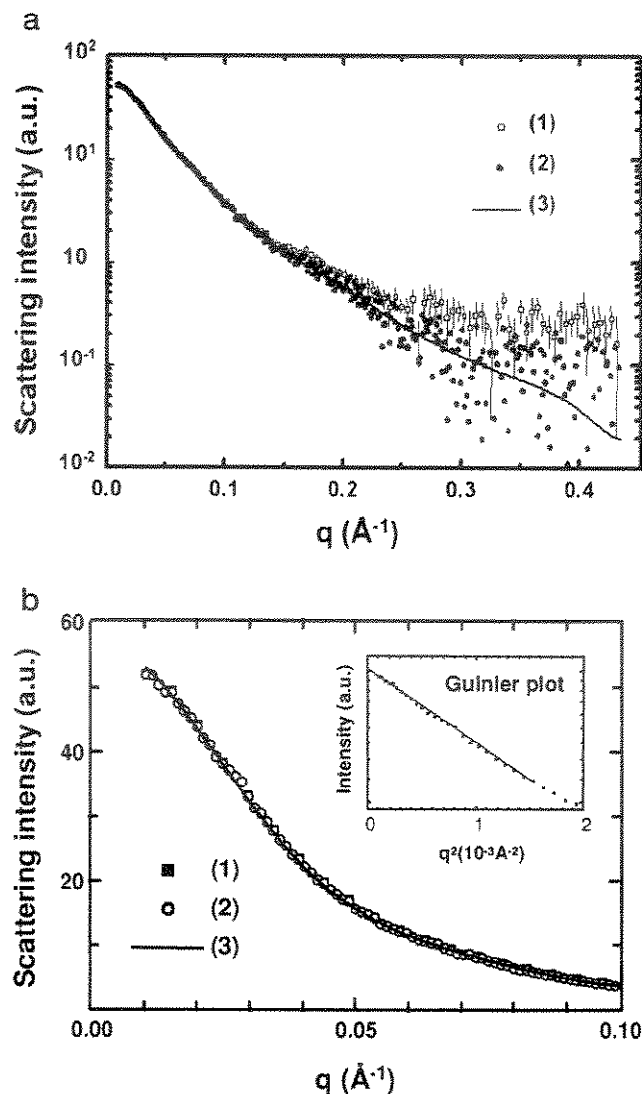


FIG. 6. Experimental SAXS curves of human GrpE in solution. *a*, plot of $\log I$ versus q (logarithmic scale) focusing on the features of SAXS curves at high q . *b*, plot of I versus q (linear scale) focusing on the features at small q . Inset, Guinier plot ($\log I$ versus q^2) within the small q range. 1, experimental curve; 2, experimental curve after subtraction of a constant value as described under "Results"; 3, scattering intensity computed from the DAM model (DAMMIN). Curves 1 and 2 are indistinguishable.

human GrpE measured either by CD or DSC differs from prokaryotic GrpE (18, 19), showing that these proteins differ in stability. The midpoint of the first thermal transition of human GrpE is at about 57 °C and is only partially reversible even if it is heated only to 50 °C. The first thermal transition of *E. coli* GrpE is at about 50 °C and is reversible if the protein is heated only up to the end of the transition at 60 °C (19). The first thermal transition in *E. coli* GrpE has been suggested to be the unfolding of the N terminus (18), a region with lower similarity to human GrpE than the C terminus (Fig. 1). Several works suggest that the N and the C termini of GrpE have different functions. The N terminus apparently acts as a temperature sensor for DnaK-DnaJ-GrpE function (18, 19, 48). We believe that the N-terminal region in human GrpE is probably more stable than in *E. coli* GrpE, thus shifting the thermal transition of this region to a higher temperature. This causes the N- and C-terminal regions to unfold simultaneously and thus not be able to refold. *E. coli* GrpE thermal unfolding is reversible

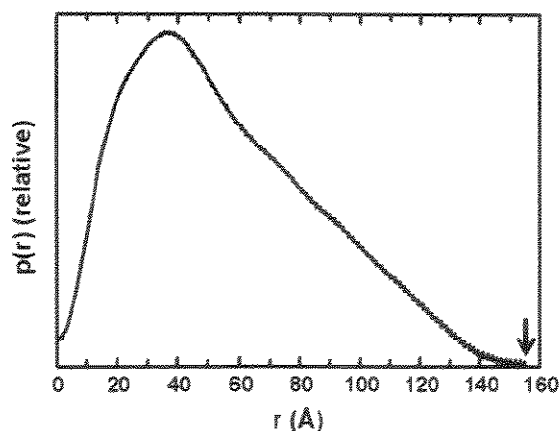


FIG. 7. Distance distribution function. Pair distance distribution, $p(r)$, of human GrpE showing the typical asymmetry that is expected for elongated (prolate) scattering objects. The $D_{\max} = 155$ Å, is indicated by an arrow.

when it is heated above the first thermal transition, where only the N-terminal domain is unfolded (18). Unfortunately, the thermal induced unfolding of human GrpE also results in aggregation, which prevents a more precise thermodynamic analysis.

