Universidade Estadual de Campinas

Celisa Caldana Costa

# Caracterização da Proteína Tif34p e do Sub-complexo Tif34p/Tif35p do Fator de Tradução eIF3 de Saccharomyces cerevisiae

Dissertação apresentada ao Instituto de Biologia da Universidade Estadual de Campinas para obtenção do título de Mestre em Genética e Biologia Molecular - Área de Concentração Genética Animal e Evolução.

## **Orientador: Nilson Ivo Tonin Zanchin**

Campinas 2004

	Costa, Celisa Caldana
C823c	Caracterização da proteína Tif34p e do sub-complexo Tif34p/Tif35p do fator de
	tradução eIF3 de Saccharomyces cerevisiae / Celisa Caldana Costa
	Campinas, SP:[s.n.], 2004.
	Orientador: Nilson Ivo Tonin Zanchin
	Dissertação (mestrado) - Universidade Estadual de Campinas.
	Instituto de Biologia.
	1. Proteínas - Estrutura I. Zanchin, Nilson Ivo Tonin II. Universidade Estadual de
	Campinas. Instituto de Biologia. III. Título.

Campinas, 01 de julho de 2004.

### **BANCA EXAMINADORA:**

Prof. Dr. Nilson Ivo Tonin Zanchin (Orientador)

Profa. Dra. Anete Pereira de Souza

Profa. Dra. Beatriz Gomes Guimarães

Prof. Dr. Marcelo Menossi Teixeira

## AGRADECIMENTOS

Ao Prof. Dr. Nilson Ivo Tonin Zanchin pela orientação e paciência.

Ao Prof. Dr. Celso Eduardo Benedetti, Profa. Dra. Anete Pereira de Souza e Profa. Dra. Thelma de Aguiar Pertinhez pelas sugestões e críticas na banca prévia.

À Profa. Dra. Beatriz Gomes Guimarães e Profa. Dra. Anete Pereira de Souza pela participação na banca.

À Profa. Dra. Íris C. L. Torriani e ao Cristiano L. P. Oliveira pela colaboração nas medidas de SAXS.

Ao Dr. Sérgio Oyama Júnior pela colaboração na construção do modelo molecular.

À Profa. Dra. Heloísa S. S. de Araújo, da Universidade Federal de São Carlos, pelas tentativas de seqüenciamento N-terminal.

Ao Prof. Dr. Francisco Javier Medrano Martin pela ajuda nos ensaios de cristalização.

À Profa. Dra. Thelma de Aguiar Pertinhez pela ajuda nas interpretações de CD.

Ao Prof. Dr. Ryland F. Young da *Texas A&M University* por ter nos cedido a cepa BL21(DE3)*slyD*<sup>-</sup>.

Às técnicas Tereza Cristina Lima Silva e Adriana Cristina Alves pela assistência no laboratório.

Às minhas amigas do BCM: Patrícia, Sandra, Flávia, Thaís, Beatriz, Juliana e ao Marcelo.

À Patrícia Roveri Marques e à Vanessa Cappovila pelo auxílio e ajuda no laboratório.

À todos os amigos do Laboratório de Biologia Molecular.

À toda estrutura oferecida pelo LNLS.

Às minhas irmãs Daniela e Camila pela amizade e apoio.

Ao meu marido Cesar pelo incentivo, compreensão e paciência.

Aos meus pais, Ana e Márcio, por me ensinar a lutar e acreditar nos meus ideais sempre.

À FAPESP pelo auxílio financeiro concedido.

# ÍNDICE

Resumo	
Abstract	
1. Introdução	
1.1 - Iniciação da Tradução em Eucariotos	
1.2 - Estudo Estrutural dos Fatores de Iniciação da Tradução de Eucariotos	09
1.3 - Fator de Iniciação da Tradução 3	10
1.4 - Família de Proteínas Formadas por Domínios WD (WD-repeat)	16
1.5 - Técnicas para a Caracterização Estrutural de Proteínas	20
2. Objetivos	29
3. Materiais e Métodos	
3.1 - Meios de Cultura e Manipulação Genética de Bactérias	30
3.2 - Plasmídeos e Cepas	32
3.3 - Amplificação de DNA por PCR	34
3.4 - Análise de DNA	35
3.5 - Ensaios de Expressão de Proteínas em E. coli	38
3.6 - Purificação de Proteínas	
3.7 - Digestão da Poli-Histidina com Trombina	42
3.8 - Concentração e Quantificação de Proteínas	
3.9 - Ensaios de Proteólise Limitada	43
3.10 - Espectroscopia de Dicroísmo Circular	44
3.11 - Espalhamento de Luz Dinâmico	45
3.12 - Ensaio de Cristalização de Proteínas	46
3.13 - Modelagem Molecular	46
3.14 - Espalhamento de Raios-X a Baixo Ângulo	47
4. Resultados	
4.1 - Vetores de Expressão	49
4.2 - Expressão e Purificação das Proteínas HisTif34p e Tif34p	
4.3 - Clivagem da Poli-Histidina com Trombina	
4.4 - Expressão e Purificação das Proteínas HisTif35p e HisTif35N	

4.5 - Co-expressão e Isolamento dos Complexos	
4.6 - Caracterização por Espalhamento de Luz Dinâmico	
4.7 - Caracterização por Proteólise Limitada	60
4.8 - Caracterização Estrutural por Dicroísmo Circular	61
4.9 - Ensaios de Cristalização de Proteínas	66
4.10 - Modelagem Molecular	66
4.11 - Caracterização Estrutural por Espalhamento de Raios-X a Baixo Ângulo	68
5. Discussão	77
6. Considerações Finais	
7. Conclusões	92
8. Referências Bibliográficas	94
9. Anexos	108

# ÍNDICE DE FIGURAS E TABELAS

FIGURAS	
Figura 1.1 - Modelo de iniciação da síntese de proteínas em eucariotos	06
Figura 1.2 - Modelo esquemático do complexo de iniciação eIF3 de S. cerevisiae	13
Figura 1.3 - Padrão geral da seqüência de aminoácidos do domínio WD	17
Figura 1.4 - Espectros característicos de dicroísmo circular	24
Figura 4.1 - Gel SDS-PAGE da expressão da HisTif34p e Tif34p	50
Figura 4.2 - Gel SDS-PAGE da purificação da HisTif34p	51
Figura 4.3 - Gel SDS-PAGE da expressão da HisTif35p	53
Figura 4.4 - Gel SDS-PAGE da purificação da HisTif35p	53
Figura 4.5 - Gel SDS-PAGE da expressão da HisTif35N	54
Figura 4.6 - Gel SDS-PAGE da purificação da HisTif35N	55
Figura 4.7 - Gel SDS-PAGE da purificação dos complexos por cromatografia de	
afinidade	56
Figura 4.8 - Perfil cromatográfico dos complexos fracionados por gel filtração	57
Figura 4.9 - Gel SDS-PAGE da purificação dos complexos por cromatografia de gel	
filtração	57
Figura 4.10 - Gel SDS-PAGE da co-sedimentação do complexo Tif34p/HisTif35N	58
Figura 4.11 - Gráficos de DLS da HisTif34p e Tif34p/HisTif35N	59
<b>Figura 4.12</b> - Gel SDS-PAGE da proteólise limitada da HisTif34p e Tif34p/HisTif35N .	60
Figura 4.13 - Gel SDS-PAGE da cromatografia de afinidade da HisTif34p digerida com	
tripsina	61
Figura 4.14 - Espectro de CD da HisTif34p	62
Figura 4.15 - Espectros de CD da HisTif35p e HisTif35N	63
Figura 4.16 - Sobreposição dos espectros de CD das proteínas separadas e do complexo	
HisTif34p/HisTif35p reconstituído com dados teóricos	64
Figura 4.17 - Sobreposição dos espectros de CD obtidos para o complexo	
HisTif34p/HisTif35p reconstituído e Tif34p/HisTif35p co-purificado	65
Figura 4.18 - Alinhamento das seqüências de aminoácidos da HisTif34p e Cdc4	
<b>Figura 4.19</b> - Traços dos $C_{\alpha}$ do modelo molecular da HisTif34p	67

Figura 4.20 - Modelo molecular da Tif34p	68	
Figura 4.21 - Gel SDS-PAGE da HisTif34p e Tif34p/HisTif35N utilizadas nas medidas		
de SAXS		
<b>Figura 4.22</b> - Curvas de SAXS da HisTif34p e Tif34p/HisTif35N	70	
Figura 4.23 - Dados de SAXS da HisTif34p obtidos pelo programa GNOM	71	
Figura 4.24 - Dados de SAXS da Tif34p/HisTif35N obtidos pelo programa GNOM	72	
Figura 4.25 - Dados de SAXS da HisTif34p obtidos pelo programa Crysol	72	
Figura 4.26 - Envelope conformacional da HisTif34p	73	
Figura 4.27 - Sobreposição do modelo molecular da Tif34p com o modelo ab initio da		
HisTif34p	73	
Figura 4.28 - Dados de SAXS da Tif34p/HisTif35N obtidos pelo programa Crysol	74	
<b>Figura 4.29</b> - Envelope conformacional do complexo Tif34p/HisTif35N	75	
Figura 4.30 - Sobreposição do modelo molecular da Tif34p com o modelo ab initio do		
complexo Tif34p/HisTif35N		
TABELAS		
Tabela 3.1 - Plasmídeos de E. coli	32	
Tabela 3.2 - Cepas de E. coli		
Tabela 3.3 - Oligonucleotídeos		
Tabela 4.1 - Distâncias do detector-amostra e concentrações das proteínas utilizadas nas		
medidas de SAXS		