The analysis of the quaternary structure of human GrpE strongly indicates that this protein exists as a dimer. Both sedimentation velocity and equilibrium sedimentation measurements of human GrpE in solution point to a protein with a molecular mass of 42–43 kDa at all concentrations tested. This molecular mass is about twice the molecular mass of the human GrpE monomer. This result is in good agreement with the analytical ultracentrifugation data generated for *E. coli* GrpE (16). The analysis of the crystal structure of *E. coli* GrpE confirms that this protein is a dimer with an elongated shape. Our hydrodynamic analysis of human GrpE also agrees with an elongated shape for this protein, and our analytical ultracentrifugation analysis leads to a sedimentation coefficient compatible with a nonglobular shape for human GrpE. SAXS data were used to generate an *ab initio* model (DAM) of human GrpE (see discussion below) which resulted in an elongated dimer model for GrpE. Together, the results obtained with analytical ultracentrifugation and SAXS data demonstrate that human GrpE has a high friction coefficient (f) for a protein of 42–43 kDa, and consequently a high translational friction ratio (P). The hydrodynamic parameters computed from this model are in good agreement with those determined experimentally, which gives confidence in both our results and the structural model presented here. The hydrodynamic analysis of human GrpE indicates that this protein has structural similarity to *E. coli* GrpE.

The elongated shape of GrpE is the most likely reason for the apparently much too high molecular mass calculated from molecular exclusion chromatography experiments (16). Several works have indicated that the elongated shape of GrpE serves the purpose of allowing its N terminus to also interact with DnaK-SBD. The deletion of the first 33 amino acid residues abolishes the release of substrate by DnaK (17), the mutation of residue Glu-53 results in a temperature-sensitive phenotype for lambda replication (49), and the GrpE_{1–89} fragment affects the release of peptide by DnaK but not the exchange rate of nucleotide (48). These results suggest that GrpE reduces DnaK affinity for substrate by interacting with SBD. Although the question of whether the GrpE N terminus binds to DnaK or not is still a matter of discussion, our data indicate a clearly elongated shape of human GrpE, the dimer being large enough and