# ABREVIATURAS

APS	Persulfato de amônio
CD	Dicroísmo circular
DLS	Espalhamento de luz dinâmico
DNA	Ácido desoxirribonucléico
DTT	Ditiotreitol
EDTA	Ácido tetra-acético dissódico (etilenodinitrilo)
eIF	Fator de iniciação de eucariotos
IF	Fator de iniciação
IPTG	Isopropil-β-D-tiogalactopiranosídeo
IRES	Sítio de ligação interna do ribossomo
LB	Luria Bertani
Met-tRNA <sub>i</sub>	Metionil-tRNA iniciador
mRNA	RNA mensageiro
Ni-NTA	Ácido níquel-nitrilotriacético
nt	Nucleotídeo
PCR	Reação de polimerase em cadeia
PDB	Protein Data Bank
PEG	Polietileno-glicol
Rg	Raio de giro
RNA	Ácido ribonucléico
rpm	Rotações por minuto
RRM	Sítio de reconhecimento de RNA
SAXS	Espalhamento de raios-X a baixo ângulo
SDS	Dodecil-sulfato de sódio
SDS-PAGE	SDS polyacrylamide gel electrophoresis
TIF	fator de iniciação da tradução
tRNA	RNA transportador
Tween 20	Monolaurato de polioxietileno (20) sorbitano
UTR	Região não traduzida

### **RESUMO**

Proteínas pertencentes à família WD-repeat estão envolvidas em uma variedade de mecanismos bioquímicos e processos celulares. Os membros desta família são formados por 4-16 cópias do domínio WD, o qual apresenta uma característica comum que é uma seqüência de aproximadamente 40 aminoácidos que tipicamente começa com o dipeptídeo GH e termina com o dipeptídeo WD. A estrutura é formada por uma  $\beta$ -propeller, a qual é composta por folhas- $\beta$ antiparalelas com quatro fitas cada. Estes domínios WD são utilizados por algumas proteínas em ligações reversíveis e regulação de proteínas alvo específicas. O fator de iniciação da tradução eIF3 é um complexo multiprotéico contendo uma proteína WD-repeat, chamada Tif34p em Saccharomyces cerevisiae e TRIP-1 em humanos. A função destas subunidades do eIF3 na síntese de proteínas está bem estabelecida, entretanto, elas também estão envolvidas em outros processos celulares. TRIP-1 interage especificamente com o receptor tipo II do TGF-β sendo fosforilada por ele. Tif34p é necessária para o controle do ciclo celular e tem um papel essencial na estabilidade do complexo eIF3. Embora tem sido proposto que a Tif34p é uma proteína necessária para a integridade do complexo, ela interage apenas com as subunidades Tif35p e Prt1p do eIF3 de levedura.

Neste trabalho, as proteínas recombinantes Tif34p, Tif35p e a uma forma truncada da Tif35p, denominada Tif35N, foram clonadas, expressas e purificadas para estudos estruturais. Foram construídos dois vetores para a expressão da Tif34p, um para expressar a proteína nativa e outro para expressar a proteína com poli-histidina (HisTif34p). A HisTif34p, Tif34p, HisTif35p, HisTif35N e os complexos Tif34p/HisTif35p e Tif34p/HisTif35N foram expressos em *Escherichia coli* BL21(DE3). Os extratos foram fracionados por cromatografia de afinidade,

1

seguida por troca-iônica. Os complexos foram posteriormente purificados em coluna Superdex 75 HR de gel filtração. A HisTif34p, HisTif35p, o complexo HisTif34p/HisTif35p reconstituído e o complexo Tif34p/HisTif35p co-purificado foram analisados por dicroísmo circular (CD). Nenhuma diferença significante na estrutura secundária foi observada quando o espectro de CD das proteínas isoladas foi comparado ao espectro do complexo reconstituído e do complexo copurificado. Análises de CD indicaram que a Tif34p é formada principalmente por folhas- $\beta$  e alças, o que é consistente com a estrutura *β-propeller* das proteínas da família *WD-repeat*. Proteínas com estrutura *β-propeller* geralmente são fortemente resistentes à ação de proteases. Reações de proteólise limitada da HisTif34p e do complexo Tif34p/HisTif35N geraram um fragmento de 30 kDa fortemente resistente à digestão com tripsina. Estes dados são consistentes com o modelo molecular da Tif34p obtido por homologia com a proteína Cdc4. De acordo com este modelo, a Tif34p é constituída por oito folhas- $\beta$  antiparalelas, as quais formam a estrutura *βpropeller*.

Foram realizadas medidas de espalhamento de raios-X a baixo ângulo (SAXS) da HisTif34p e Tif34p/HisTif35N. Os modelos *ab initio* revelaram que a HisTif34p possui uma região central que corresponde ao  $\beta$ -propeller e uma extensão que deve corresponder aos trinta e dois aminoácidos do N-terminal, o qual contém a poli-histidina e o sítio de reconhecimento da trombina. Esta região não faz parte da  $\beta$ -propeller e pode estar desenovelada em solução. O modelo *ab initio* da Tif34p/HisTif35N mostrou uma região de sobreposição com o modelo *ab initio* da HisTif34p, correspondendo a Tif34p. No modelo do complexo obtido em solução, a HisTif35N está localizada em uma extremidade da  $\beta$ -propeller. Concluindo, a estrutura de baixa resolução obtida por SAXS para a HisTif34p é consistente com a  $\beta$ -propeller encontrada em todas as proteínas *WD-repeat* caracterizadas até o momento. Este tipo de estrutura parece agir como um suporte para interações proteína-proteína e formação de complexos.

### ABSTRACT

WD-repeat proteins comprise a diverse family involved in a variety of biochemical mechanisms and cellular processes. Members of this family are formed by 4-16 copies of the WD motif, whose common feature is an amino acid stretch typically beginning with a GH pair and ending with a WD pair. The structure forms a  $\beta$ -propeller with several blades where each blade is composed of a four-stranded anti-parallel  $\beta$ -sheets. These motifs constructed from WD repeats are utilized by many proteins in the reversible binding and regulation of specific protein targets. The eukaryotic translation initiation factor eIF3 is a multi-subunit complex containing a WD-repeat protein, named Tif34p in *Saccharomyces cerevisiae* and TRIP-1 in humans. The role of these eIF3 subunits in protein synthesis is well established; however they are also involved in other cellular processes. TRIP-1 interacts specifically with and is phosphorylated by the transforming growth factor-  $\beta$  receptor type II. Tif34p is required for cell cycle control and plays an essential role in the stabilization of the eIF3 complex. Although it has been proposed that Tif34p is a scaffold protein required for eIF3 complex integrity, it interacts only with the Tif35p and Prt1p subunits of the yeast eIF3.

In this work we cloned, expressed and purified recombinant Tif34p, Tif35p and a truncated form of Tif35p for structural studies. Two vectors were constructed for Tif34p, one expressing the native protein (Tif34p) and another expressing a His-tagged protein (HisTif34p). HisTif34p, Tif34p, HisTif35p, HisTif35N and the complexes Tif34/HisTif35N and Tif34p/HisTif35p were expressed in *Escherichia coli* BL21(DE3). Extracts were fractioned by Ni-NTA affinity chromatography, followed by ion-exchange chromatography. Complexes were further purified on a Superdex 75 HR molecular exclusion column. HisTif34p, HisTif35p, the HisTif34p/HisTif35p reconstituted complex and the Tif34p/HisTif35p co-purified complex were analyzed by circular

dichroism (CD). No significant change in secondary structure was observed when CD spectra of isolated proteins are compared to the CD spectra of both the reconstituted complex and of the complex purified from cells co-expressing Tif34p and HisTif35p. CD analysis indicated that Tif34p is formed mainly by  $\beta$ -sheets and loops, which is consistent with the predicted  $\beta$ -propeller structure of the WD-repeat protein family.  $\beta$ -propeller proteins usually show a strong resistance to protease attack. Partial proteolysis of the Tif34p and the Tif34p/Tif35N complex produced a 30 kDa fragment fairly resistant to trypsin digestion. These data are consistent with a molecular model of Tif34p obtained by using the Cdc4 protein as template. According to this model, Tif34p comprises eight anti-parallels  $\beta$ -sheets which form a  $\beta$ -propeller structure.

Small angle X-ray scattering (SAXS) studies were performed with HisTif34p and the complex Tif34/HisTif35N. *Ab initio* modelling of the SAXS data revelaed that HisTif34p shows a central region that corresponds to a  $\beta$ -propeller and an extension that may correspond to the 32 amino acids of the N-terminal, which contains the histidine-tag and the thrombin recognition site. This region is not part of the  $\beta$ -propeller and may be unfolded in solution. *Ab initio* modelling of the Tif34/HisTif35N showed an overlapping region with the HisTif34p model, corresponding to Tif34p. In the model of the complex obtained in solution, HisTif35N is located in one side of the  $\beta$ -propeller. In conclusion, the low resolution structure obtained for Tif34p in solution by SAXS is consistent with the expected  $\beta$ -propeller found in all WD-containing protein structures determined so far. This type of structure appears to function as a support for protein-protein interaction and complex formation.

## 1. INTRODUÇÃO

### 1.1 - Iniciação da Tradução em Eucariotos

A iniciação da síntese de proteínas consiste no recrutamento do complexo ribossomo-tRNA iniciador para o códon de iniciação do mRNA. Em procariotos, três fatores peptídicos (IF1, IF2 e IF3) são necessários para a iniciação da tradução. Já em eucariotos, este é um processo complexo envolvendo interações proteína-RNA e proteína-proteína, o qual requer a participação coordenada de inúmeros fatores de iniciação da tradução (eIFs) (Pestova et al, 2001) para o correto funcionamento da maquinaria de síntese de proteínas. No mRNA, o códon utilizado como sítio de iniciação para a síntese de proteínas é precedido por uma região 5' não traduzida, na qual a extensão, a composição de nucleotídeos e a estrutura podem determinar a eficiência e o mecanismo pelo qual o mRNA é traduzido (Hershey e Merrick, 2000).

A iniciação da tradução da maioria dos mRNAs eucarióticos é dependente do recrutamento da subunidade ribossomal 40S através da extremidade 5' do mRNA num processo denominado de *scanning model* ou "modelo de varredura" (Kozak, 1989; Hershey e Merrick, 2000), o qual está representado na Figura 1.1.



Figura 1.1 - Modelo de iniciação da síntese de proteínas em eucariotos (Hershey e Merrick, 2000). 1: eIF1; 1A: eIF1A; 2: eIF2; 3: eIF3; 4A: eIF4A; 4B: eIF4B; 4E: eIF4E; 4G: eIF4G. 5: eIF5; m7G: 7-metil-guanosina.

Este modelo é baseado em evidências genéticas e bioquímicas, as quais têm mostrado que um complexo ternário é formado entre eIF2, GTP e o metionil-tRNA iniciador (Met-tRNA<sub>i</sub>), enquanto a subunidade ribossomal 40S livre é ligada ao eIF3 e eIF1A (Merrick e Hershey, 2000). A subunidade 60S livre similarmente se liga ao eIF6 monomérico. Juntos, eIF3 e eIF6 impedem a associação prematura das subunidades ribossomais 40S e 60S, respectivamente. O complexo ternário é transferido para eIF3/40S juntamente com eIF1 e eIF1A, para formar o complexo de pré-iniciação 43S. eIF3 pode agora se ligar ao complexo eIF4F, que está associado com a extremidade 5' do mRNA, deste modo unindo o ribossomo 40S ao mRNA e gerando o complexo de pré-iniciação 48S.

O fator eIF4F desempenha um papel central no recrutamento do complexo 43S pela extremidade 5' do mRNA. Este fator consiste em três componentes, eIF4G, eIF4E e eIF4A, sendo que o eIF4G se liga ao eIF3 e age como um adaptador para eIF4E e eIF4A. eIF4E reconhece especificamente o metil-GTP e se liga à estrutura 5' cap do mRNA; o eIF4A se liga a estruturas secundárias da região 5' não traduzida (UTR) e as desfaz com sua atividade de RNA helicase. Além disso, existe o fator eIF4B que pode estimular a atividade do eIF4A. A cauda poli(A) do mRNA interage com a proteína PABP [poli(A) binding protein] que também se liga ao eIF4G. No modelo de varredura, o complexo de pré-iniciação 48S se posiciona na extremidade 5' do mRNA e o percorre até encontrar o códon de iniciação AUG. Com o Met-tRNA<sub>i</sub> posicionado neste códon, o eIF5 interage com o complexo de pré-iniciação via eIF2 e eIF3. O eIF2 ligado a GTP é hidrolisado e eIF2-GDP é liberado. A hidrólise de um segundo GTP ligado ao fator de iniciação eIF5B é ativada pela associação da subunidade ribossomal 60S. Estas duas hidrólises sucessivas e a simultânea liberação do eIF3, são essenciais para a união da subunidade 60S à 40S. Um ribossomo funcional 80S é conseqüentemente formado, que em seguida entra na fase de elongação da cadeia polipeptídica. Na etapa de elongação, há a contínua adição de aminoácidos na extremidade C-terminal do novo polipeptídio e a fase de terminação compreende o

reconhecimento do códon de terminação pelo ribossomo, liberação da cadeia polipeptídica e a dissociação entre o ribossomo e o mRNA.

Um segundo mecanismo de iniciação da tradução tem sido proposto para os mRNAs que possuem região 5' não traduzida longa e com estruturas secundárias estáveis. Alguns grupos de vírus como os picornavírus e os poliovírus, têm propriedades que são incompatíveis com a iniciação da tradução dependente da estrutura *cap*. Em 1988, foi descoberto que os mRNAs de picornavírus são traduzidos por um mecanismo que permite os ribossomos iniciarem a tradução em regiões altamente estruturadas localizadas na região 5' UTR (Jang et al, 1988; Pelletier e Sonenberg, 1988). Estas regiões foram chamadas de sítios de ligação interna de ribossomos (*internal ribosome entry site*, IRES). Os IRES com suas diferentes estruturas têm revelado que a subunidade ribossômica 40S pode ser recrutada por alguns mecanismos diferentes e independentes da estrutura *cap*: no vírus EMCV (vírus da encefalomiocardite) a interação do eIF4G/eIF4A com o IRES, auxiliada pelo eIF2, eIF3 e eIF4B, é necessária para o recrutamento da subunidade 40S (Jang et al, 1988). O IRES do vírus da hepatite C (HCV) pode se ligar à subunidade 40S livre de eIFs, sem a helicase eIF4A e ainda recrutar o complexo ternário para a posição do tRNA<sub>i</sub> no sítio P ribossomal (Tsukiyama-Kohara et al, 1992).

Os fatores de tradução podem ser regulados em vários níveis. Entretanto, a regulação ocorre com mais eficiência e freqüência na iniciação da tradução. Dois fatores têm um papel crítico na regulação desta etapa da tradução: o eIF2 $\alpha$  e o eIF4E. A atividade do eIF2 $\alpha$  é regulada pela fosforilação de sua serina 51 catalisada por quatro proteínas-quinases específicas (HRI, PKR, PERK e Gcn2p), sendo que mutações no sítio de fosforilação ou inativação destas quinases, resultam no crescimento celular desordenado (Samuel, 1993). Um aumento do nível celular de eIF4E resulta na perda do controle do crescimento celular e na transformação celular malígna (Raugt e Gingras, 1999). Este fator também é regulado por fosforilação, sendo que sua atividade é aumentada quando o grupamento fosfato é covalentemente ligado à serina 209.

#### 1.2 - Estudo Estrutural dos Fatores de Iniciação da Tradução

Progressos significativos têm ocorrido nos estudos estruturais dos fatores de iniciação da tradução. Estes estudos estão revelando o uso repetido de grupos estruturais comuns, evidenciando a conservação do aparato de tradução e promovendo maiores conhecimentos sobre a função e regulação destes fatores (Sonenberg e Dever, 2003). A maioria dos eIFs teve apenas parte de sua estrutura resolvida, seja por cristalografia de raios-X ou por ressonância magnética nuclear. Os únicos fatores de iniciação que não tiveram sua estrutura determinada foram o eIF2B e o eIF3, os quais são formados por várias subunidades. Recentemente, uma das cinco subunidades do eIF2B (eIF2Bε) teve a estrutura cristalográfica de um fragmento catalítico resolvida (Boesen et al, 2004), a qual está sendo usada para estudos envolvendo a presença de mutações responsáveis pela leucoencefalopatia, uma doença genética rara de humanos (van der Knaap et al, 2002).

A estrutura dos três fatores de iniciação da tradução de procariotos (IF1, IF2 e IF3) foi determinada recentemente (Boelens e Gualerzi, 2002). A estrutura do eIF1 resolvida por RMN (Fletcher et al, 1999), assemelha-se com a estrutura do domínio C-terminal do IF3. Os fatores eIF1A e eIF5B são estruturalmente homólogos ao IF1 e IF2, respectivamente (Roll-Mecak et al, 2001). Comparações entre as estruturas do eIF5B-GTP ativo e o eIF5B-GDP inativo revelaram que existe uma modesta mudança conformacional no sítio de ligação ao GTP (Roll-Mecak et al, 2000) e recentes análises mutacionais indicaram que esta mudança conformacional regula a afinidade do ribossomo ao fator de iniciação eIF5B (Shin et al, 2002).

A porção N-terminal do fator eIF2 $\alpha$  de humanos, teve sua estrutura resolvida por cristalografia e revelou ser formado por dois domínios (Nonato et al, 2002). O sítio de fosforilação na serina 51, principal alvo de regulação do eIF2 $\alpha$ , está localizado em uma alça que não teve sua estrutura resolvida, indicando que ele é altamente flexível. A estrutura

cristalográfica do eIF4E humano ligado ao análogo da estrutura *cap*, m<sup>7</sup>GpppA, foi determinada (Tomoo et al, 2002) e comparada com a estrutura obtida do eIF4E de camundongo (Marcotrigiano et al, 1997). Apesar das estruturas centrais do eIF4E humano e de camundongo serem altamente conservadas, a porção C-terminal das estruturas são significativamente diferentes. A estrutura tridimensional do eIF4E também trouxe novos progressos para o entendimento da possível função da fosforilação da serina 209 na regulação da tradução. Com base na co-cristalização do eIF4E-m<sup>7</sup>GDP (Marcotrigiano et al, 1997), foi sugerido que a serina 209 fosforilada forma uma ponte com a lisina 159, que fica justamente sobre o encaixe que forma a trajetória do mRNA. Esta ponte foi caracterizada como sendo um grampo que estabiliza a interação mRNA/eIF4B, e isto é consistente com os dados de Minich e colaboradores (1994) que mostraram que o eIF4E fosforilado se liga com maior afinidade à estrutura *cap* do que quando não está fosforilado.