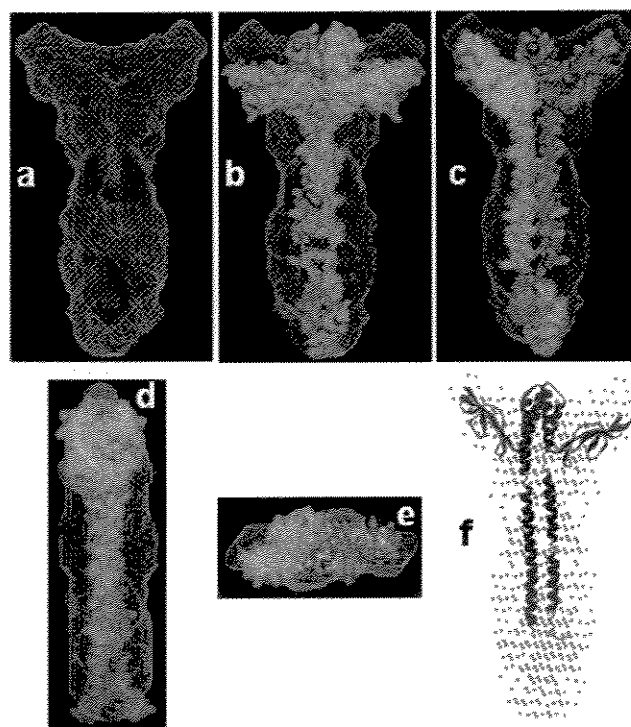


FIG. 8. Low resolution DAM models derived from the results of the SAXS experiment and molecular modeling for human GrpE. *a*, front view of human GrpE DAM model showing the elongated leaf-like shape. Grid spacing is 20 Å. *b*, front view of human GrpE dimer structure modeled from proximal monomer of *E. coli* GrpE crystallographic structure (17) fitted into the DAM model (for details, see “Experimental Procedures”). In the GrpE dimer, the proximal monomer is bound to DnaK-NBD, whereas the distal monomer is not. This figure indicates that the fit of the β -sheet domains is not satisfactory. *c*, front; *d*, side; and *e*, top view of human GrpE dimer structure modeled from the distal monomer of *E. coli* GrpE fitted into the DAM model (for details, see “Experimental Procedures”). The β -domains and the α -helical tail are well fit by the model. Comparison of (*b*) and (*c*) indicates that, in solution, human GrpE is in the distal monomer conformation. These models suggest that the binding of DnaK moves the β -domain away from the helical axis and that GrpE is an asymmetric dimer when not bound to DnaK. *f*, superposition of the *E. coli* GrpE dimer structure solved by single crystal x-ray diffraction (17) into the low resolution DAM model for the human GrpE. The proximal monomer (*right*), which binds to DnaK-NBD (17), does not yield a satisfactory fit, whereas the distal monomer (*left*) evidences a satisfactory fit. The model was displayed with the programs MASSHA and PYMOL (pymol.sourceforge.net).

sufficiently flexible to accomplish this function.

The high resolution structure of *E. coli* GrpE bound to NBD of DnaK shows that this protein is a dimer with a cruciform format (17). The C terminus is composed mainly of β -sheet in a rounded form, and the N terminus is composed of elongated α -helices that almost coil around each other. GrpE interacts with DnaK in 2:1 stoichiometry (16), and the crystallographic data (17) show that this interaction is asymmetric: only one GrpE monomer interacting with NBD of DnaK. The overall structure shown herein for human GrpE in solution also has a cruciform format and fits well with the *E. coli* GrpE high resolution structure (17). Together, all of the results indicate that the human GrpE structure is quite similar to the *E. coli* homolog. The fact that human and *E. coli* GrpE have similar global structures is expected because these proteins share about 53% similarity. Conservation of the amino acid sequence in this family suggests that the structure is also likely to be conserved.

To test our bead model we generated two possible human GrpE structures from each monomer of *E. coli* GrpE. The monomer of *E. coli* GrpE that is bound to DnaK has the C-

terminal domain more "open," compared with the unbound monomer (Fig. 8f) (17). The human GrpE structure generated using the bound monomer as a template does not fit well to the bead model generated from SAXS results (Fig. 8b), whereas the structure generated using the unbound monomer as a template exhibits a much better fit (compare Fig. 8, b and c). This suggests that in solution, when not bound to DnaK, GrpE is a symmetric dimer with each monomer in a "closed" conformation. The analysis of the *E. coli* GrpE structure (17) shows that the monomer and the dimer have perpendicular asymmetry, the C terminus has a different shape than the N terminus. The crystal structure of *E. coli* GrpE bound to DnaK (17) also shows an axial asymmetry because the C terminus of one of the monomers bends away from the axis (this monomer is known as the proximal monomer). The binding of DnaK then causes the C-terminal domain of the proximal monomer to open, probably to expose residues that will bind to DnaK. It is interesting to note that Harrison *et al.* (17) showed that part of the surface between the β -domain and the four-helix bundle domain makes contact with DnaK. This part is more accessible in an open conformation. The crystal structure resolution of GrpE was done with a protein deleted of the first 33 residues, opening the question of whether this region has a compact or loose structure. Although our low resolution structure cannot define the degree of compactness in this region, it indicates that this region is not totally loose.

The superimposed models display a consistent overall structure that justifies the *ab initio* modeling. It should be pointed out that, although the derived structure model generates a SAXS intensity curve in very good agreement with the experimental results, this does not in general guarantee that the model is unique. Our biophysical characterization of human GrpE in solution, using several experimental techniques, clearly indicates that the protein is in a dimeric conformation with an elongated overall shape. The low resolution structure generated from our SAXS data also suggests that GrpE is a symmetric dimer when not bound to DnaK.

Acknowledgments—We thank the technical staff of BFM and SAS groups at LNLs for valuable assistance.