### 1.3 - Fator de Iniciação da Tradução 3

O fator de iniciação da tradução 3 (eIF3) foi primeiramente isolado e purificado como um complexo de alto peso molecular de reticulócitos de coelho (Bene e Hershey, 1976; Safer et al, 1976; Schreier et al, 1977). Desde sua caracterização inicial, a estrutura e função permanecem pouco entendidas. A composição e estequiometria das subunidades não têm sido determinadas rigorosamente, estruturas de alta resolução não estão disponíveis e a função de subunidades individuais ainda estão sendo elucidadas. O eIF3 de mamíferos possui massa molecular de aproximadamente 600 kDa sendo formado por 12 subunidades, nomeadas de acordo com a ordem decrescente de peso molecular (Browning et al, 2001): eIF3a, eIF3b, eIF3c, eIF3d, eIF3l, eIF3e, eIF3f, eIF3g, eIF3h, eIF3i, eIF3j e eIF3k (Hershey e Merrick, 2000; Morris-Desbois et al,

2001). A função do eIF3 de mamíferos tem sido elucidada por uma variedade de experimentos *in vitro*. Ele se liga diretamente à subunidade ribossomal 40S na ausência de outros componentes da iniciação da tradução e afeta a associação/dissociação dos ribossomos (Bene e Hershey, 1976; Thompson et al, 1977; Goss et al, 1988; Chaudhuri et al, 1999). Além disso, o eIF3 promove a ligação do Met-tRNA<sub>i</sub> e mRNA à subunidade ribossomal 40S (Hershey e Merrick, 2000) e se liga diretamente ao eIF1, eIF4B, eIF4G e eIF5 (Flatcher et al, 1999; Méthot et al, 1996; Lamphear et al, 1995; Bandyopadhyay e Maitra, 1999). De fato, o eIF3 tem um papel central na via de iniciação, organizando estruturalmente outros componentes traducionais na superfície da subunidade ribossomal 40S.

Das doze subunidades do eIF3 de mamíferos, cinco têm sua expressão alterada em alguns tipos de câncer humano. A eIF3a é a maior subunidade e é superexpressa em uma variedade de cânceres, como o de mama, esôfago e pulmão (Lin et al, 2001; Pincheira et al, 2001). A subunidade eIF3b também teve a sua superexpressão identificada em câncer de mama humano (Lin et al, 2001). Um aumento dos níveis de transcrição da eIF3c foi observado em todos os tumores testiculares examinados em estudos recentes (Rothe et al, 2000). Além disso, em um ensaio para amplificar genes relacionados com câncer de mama e pulmão, os níveis de mRNA da subunidade eIF3h foram encontrados superexpressos em aproximadamente 30% dos tumores de mama e em 20% dos carcinomas de mama (Nupponen et al, 1999). A subunidade eIF3e foi recentemente encontrada em níveis reduzidos em 40% dos carcinomas mamários e 30% dos cânceres de pulmão. A subunidade eIF3e de camundongos é codificada pelo gene Int6, que foi identificado como sendo um dos sítios de inserção do vírus do tumor de glândula mamária (MMTV). A disrupção deste gene está relacionada com formação de tumor (Marchetti et al, 2001) e tem sido relatado que eIF3e talvez seja também um regulador negativo do eIF3. A subunidade pode interagir com a proteína p56 o que resulta na inibição da síntese de proteínas in

*vitro* e *in vivo* (Guo et al, 2000). Níveis reduzidos de eIF3e podem, portanto, servir para aumentar a atividade do eIF3 e conseqüentemente o potencial para transformação.

A levedura Saccharomyces cerevisiae tem servido como ótimo modelo para o estudo do eIF3. Atualmente, análises de genética molecular têm permitido avanços significativos no entendimento da função do eIF3 neste organismo (Hinnebusch, 2000). A análise das interações entre as subunidades do eIF3 de levedura foram feitas através de ensaios de duplo-híbrido, ensaios de ligação in vitro e purificação por afinidade do complexo formado in vivo (Verlhac et al, 1997; Asano et al, 1998, 1999, 2000; Phan et al, 1998; Vernlocher et al, 1999; Valásek et al, 2001, 2002). Através destes estudos, foi identificado um grupo de cinco subunidades centrais componentes do eIF3 (Phan et al, 1998). Estas subunidades, Tif32p, Prt1p, Nip1p, Tif34p e Tif35p e uma subunidade subestequiométrica Hcr1p, são todas essenciais para o crescimento celular e, além disso, são ortólogas às subunidades eIF3a, eIF3b, eIF3c, eIF3g, eIF3i, eIFj de mamíferos, respectivamente (Hershey e Merrick, 2000; Hinnebusch, 2000). Além destas, ainda existem as subunidades Tif31p, Gcd10p e Sui1p, as quais se associam ao complexo central do eIF3 de S. cerevisiae (Hershey e Merrick, 2000). Os estudos de interação entre as proteínas componentes do eIF3 (sumarizados na Figura 1.2), sugerem que cada uma das três maiores subunidades do eIF3 (Tif32p, Nip1p e Prt1p) têm um sítio de ligação para as outras duas subunidades e que a extremidade C-terminal da Prt1p adicionalmente interage com as subunidades Tif34p e Tif35p. A Tif34p interage simultaneamente com a Tif35p e Prt1p. Hcr1p se liga simultaneamente ao domínio N-terminal da Prt1p e ao domínio C-terminal da Tif32p, sendo uma importante subunidade para a formação do complexo estável eIF3-40S (Fraser et al, 2003). Com os outros fatores de iniciação também já existem detalhes da interação. O eIF1 interage com o complexo central eIF3 através da extremidade C-terminal da Tif32p e da porção N-terminal da Nip1p. Em adição ao eIF1, a porção N-terminal da Nip1p também se liga ao domínio C-terminal

do eIF5. A subunidade  $\alpha$  do eIF2 faz dois contatos com o eIF3, uma interação direta com a porção C-terminal da Tif32p e uma associação indireta com a porção N-terminal da Nip1p via eIF5. A porção N-terminal da Tif32p, Nip1p e eIF5 compreendem uma unidade mínima de ligação à subunidade ribossomal 40S, que é suficiente para a ligação do 40S *in vivo* e *in vitro* (Valásek et al, 2003).



**Figura 1.2** - Modelo esquemático do complexo de iniciação eIF3 de *S. cerevisiae*, baseado em interações binárias entre subunidades isoladas (Valásek et al, 2002). O tamanho das proteínas representado na figura é proporcional ao peso molecular de cada. **TC**: complexo ternário (eIF2:GTP:tRNA<sub>i</sub>). *ctd*: domínio C-terminal do eIF5. *hld*: domínio tipo HRC1. **rrm**: sítio de reconhecimento de RNA.

O gene *TIF34* (*translation initiation factor 3, subunit 4*), isolado de *S. cerevisiae*, codifica uma proteína de aproximadamente 39 kDa que corresponde à subunidade Tif34p do eIF3 (Naranda et al, 1997). A proteína Tif34p é requerida para a viabilidade celular e tem um importante papel, embora não necessariamente direto, na iniciação da tradução, possivelmente

porque ela é requerida para a montagem e manutenção do complexo eIF3. Este possível envolvimento na síntese de proteínas foi sugerido pela análise de células mutantes (Verlhac et al, 1997). Para definir mais precisamente o papel da Tif34p em leveduras, foram gerados alelos sensíveis à temperatura do gene *TIF34* (Verlhac et al, 1997), demonstrando assim, que a progressão através do ciclo celular requer a Tif34p funcional. Esta proteína é essencial para a passagem através de G1 e G2/M do ciclo celular, mas não é necessária na fase S, uma vez que a replicação já foi iniciada. O requerimento da Tif34p para proceder através do ciclo celular, está provavelmente associado com sua função na iniciação da síntese de proteínas. Estudos anteriores (Shilo et al, 1979; Burke e Church, 1991) utilizando inibidores da síntese de proteínas, como a ciclohexamida, têm mostrado que o ponto de início na fase G1 e o início da mitose são muito sensíveis a uma pequena redução na síntese de proteínas. Ao passo que a replicação de DNA, uma vez iniciada, pode ocorrer sem a síntese de proteínas.

Experimentos de superexpressão da Tif34p não resultaram em níveis aumentados do eIF3, entretanto, a inibição da expressão do gene *TIF34* por glicose resultou na degradação de todas a subunidades do eIF3 numa velocidade mais rápida do que o normal, o que é consistente com o papel da Tif34p na integridade, estabilidade e manutenção do complexo eIF3 (Naranda et al, 1997). Utilizando o método do duplo-híbrido em leveduras e cepas mutantes, foi determinado que a proteína Tif34p interage com a porção N-terminal da proteína Tif35p e com a porção C-terminal da proteína Prt1p (Asano et al, 1998). O gene essencial *TIF35* de *S. cerevisiae* codifica a subunidade Tif35p do eIF3, de 30 kDa que é requerida *in vivo* na iniciação da síntese de proteínas (Hinachi et al, 1999). Esta proteína possui um domínio de ligação ao RNA (RRM – *RNA recognition motif*) na sua porção C-terminal e tem 33% de identidade e 43% de similaridade com a subunidade eIF3g do eIF3 humano (Hinachi et al, 1999). Um possível papel da Tif35p seria na organização estrutural e integridade funcional do complexo eIF3 (Verlhac et al, 1997). Já o gene

*PRT1*, que também é essencial para a célula, codifica a subunidade Prt1p do eIF3 de levedura e contém um RRM em sua porção N-terminal (Danaie et al, 1995). Prt1p é requerida *in vivo* para a iniciação da tradução e em ensaios envolvendo mutantes *prt1* termo-sensíveis, não houve a ligação do complexo ternário à subunidade 40S (Phan et al, 1998). Todos estes experimentos demonstraram que a Tif34p pertence ao complexo central do eIF3 juntamente com a Tif35p e a Prt1p e tem um importante papel na integridade do sistema de iniciação da tradução. Estudos ainda serão necessários para se caracterizar melhor o papel da Tif34p na iniciação da tradução, assim como sua interação com as subunidades Tif35p e Prt1p, a qual ainda não foi bem estabelecida.

A proteína Tif34p de S. cerevisiae apresenta 46% de identidade e 64% de similaridade de aminoácidos com a proteína TRIP-1 (TGF- $\beta$ -receptor interacting protein 1) de origem humana (Naranda et al, 1997). A estrutura primária da TRIP-1 é idêntica a da subunidade eIF3i do eIF3 de mamíferos (Asano et al, 1997), estando envolvida na iniciação da tradução. Além disso, essa proteína liga-se ao domínio citoplasmático do receptor tipo II do TGF-β (transforming growth *factor-\beta*) sendo fosforilada em resíduos serina/treonina pelo receptor, o qual possui atividade de proteína-quinase (Chen et al, 1995). A sinalização via TGF-\u03b3 é mediada por dois tipos de receptores quinase (serina/treonina), os receptores tipo I e tipo II, que formam um complexo heterotrimérico. Foi elaborado um modelo de sinalização via TGF-β, no qual um ligante interage com o receptor tipo II e estimula a união do receptor tipo I e tipo II, formando o complexo. O receptor tipo II fosforila o tipo I que por sua vez fosforila fatores de transcrição chamados SMADs. Fosforiladas, estas proteínas migram para o núcleo da célula, se associam com outras proteínas SMADs e ativam ou reprimem a transcrição de genes alvos (Massague, 2000). Já foram caracterizadas várias moléculas envolvidas neste modelo de transdução de sinal, mas ainda pouco se sabe como elas se relacionam durante a sinalização. As proteínas pertencentes à família TGF- $\beta$  estão envolvidas em várias funções celulares importantes, entre elas, a regulação da proliferação celular, determinação de linhagem, diferenciação e morte das células (Massague, 1998).

A expressão de mutantes da TRIP-1 em células de mamíferos reforçaram a hipótese de que ela pode estar relacionada com mecanismos de transdução de sinal mediado pelo TGF-β (Choy e Derynck, 1998). Além disso, utilizando a técnica de *phage display*, Sheu e colaboradores (2003) identificaram a TRIP-1 como sendo uma proteína regulatória que interage com uma enzima chamada TRAP (*tartrate resistant acid phosphatase*) e modula a diferenciação dos osteoblastos, através da sinalização de proteínas SMADs. Estes autores demonstraram que TRIP-1 interage com TRAP no citoplasma, sendo esta interação necessária para a ativação do receptor tipo II TGF-β pela TRIP-1. Não se sabe ainda, quais são os alvos subseqüentes de interação da TRIP-1 e nem o significado da fosforilação para a sua função, sendo assim, serão necessários estudos mais detalhados para determinar o papel da TRIP-1 na transdução de sinal.

### 1.4 - Família de Proteínas Formadas por Domínios WD (WD-repeat)

Por similaridade de seqüência primária, as proteínas Tif34p e TRIP-1 pertencem à família de proteínas *WD-repeat*. As proteínas *WD-repeat* são encontradas no citoplasma e núcleo, ligadas ao citoesqueleto ou associadas com membranas (van Nocker e Ludwig, 2003). Os membros desta família exercem uma grande variedade de funções (Li e Roberts, 2001), como: transdução de sinal (exemplo: *RACK1* e proteína G), processamento e síntese de RNA (*TAFs* e *TUP1*), montagem da cromatina (*CAF-1* e *HIR1*), transporte vesicular (*SEC13* e *COP*), montagem do citoesqueleto (*MAP* e *Arp2*), regulação do ciclo celular (*CDC4* e *CDC20*), apoptose (*Apaf-1* e *Dark*) e além destas funções importantes, recentemente foram identificadas proteínas

relacionadas com doenças humanas (Li e Roberts, 2001). Mutações em genes que codificam algumas das proteínas desta família podem levar a doenças como, por exemplo, as síndromes de Cockayne (Henning et al, 1995) e do Triplo-A (Tullio-Pelet et al, 2000).

A característica comum destas proteínas é o domínio WD (também chamado de Trp-Asp ou WD-40), o qual compreende de 4 a 16 repetições conservadas de aproximadamente 40 aminoácidos, que são iniciadas com o dipeptídeo glicina-histidina (GH) 11 a 24 resíduos da porção N-terminal e apresentam o dipeptídeo triptofano-aspartato (WD) na sua porção C-terminal (Smith et al, 1999) (Figura 1.3). Seqüências contidas dentro destes domínios são divergentes e podem conter inserções que complicam sua identificação com base somente em informações da estrutura primária (Madrona e Wilson, 2004). A caracterização estrutural destas proteínas está começando a revelar a possível diversidade dentro destas repetições e irá fornecer uma base para se fazer predições mais precisas no futuro.



**Figura 1.3** - Padrão geral da seqüência de aminoácidos do domínio WD. Cada domínio WD tem uma seqüência inicial variável, seguida por uma seqüência central de tamanho mais ou menos constante que começa tipicamente com GH (glicina e histidina) e termina com WD (triptofano e aspartato). X: qualquer aminoácido. N<sub>4-16</sub>: número de repetições do domínio WD (Garcia-Higuera et al, 1996; Smith et al, 1999).

Atualmente é aceito que proteínas com domínios WD agem como sítios para interação com outras proteínas. Esta característica permite três papéis funcionais gerais (van Nocker e Ludwig, 2003). Primeiro, as repetições WD dentro de uma proteína podem promover sítios de ligação para outras duas ou mais proteínas e favorecer interações transientes entre elas. Dois exemplos que ilustram melhor são a proteína G humana, uma GTPase que se associa com diversos receptores de membrana para propagar a resposta celular (van der Voorn e Ploegh, 1992), e a Hat2 de S. *cerevisiae*, que é requerida para uma interação eficiente da subunidade catalítica Hat1 com a histona alvo (Ruiz-Garcia et al, 1998). Um segundo potencial papel para proteínas WD-repeat, é como um componente integral de complexos de proteínas. Esta função é melhor ilustrada pela proteína de levedura Pfs2, a qual é associada com a poli(A) polimerase Pap1 e com vários outros fatores em um grande complexo requerido para o processamento e poliadenilação da extremidade 3'do pré-RNA (Zhao et al, 1999). O terceiro papel reconhecido para proteínas contendo domínios WD, é o de agir como domínio de interação modular de proteínas maiores. Neste caso, elas conduzem a proteína e seus domínios associados para seus alvos. Dois exemplos em plantas são as proteínas COP1 e SPA1, que justapõe o domínio WD C-terminal com um domínio ring-finger N-terminal ou domínio tipo quinase, respectivamente (Ang et al, 1998; Hoecker e Quail, 2001).

A primeira proteína WD que teve sua estrutura tridimensional determinada foi a subunidade  $\beta$  da proteína G heterotrimérica humana (Wall et al, 1995; Lambright et al, 1996; Sondek et al, 1996), a qual revelou que os domínios WD formam folhas- $\beta$  antiparalelas conectadas por alças e compostas predominantemente por quatro fitas cada. Estas folhas- $\beta$  se dobram em uma estrutura altamente ordenada e simétrica denominada  $\beta$ -propeller, a qual é visualizada como um pequeno disco onde as fitas formam as paredes (Smith et al, 1999). O primeiro e o último domínio WD (isto é, o amino e o carboxi-terminal) interagem, fechando e reforçando a estrutura circular (Smith et al, 1999). Acredita-se que quatro domínios WD, no

mínimo, são necessários para formar a  $\beta$ -propeller (Chothia et al, 1997), entretanto, a existência de domínios WD não é fator limitante para a formação deste tipo de estrutura (Neer e Smith, 1996). Outras famílias que não apresentam similaridade de estrutura primária com as proteínas da família *WD-repeat* formam a  $\beta$ -propeller idêntica à das proteínas *WD-repeat*.

Foram identificadas com base na seqüência de aminoácidos, 136 proteínas WD-repeat em humanos, 98 em Drosophila melanogaster, 72 em Caenorhabditis elegans, 56 em Saccharomyces cerevisiae (Li e Roberts, 2001) e 237 em Arabidopsis thaliana (van Nocker e Ludwig, 2003). De toda esta gama de proteínas, poucas tiveram sua estrutura tridimensional resolvida por cristalografia de raios-X. De eucariotos, pode-se citar além da subunidade  $\beta$  da proteína G humana já discutida, o domínio C-terminal da proteína Tup1 (Sprague et al, 2000), a proteína Aip1p (Voegtli et al, 2003), a Cdc4 (Orlicky et al, 2003) e a Ski8p (Madrona e Wilson, 2004), ambas de S. cerevisiae, o domínio C-terminal da Groucho/TLE1 humana (Pickles et al, 2002) e a subunidade p40 do complexo Arp2/3 bovino (Robinson et al, 2001). Das sete proteínas da família WD-repeat identificadas em procariotos (Li e Roberts, 2001), apenas a TolB teve sua estrutura determinada por cristalografia (Carr et al, 2000). Apesar de basicamente todas as estruturas resolvidas apresentarem a formação da  $\beta$ -propeller, variando-se o número de repetições dos domínios WD, algumas características puderam ser evidenciadas. Como exemplo pode-se citar a presença de uma folha-β formada por seis fitas no domínio C-terminal da proteína Tup-1 e na Groucho/TLE1 e, além disso, foi possível determinar a presença de mais um domínio na seqüência da proteína Tup-1 e mais dois domínios WD na Ski8p, os quais não foram identificados na següência primária destas proteínas. Os dados estruturais aliados aos estudos funcionais têm trazido importantes informações sobre os mecanismos pelos quais as proteínas exercem sua função na célula. Smith e colaboradores (1999) afirmaram que todas as proteínas WD-repeat pertencem a uma classe estruturalmente semelhante, pois há uma correspondência entre as seqüências WD e as repetições estruturais. Apesar de serem importantes estruturalmente, o que se sabe até o momento, baseado em estudos funcionais, é que as repetições WD são importantes para a interação entre proteínas, servindo como base para diversos processos celulares vitais.