REFERENCES

- Martin, J., and Hartl, F. U. (1997) *Curr. Opin. Struct. Biol.* **7**, 41–45
- Fink, A. L. (1999) *Physiol. Rev.* **79**, 425–449
- Friedman, D. I., Olson, E. R., Georgopoulos, C., Tilly, K., Herskowitz, I., and Banuett, F. (1984) *Microbiol. Rev.* **48**, 299–325
- Skowrya, D., Georgopoulos, C., and Zylicz, M. (1990) *Cell* **62**, 939–944
- Gething, M. J., and Sambrook, J. (1992) *Nature* **355**, 33–45
- Schroder, H., Langer, T., Hartl, F. U., and Bukau, B. (1993) *EMBO J.* **12**, 4137–4144
- Mayer, M. P., Brehmer, D., Gässler, C. S., and Bukau, B. (2001) *Adv. Protein Chem.* **59**, 1–44
- Flaherty, K. M., DeLuca-Flaherty, C., and McKay, D. B. (1990) *Nature* **346**, 623–628
- Pelham, H. R. (1986) *Cell* **46**, 959–961
- Johnson, C., Chandrasekhar, G. N., and Georgopoulos, C. (1989) *J. Bacteriol.* **171**, 1590–1596
- Liberek, K., Marszalek, J., Ang, D., Georgopoulos, C., and Zylicz, M. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 2874–2878
- Szabo, A., Langer, T., Schroder, H., Flanagan, J., Bukau, B., and Hartl, F. U. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 10345–10349
- Ang, D., Chandrasekhar, G. N., Zylicz, M., and Georgopoulos, C. (1986) *J. Bacteriol.* **167**, 25–29
- Langer, T., Lu, C., Echols, H., Flanagan, J., Hayer, M. K., and Hartl, F. U. (1992) *Nature* **356**, 683–689
- Osipiuk, J., Georgopoulos, C., and Zylicz, M. (1993) *J. Biol. Chem.* **268**, 4821–4827
- Schönfeld, H. J., Schmidt, D., Schröder, H., and Bukau, B. (1995) *J. Biol. Chem.* **270**, 2183–2189
- Harrison, C. J., Hayer-Hartl, M., Di Liberto, M., Hartl, F. U., and Kuriyan, J. (1997) *Science* **276**, 431–435
- Gelinas, A. D., Langsetmo, K., Toth, J., Bethoney, K. A., Stafford, W. F., and Harrison, C. J. (2002) *J. Mol. Biol.* **323**, 131–142
- Grimshaw, J. P., Jelesarov, I., Schönfeld, H. J., and Christen, P. (2001) *J. Biol. Chem.* **276**, 6098–6104
- Naylor, D. J., Stines, A. P., Hoogenraad, N. J., and Hoj, P. B. (1998) *J. Biol. Chem.* **273**, 21169–21177
- Schlicher, T., and Soll, J. (1997) *Plant Mol. Biol.* **33**, 181–185
- Borges, J. C., Peroto, M. C., and Ramos, C. H. I. (2001) *Gen. Mol. Biol.* **24**, 85–92
- Choglay, A. A., Chapple, J. P., Blatch, G. L., and Cheetham, M. E. (2001) *Gene (Amst.)* **267**, 125–134
- Laemmli, U. K. (1970) *Nature* **227**, 680–685
- Edelhock, H. (1967) *Biochemistry* **6**, 1948–1954
- Gill, S. C., and von Hippel, P. H. (1989) *Anal. Biochem.* **182**, 319–326
- van Holde, K. E., and Weisheit, W. O. (1978) *Biopolymers* **17**, 1397–1403
- Goldberg, R. J. (1953) *J. Phys. Chem.* **57**, 194–202
- Stafford, W. F. (1994) *Methods Enzymol.* **240**, 478–501
- Johnson, M. L., Correia, J. J., Yphantis, D. A., and Halvorson, H. R. (1981) *Biophys. J.* **36**, 575–588
- Kellermann, G., Vicentin, F., Tamura, E., Rocha, M., Tolentino, H., Barbosa, A., Craievich, A., and Torriani, I. (1997) *J. Appl. Crystallogr.* **30**, 880–883
- Svergun, D. I. (1992) *J. Appl. Crystallogr.* **25**, 495–503
- Svergun, D. I., and Stuhmann, H. B. (1991) *Acta Crystallogr. Sect. A* **47**, 736–744
- Porod, G. (1982) in *Small-angle X-ray Scattering* (Glatter, O., and Kratky, O., eds) pp. 17–51, Academic Press, London
- Svergun, D. I. (1999) *Biophys. J.* **76**, 2879–2886
- Konarev, P. V., Petoukhov, M. V., and Svergun, D. I. (2001) *J. Appl. Crystallogr.* **34**, 527–532
- Svergun, D., Barberato, C., and Koch, M. H. (1995) *J. Appl. Crystallogr.* **28**, 768–773
- Kozin, M. B., and Svergun, D. I. (2001) *J. Appl. Crystallogr.* **34**, 33–41
- Garcia de la Torre, J., Huertas, M. L., and Carrasco, B. (2000) *Biophys. J.* **78**, 719–730
- Garcia de la Torre, J., Carrasco, B., and Harding, S. E. (1997) *Eur. Biophys. J.* **25**, 361–372
- Guex, N., and Peitsch, M. C. (1997) *Electrophoresis* **18**, 2714–2723
- Peitsch, M. C. (1996) *Biochem. Soc. Trans.* **24**, 274–279
- Peitsch, M. C. (1995) *Bio/Technology* **13**, 658–660
- Baldi, P., Brunak, S., Frasconi, P., Pollastri, G., and Soda, G. (1999) *Bioinformatics* **15**, 937–946
- Feigin, L. A., and Svergun, D. I. (1987) *Structure Analysis by Small-angle X-ray and Neutron Scattering*, pp. 83–87, Plenum Press, New York
- Zylicz, M., Ang, D., and Georgopoulos, C. (1987) *J. Biol. Chem.* **262**, 17437–17442
- Zylicz, M., Ang, D., Liberek, K., and Georgopoulos, C. (1989) *EMBO J.* **8**, 1601–1608
- Mehl, A. F., Heskett, L. D., and Neal, K. M. (2001) *Biochem. Biophys. Res. Commun.* **282**, 562–569
- Wu, B., Ang, D., Snively, M., and Georgopoulos, C. (1994) *J. Bacteriol.* **176**, 6965–6973