A proteína Tif34p de *S. cerevisiae* como já dito, é componente do eIF3, sendo responsável pela integridade e manutenção deste fator de iniciação da tradução. Não existem estudos sobre a participação dos domínios WD na interação com as outras subunidades do complexo, porém estes devem ter alguma participação, baseando-se nas informações obtidas sobre a função dos domínios WD de outras proteínas da família. Foram identificados cinco domínios WD conservados na seqüência de aminoácidos da proteína Tif34p. Entretanto, existem espaços onde podem existir outros domínios e que podem formar folhas- $\beta$ . O estudo estrutural desta proteína irá estabelecer se realmente há correspondência entre as repetições WD com as folhas- $\beta$  e, além disso, será possível estudar em detalhe o papel desta proteína na iniciação da tradução.

### 1.5 - Técnicas para a Caracterização Estrutural de Proteínas

Existem várias técnicas disponíveis para o estudo de proteínas em solução. Neste trabalho, foram utilizadas as técnicas de espalhamento de luz dinâmico (DLS), dicroísmo circular (CD) e espalhamento de raios-X a baixo ângulo (SAXS), as quais serão comentadas a seguir.

#### **1.5.1 - Espalhamento de Luz Dinâmico**

Recentemente, esta técnica não-destrutiva tem obtido popularidade por explorar algumas propriedades, como tamanho e forma, de moléculas em solução (Murphy, 1997). O espalhamento de luz dinâmico (DLS) é uma técnica que mede o coeficiente de difusão de uma partícula em solução e o seu raio hidrodinâmico, o qual é geralmente calculado baseado em modelos esféricos. A maior vantagem do espalhamento de luz, além de fornecer informações estruturais, é o curto tempo necessário para a obtenção dos dados, a modesta quantidade de proteína utilizada (de 1 a 10 mg/mL) e pequenos volumes de amostra requeridos (60 µL), além de sua relativa simplicidade de manipulação. Pode-se obter informações estruturais sobre partículas de 1-1000 nm de tamanho (Murphy, 1997) e, além disso, esta técnica é extremamente sensível a pequenas quantidades de agregados (Papish et al, 2002). Devido à sua velocidade na obtenção de dados, o DLS permite o estudo de mudanças conformacionais de proteínas ou em complexos de proteínas em solução. Esta técnica tem diversas aplicações como: caracterização de proteínas incluindo agregação, conformação, estabilidade estrutural, desnaturação e estabilidade térmica; complexos macromoleculares; polímeros biológicos e sintéticos.

No DLS, um feixe de luz monocromático de 827,6 nm incide sobre a amostra e a flutuação da intensidade de luz espalhada (escala de tempo em microsegundos) é detectada por um fotodiodo a 90° (Murphy, 1997). A posição das partículas em relação ao detector influencia na intensidade de espalhamento medida pelo detector. Partículas em solução movem-se ao acaso (movimento Browniano) e este movimento causa flutuações de concentração em um dado volume da solução em função do tempo. No espalhamento de luz dinâmico, a intensidade de espalhamento (em curtos intervalos de tempo) é registrada como função do tempo, e flutuações nessa intensidade de espalhamento são observadas como variações de concentração no volume de espalhamento. O tempo necessário para as flutuações na intensidade de luz espalhada

acontecerem, é a característica mais importante do sinal, uma vez que esse tempo contém informações sobre as propriedades dinâmicas das moléculas do soluto. Uma partícula maior terá o coeficiente de difusão menor do que uma partícula pequena e as flutuações ocorrerão mais vagarosamente (Pecora, 1985). A informação mais simples obtida é o coeficiente de difusão translacional que, para uma partícula esférica, é relacionado ao raio hidrodinâmico. Através do coeficiente de difusão calculado, pode-se determinar o tamanho da partícula esférica (raio hidrodinâmico), o qual relaciona difusão x tamanho, o que leva a concluir que o espalhamento é diretamente proporcional à massa molecular, pois o aumento da massa leva ao aumento do tamanho do raio. A função de auto-correlação do sinal de intensidade de espalhamento é a convergência da intensidade do sinal em função do tempo e o gráfico típico, apresenta para cada população monodispersa de partículas, um decaimento exponencial. Assim pode-se analisar se a amostra apresenta monodispersão através da avaliação da curva de auto-correlação e do cálculo do coeficiente de difusão.

### 1.5.2 - Dicroísmo Circular

O método espectroscópico de dicroísmo circular (CD) é bastante utilizado nas análises conformacionais de proteínas, onde o espectro obtido é determinado pela quantidade e composição de estrutura secundária (Fasman, 1996). Moléculas assimétricas em solução são opticamente ativas, isto é, quando elas interagem com uma luz circularmente polarizada elas absorvem luz polarizada à direita e à esquerda de maneira desigual. Esta diferença na absorção gera luz elipticamente polarizada que pode ser medida por um espectropolarímetro de CD, que converte este sinal em elipticidade ( $\theta$ ) expresso em miligraus. O sinal de CD será positivo ou negativo dependendo da absorção preferencial da luz polarizada à direita ou à esquerda. Os cromóforos mais numerosos e característicos que proteínas e peptídeos possuem para luz

polarizada circularmente na faixa do UV-distante (170 a 260 nm) são as ligações peptídicas (Kelly e Price, 1997). Os espectros de CD produzidos por esses cromóforos são característicos para os diversos tipos de estrutura secundária. As faixas de absorção dos resíduos aromáticos no UV-distante são pequenas quando comparadas às das ligações peptídicas. Contudo, no UV-próximo (260 a 320 nm), o espectro de CD passa a ser dominado pela contribuição das cadeias laterais aromáticas, juntamente com as pontes dissulfeto e grupamentos prostéticos. Assim, enquanto o UV-distante detecta estrutura secundária, o UV-próximo é sensível à estrutura terciária, o que compõe a base para a análise da conformação de proteínas através de dicroísmo circular (Kelly e Price 1997). Experimentos de dicroísmo circular têm sido muito utilizados para a investigação de: estrutura secundária de peptídeos e proteínas; mudanças conformacionais em proteínas; desenovelamento de proteínas; desenho de peptídeos e proteínas; proteínas de membrana; interação entre domínios de proteínas e ligação de substratos (Kelly e Price, 2000).

Por esta técnica ser sensível à quantidade total de estrutura secundária de uma proteína, ela é usada para investigar a quantidade de  $\alpha$ -hélice, folha- $\beta$  e estrutura randômica que a compõe (Fasman, 1996). As  $\alpha$ -hélices e folhas- $\beta$  são estruturas bem compactas presentes nas proteínas enoveladas e a estrutura randômica está presente nas proteínas desenoveladas e em alças que conectam os elementos de estrutura secundária. Proteínas que possuem sua estrutura secundária formada predominantemente por  $\alpha$ -hélices exibem picos negativos a 208 e 222 nm e um pico positivo entre 190 e 195 nm. As proteínas formadas predominantemente por folhas- $\beta$  apresentam espectro de CD contendo um pico negativo em aproximadamente 216 nm e um positivo a 196 nm. A estrutura randômica é representada pelo espectro contendo um pico negativo próximo de 195 nm e um pico positivo a 212 nm (Pelton e McLean, 2000). A Figura 1.4 traz a representação dos espectros de CD característicos de cada tipo de estrutura secundária.



**Figura 1.4** - Representação dos espectros característicos de cada tipo de estrutura secundária obtidos por CD (http://www.ap-lab.com/circular\_dichroism.htm).

Os dados obtidos por CD podem ser interpretados e convertidos em porcentagem de estrutura secundária, levando-se em consideração a concentração e o número de aminoácidos da proteína estudada. Para isso, existem programas computacionais de predição de estrutura que geram informações detalhadas sobre a quantidade e composição de estrutura secundária da amostra analisada, baseados em bancos de dados de espectros de proteínas que já possuem sua estrutura cristalográfica determinada. Como exemplo de programa, pode-se citar o CDNN (Bohm et al, 1992), CONTIN (Provencher e Glöckner, 1981), entre outros.

### 1.5.3 - Espalhamento de Raios-X a Baixo Ângulo

O espalhamento de raios-X a baixo ângulo (SAXS) é uma ferramenta fundamental no estudo de macromoléculas biológicas em solução que variam entre poucos quilodaltons a vários megadaltons (Feigin e Svergun, 1987). A maior vantagem deste método está na sua habilidade de dar informações estruturais sobre sistemas parcialmente ou completamente desordenados (Svergun e Koch, 2003). O SAXS permite estudar a estrutura de partículas nativas em ambientes próximos ao fisiológico e analisar mudanças estruturais em resposta a variações nas condições externas (Svergun e Koch, 2003). Devido ao requerimento de bons cristais para cristalografia de raios-X e a necessidade de proteínas de baixo peso molecular para ressonância magnética nuclear, uma fração significante de proteínas não pode ser analisada usando estes dois métodos. Além disso, a função das proteínas está relacionada não apenas ao arranjo tridimensional dos peptídeos, mas também à sua mobilidade intrínseca. A cristalografia de raios-X pode trazer informações em relação à posição de grupos atômicos individuais dentro de uma macromolécula, mas regiões flexíveis e desordenadas podem parecer ausentes (Petoukhov et al, 2002). Por isso, progressos na instrumentação de SAXS e novos métodos de análise, que substancialmente melhoram a resolução e a realidade dos modelos estruturais, estão fazendo do método uma importante ferramenta (de baixa resolução) para estudo estrutural.

Esta técnica permite determinar várias características das macromoléculas, tais como raio de giro, peso molecular, volume, grau de hidratação, diâmetro máximo e forma (Glatter e Kratki, 1982). Nos experimentos de espalhamento, a solução de macromoléculas é exposta aos raios-X e o padrão de SAXS resulta de uma média de espalhamento de um conjunto de partículas orientadas randomicamente na amostra, reduzindo a resolução do método (Petoukhov et al, 2002). A intensidade espalhada (I) é registrada como uma função do ângulo de espalhamento (Svergun et al, 2001). Tradicionalmente, apenas limitadas porções do padrão de espalhamento

podem ser usadas para a construção de modelos tridimensionais (Svergun et al, 2001). A baixos ângulos (resolução de 2 a 3 nm), os raios-X são insensíveis à estrutura interna e o espalhamento é essencialmente dado pela forma da partícula. A intenção da análise *ab initio* obtida através dos dados de SAXS é recuperar a estrutura tridimensional através do padrão de espalhamento unidimensional (Svergun e Koch, 2002).

O experimento para obtenção dos dados consiste em direcionar um feixe de radiação para a amostra e medir a variação de intensidade espalhada em função do ângulo de espalhamento. Essa radiação deve possuir comprimento de onda bem definido na faixa de raios-X, entre 0,5 e 2 Å, ter uma boa colimação e alta intensidade (Linder e Zemb, 1991). Esta técnica fornece informações estruturais de flutuações ou heterogeneidades da densidade eletrônica com dimensões características da ordem de 10 a 1000 Å.

O espalhamento de raios-X a baixos ângulos somente é observado, quando as amostras analisadas apresentam heterogeneidade em sua densidade eletrônica (Glatter e Kratky, 1982). A densidade eletrônica  $\rho$  é definida como a quantidade em mols de elétrons por unidade de volume. Quando partículas estão no vácuo, a amplitude de luz espalhada é proporcional à densidade eletrônica das partículas. Porém, quando partículas analisadas estão dispersas em uma matriz, a diferença entre a densidade eletrônica da partícula e da matriz é que influencia na amplitude de luz espalhada. Essa diferença é chamada de contraste de densidade eletrônica e é um dos principais parâmetros da amostra quando a técnica de SAXS é utilizada (Feigin e Svergun, 1987). Como a técnica de SAXS é sensível à heterogeneidade em escala maior que a atômica, em muitos casos as macromoléculas espalhadoras podem ser descritas como tendo uma densidade eletrônica constante e única, sem considerar as variações da densidade eletrônica em escala atômica. As proteínas podem estar muito diluídas ou não na solução, possuindo assim uma certa correlação espacial que influencia a intensidade espalhada. A forma das curvas de SAXS contém

informações sobre o tamanho e a forma de partículas e também sobre a correlação espacial das mesmas.

A função p(r) é chamada de função de distribuição de distâncias (Glatter e Kratky, 1982). Ela possui uma definição geométrica que pode ser entendida dividindo-se uma partícula em um grande número de elementos de volumes pequenos e idênticos. A função p(r) descreve a distribuição de distâncias r que podem ser encontradas a partir de combinação de qualquer par de elementos de volume (Glatter e Kratky, 1982). Em princípio, a função de distribuição de distâncias p(r) contém a mesma informação da intensidade de espalhamento I(s), mas a representação do espaço real é mais intuitiva e a informação sobre a forma da partícula pode freqüentemente ser deduzida diretamente por uma inspeção visual da p(r) (Glatter e Kratky, 1982).

A técnica de SAXS permite o estudo de sistemas monodispersos e polidispersos. Para sistemas monodispersos é possível determinar a forma e o tamanho das partículas observadas. Estas condições são normalmente encontradas em soluções de macromoléculas biológicas a diluições infinitas, nas quais a intensidade total de espalhamento é a somatória do espalhamento de cada uma das N partículas existentes na amostra irradiada (Linder e Zemb, 1991). Variações na forma, tamanho e estrutura das partículas do sistema analisado, são refletidas na função de espalhamento I(q) (a intensidade de onda espalhada em função do espaço recíproco) e na função de distribuição de distâncias p(r); portanto, a monodispersividade pode então ser determinada a partir das curvas de espalhamento obtidas. Para sistemas polidispersos, a distribuição de tamanhos pode ser analisada, assumindo-se que todas as partículas possuem a mesma forma (Linder e Zemb, 1991).

O raio de giro (Rg) é definindo como o raio de uma fina camada esférica de mesma massa e momento de inércia (Glatter e Kratky, 1982). Este parâmetro é uma medida da distribuição das cargas em relação ao centro de massa eletrônico de uma partícula e pode ser obtido através da
função de distribuição de distâncias p(r) ou diretamente pela curva de espalhamento (Glatter e Kratky, 1982). O conhecimento do raio de giro pode ser muito útil na investigação estrutural de uma partícula em solução, pois um aumento ou diminuição do mesmo, provocado por qualquer variação nas condições da solução, indicaria uma mudança nas posições relativas entre os átomos que compõe a partícula.

Na maioria dos casos, o efeito de interferência interpartículas é tido como uma perturbação do experimento, no qual a informação que interessa é a função de espalhamento da partícula. Qualquer efeito de concentração pode levar a erros indesejáveis nas determinações dos parâmetros das partículas em análise (Linder e Zemb, 1991). O problema pode ser solucionado através de medidas de funções de espalhamento I(q) para amostras em diferentes concentrações, normalmente de 1-2 mg/mL até 50 mg/mL. O efeito da concentração pode então ser visto num gráfico das curvas de espalhamento divididas pela concentração utilizada em cada determinação. Em alguns casos pode ser impossível fazer experimentos variando-se a concentração das amostras, já que a estrutura de algumas partículas pode ser dependente de sua concentração. Nesses casos, é possível encontrar os parâmetros das partículas negligenciando-se a parte das funções de espalhamento que são influenciadas pelo efeito de concentração (Linder e Zemb, 1991).

A visualização e manipulação de estruturas atômicas e modelos de baixa resolução, a partir de modelos de alta resolução, é obtida através de programas como, por exemplo, o CRYSOL (Svergun et al, 1995) e o GNOM (Svergun, 1992), entre outros. Quando não se têm modelos de alta resolução para comparação, parte-se para uma análise mais cuidadosa, utilizando-se a função p(r) para inferir a curva experimental dada pelos programas CREDO (Petoukhov et al, 2002), DAMAVER (Svergun et al, 2001), DAMIM (Svergun, 1999), entre outros.

# **2. OBJETIVOS**

Este trabalho visa a caracterização estrutural da proteína Tif34p de *Saccharomyces cerevisiae* e do complexo Tif34p/Tif35p, envolvendo três objetivos específicos:

1°.)Clonagem, expressão em *Escherichia coli* e purificação das proteínas Tif34p e Tif35p.

2°.)Caracterização do complexo por cromatografia de gel filtração e espalhamento de luz dinâmico.

3°.)Caracterização estrutural da Tif34p e do complexo Tif34p/Tif35p.

- Caracterização da estrutura secundária por espectroscopia de dicroísmo circular.

- Caracterização da estrutura terciária por espalhamento de raios-X a baixo ângulo.

# **3. MATERIAIS E MÉTODOS**

# 3.1 - Meios de Cultura e Manipulação Genética de Bactérias

# 3.1.1 - Meios de Cultura para Manutenção e Seleção

As bactérias foram crescidas em meio LB (Sambrook et al, 1989), composto por 1% de peptona, 0,5% de extrato de levedura e 1% de cloreto de sódio. A preparação do meio sólido foi feita adicionando-se 2% de ágar bacteriológico. O meio foi autoclavado por 20 min a 115°C e o sólido, distribuído em placas de Petri descartáveis e estéreis.

### 3.1.2 - Transformação de E. coli

## 3.1.2.1 - Transformação Mediada por Polietileno-Glicol (PEG)

Para tornar as cepas DH5 $\alpha$  e BL21(DE3) competentes para transformação por PEG, inoculou-se 200 mL de meio contendo 20 g/L triptona, 5 g/L extrato de levedura, 5,8 g/L NaCl, 2,5 g/L MgSO<sub>4</sub> com 50 mL de uma pré-cultura, a 37°C, com agitação de 200 rpm até atingir densidade óptica a 600 nm igual a 0,6. As culturas foram submetidas a choque térmico em banho de água gelada e centrifugadas a 4.000 g por 15 min. As células foram ressuspendidas em 100 mL de tampão contendo 2,94 g/L acetato de potássio, 9,9 g/L MnCl<sub>2</sub>, 7,46 g/L KCl, 1,47 g/L CaCl<sub>2</sub>, 15% glicerol e centrifugadas a 4.000 g por 8 min. O precipitado foi ressuspendido em 10 mL de tampão contendo 2,09 g/L MOPS pH 7,0; 11,02 g/L CaCl<sub>2</sub>; 0,75 g/L KCl; 15% glicerol e a suspensão armazenada em alíquotas de 60  $\mu$ L a -80°C. Todas as soluções utilizadas foram mantidas a 4°C. As cepas DH5 $\alpha$  e BL21(DE3) foram transformadas pelo método PEG (Sambrook et al, 1989) onde 100 ng de DNA do vetor ou da ligação foram misturados com 100  $\mu$ L de células competentes, juntamente com 80  $\mu$ L de tampão de transformação (1 M KCl; 0,3 M CaCl<sub>2</sub>; 0,5 M MgCl<sub>2</sub>; 10% PEG). A reação foi incubada no gelo por 30 min e a seguir à temperatura ambiente por 10 min. Subseqüentemente, 1 mL de meio LB foi adicionado à reação e a cultura foi incubada por 50 min a 37°C. Duzentos  $\mu$ L da suspensão foram plaqueados em meio LB contendo o antibiótico necessário para a seleção dos transformantes [canamicina (50  $\mu$ g/mL) para o vetor pET28a; ampicilina (50  $\mu$ g/mL) para o vetor pET23a]. As placas foram mantidas por 16 h a 37°C.