9 CURRICULUM VITAE

Júlio César Borges (Borges, J.C.)

Nascido em 30 de Junho de 1976, em Santo Antônio do Monte, MG, Brasil.

Formação acadêmica

1- Doutor em Biologia Funcional e Molecular ênfase em Bioquímica, Instituto de Biologia, UNICAMP, Campinas, SP, Brasil.

Título da Tese: “ESTUDOS ESTRUTURAIS DO SISTEMA CHAPERONE MOLECULAR HSP70 HUMANO”, sob a orientação do Prof. Dr. Carlos H.I. Ramos, em desenvolvimento nas dependências do Centro de Biologia Molecular Estrutural do Laboratório Nacional de Luz Síncrotron – CeBiME/LNLS.

Período: Agosto de 2000 a Outubro de 2004.

Auxílio Financeiro: Bolsa de Doutorado da FAPESP.

2- Bacharel em Farmácia, Faculdade de Farmácia, UFMG, Belo Horizonte, MG, Brasil.

Período: Fevereiro de 1996 a Janeiro de 2000.

Outras Atividades: Bolsista de Iniciação científica do programa PIBIC/CNPq no Laboratório de Biofísica, Depto. de Fisiologia e Biofísica, ICB/UFMG, sob a orientação do Prof. Dr. Jorge Luiz Pesquero.

Áreas de interesse

Química de Macromoléculas – Proteínas;

Bioquímica – Estrutura e função de proteínas;

Biofísica – Enovelamento de proteínas e estabilidade;

Farmácia – Farmacologia e química.

Publicações Científicas

- Borges, J.C.**, Fischer, H., Craievich, A.F. & Ramos, C.H.I. (2004) Low-resolution structural study of two human Hsp40 chaperones in solution. DjA1 from subfamily A and DjB4 from subfamily B, have different quaternary structures. *J. Biol. Chem.*, (Submetido).
- Borges, J.C.** & Ramos, C.H.I. (2004) Protein folding assisted by chaperones. *Protein Pept. Lett.* (submetido).
- Cagliari, T.C., Tiroli, A.O., **Borges J.C.** & Ramos, C.H.I. (2004) Identification and expression pattern analysis of eucalyptus expressed sequencing tags (ESTs) encoding molecular chaperones. *Genet. Mol. Biol.* (Submetido).
- Borges, J.C.**, Fischer, H., Craievich, A.F., Hansen, L.D., & Ramos, C.H.I. (2003) Free human mitochondrial GrpE is a symmetric dimer in solution. *J. Biol. Chem.*, 278, 35337-35344.
- Borges, J.C.**, Silva, Jr., J.A., Gomes, M.A., Lomez, E.S.L., Leite, K.M., Bader, M., Pesquero, J.B. & Pesquero, J.L. (2003) Tonin in rat heart with experimental hypertrophy. *Am. J. Physiol. Heart Circ. Physiol.* 284, H2263-H2268.
- Borges, J.C.**, Peroto, M.C. & Ramos, C.H.I. (2001) Molecular Chaperone Genes in the SugarCane Expressed Sequence Database (SUCEST). *Genet. Mol. Biol.*, 24, 85-92.

Publicações em Anais de congressos e reuniões científicas

Publicou 14 trabalhos em Congressos e reuniões científicas, sendo 12 em forma de painel e duas em forma oral. Publicou 3 trabalhos de caráter científico em anais científicos.