# 3.1.2.2 - Transformação por Eletroporação

Para tornar a cepa mutante BL21(DE3)*slyD*<sup>-</sup> competente para transformação por eletroporação, inoculou-se 1 L de meio LB contendo 10 mL de tetraciclina (5 mg/mL) com 10 mL de uma pré-cultura contendo a cepa BL21(DE3)*slyD*<sup>-</sup>, a 37°C sob agitação constante de 200 rpm até atingir OD<sub>600</sub> igual a 0,6. Transferiu-se a cultura para tubos de centrífuga, os quais foram incubados no gelo por 30 min. As células foram sedimentadas por 10 min a 4.000 g e ressuspendidas em 1 L de água Milli-Q autoclavada. Uma outra etapa de centrifugação e lavagem foi realizada, seguida por nova centrifugação. As células foram então ressuspendidas em 20 mL de glicerol 10%, sedimentadas por centrifugação e ressuspendidas em 4 mL de glicerol 10%. Alíquotas de 50 µL foram armazenadas a -80°C. Todas as soluções utilizadas estavam a 4°C. A transformação da cepa BL21(DE3)*slyD*<sup>-</sup> foi realizada por eletroporação, utilizando-se o sistema Gene Pulser II (BioRad). Adicionou-se 2 µL de DNA do vetor em 40 µL de células competentes. Após a aplicação de um pulso de 2,5 kV por 4,5 mseg, as células foram cultivadas em 1 mL de LB por 1 h a 37°C.

# 3.2 - Plasmídeos e Cepas

Os plasmídeos e cepas de *Escherichia coli* utilizados neste trabalho estão listados nas Tabelas 3.1 e 3.2, respectivamente. No Anexo I está o mapa dos plasmídeos pET28a e pET23a.

Nome	Plasmídeo	Inserto	Oligonucleotídeos	Fusão	Referência
pET28a	Base -	-	-	-	Novagen
pET23a	-	-	-	-	Novagen
pET28a+HisTIF34	pET28a	TIF34* nt 1 a 1044	5' TIF34-Bam 3' TIF34-Hind	6-His N-terminal	Este trabalho
pET23a+TIF34	pET23a	TIF34* nt 1 a 1044	5' TIF34-Nde 3' TIF34-Hind	-	Este trabalho
pET28a+HisTIF35	pET28a	TIF35** nt 1 a 825	5' TIF35-Bam 3' TIF35-Sal	6-His N-terminal	Este trabalho
pET28a+HisTIF35N	pET28a	TIF35N** nt 1 a 456	5' TIF35-Bam 3' TIF35-Xho	6-His N-terminal	Este trabalho

Tabela 3.1 - Plasmídeos de E. coli.

Código de acesso da seqüência depositada no GenBank:

\* U56937

\*\* AF004913

Сера	Genótipo / Descrição	Referência
DH5α	<i>sup</i> E44 ∆ <i>lac</i> U169 ( <i>ø</i> 80 <i>lac</i> Z∆M15) <i>hsdR</i> 17 <i>recA</i> 1 <i>endA</i> 1 <i>gyrA</i> 96 <i>thi</i> -1 <i>relA</i> 1	Hanahan, 1983
BL21(DE3)	hsdS gal (\lambda clts857 ind1 Sam7 nin5 lacUV5-T7 gene1)	Studier e Moffatt, 1986
BL21(DE3) <i>slyD</i> <sup>-</sup>	BL21(DE3) contendo gene <i>slyD</i> deletado	Yan et al, 2001
BL21-H34	BL21(DE3) transformada com o plasmídeo pET28a+HisTIF34	Este trabalho
BL21-34	BL21(DE3) transformada com o plasmídeo pET23a+TIF34	Este trabalho
BL21-H35	BL21(DE3) <i>slyD</i> <sup>-</sup> transformada com o plasmídeo pET28a+HisTIF35	Este trabalho
BL21-H35N	BL21(DE3) <i>slyD</i> <sup>-</sup> transformada com o plasmídeo pET28a+HisTIF35N	Este trabalho
BL21-34H35	BL21(DE3) <i>slyD</i> <sup>-</sup> co-transformada com os plasmídeos pET23a+TIF34 e pET28a+HisTIF35	Este trabalho
BL21-34H35N	BL21(DE3) <i>slyD</i> <sup>-</sup> co-transformada com os plasmídeos pET23a+TIF34 e pET28a+HisTIF35N	Este trabalho

Tabela 3.2 - Cepas de E. coli.

#### 3.3 - Amplificação de DNA por PCR

As reacões de PCR contém: 20 mM Tris-HCl pH 8,4; 50 mM KCl; 1,5 mM MgCl<sub>2</sub>; 0,2 mM de cada um dos quatro deoxiribonucleotídeos trifosfato, oligonucleotídeos específicos na concentração final de 0,5 µM; 20 ng de DNA template e 2,5 unidades de enzima Taq DNA Polimerase (Life Technologies) em um volume final de 50 µL. Foi utilizada a ciclagem de 2 min a 94°C, seguindo-se 25 ciclos de 1 min a 94°C, 2 min na temperatura de anelamento, 1 min à 68°C e 5 min à 68°C para extensão.

Os oligonucleotídeos utilizados nas reações de PCR estão listados na Tabela 3.3. As seqüências de nucleotídeos sublinhadas correspondem aos sítios de restrição.

Oligonucleotídeo	Seqüência	Sítio de restrição
5' TIF34-Bam	5` GACA <u>GGATCC</u> ATGAAGGCTATCAAATTAAC 3`	<i>Bam</i> HI
3' TIF34-Hind	5` ATGT <u>AAGCTT</u> ACTATACATCAACATCAGC 3`	HindIII
5' TIF34-Nde	5´ AAGG <u>CATATG</u> AAGGCTATCAAATTAACAG 3'	Ndel
5' TIF35-Bam	5` ACCA <u>GGATCC</u> ATGAGTGAAGTTGCACCA 3`	BamHI
3' TIF35-Sal	5` TTCT <u>GTCGAC</u> CTATTCCTTAACCTTAGGTT 3`	Sall
3' TIF35-Xho	5' CTTCCTCGAGTTAGCCTGCCTTCTCTC 3'	Xhol

 Tabela 3.3 - Oligonucleotídeos utilizados neste trabalho.

A sequência codificadora da HisTif34p foi amplificada por PCR usando-se os oligonucleotídeos 5'TIF34-Bam e 3'TIF34-Hind, juntamente com a enzima Taq DNA Polimerase. Os sítios de restrição BamHI e HindIII, indicados na Tabela 3.3, foram necessários para a clonagem no vetor pET28a, dando origem ao vetor pET28a+HisTIF34 (Tabela 3.1). Uma segunda reação foi feita com os oligonucleotídeos 5'TIF34-Nde e 3'TIF34-Hind, juntamente com a enzima Taq DNA Polimerase para a clonagem da seqüência da Tif34p nos sítios de NdeI e *Hind*III do vetor pET23a, gerando o vetor pET23a+TIF34 (Tabela 3.1). Nas duas reações descritas acima, foi utilizado como *template* o DNA genômico da cepa W303-1a de *S. cerevisiae*.

A seqüência codificadora da proteína HisTif35p foi amplificada por PCR utilizando-se os oligonucleotídeos 5'TIF35-Bam e 3'TIF35-Sal (Tabela 3.3). O produto da PCR foi inserido nos sítios *Bam*HI e *Sal*I do vetor pET28a, originando o vetor pET28a+HisTif35 (Tabela 3.1). Uma seqüência de 456 nucleotídeos correspondentes à extremidade N-terminal do gene *TIF35*, a qual codifica a proteína HisTif35N, foi amplificada por PCR e os sítios de restrição para as enzimas *Bam*HI e *Xho*I foram inseridos nos oligonucleotídeos 5'TIF35-Bam e 3'TIF35-Xho (Tabela 3.3), respectivamente. O produto da PCR foi inserido nos sítios *Bam*HI e *Xho*I do vetor pET28a, dando origem ao vetor pET28a+HisTIF35N (Tabela 3.1). Nestas duas reações, utilizou-se a enzima *Taq* DNA Polimerase juntamente com o DNA genômico da cepa W303-1a de *S. cerevisiae* como *template*. Toda a análise de DNA está descrita a seguir.

### 3.4 - Análise de DNA

# **3.4.1 - Eletroforese**

Todas as análises de DNA foram feitas por eletroforese em gel de agarose 0,8% contendo 40 mM Tris-acetato pH 8,0 e 1 mM EDTA (Sambrook et al, 1989). As amostras foram preparadas em 2,5% ficoll tipo 400; 0,015% xileno-cianol e 0,015% azul de bromofenol e aplicadas no gel. A eletroforese foi feita em cubas modelo Horizon (Life Technologies), com tampão contendo 40 mM Tris-acetato pH 8,0 e 1 mM EDTA, sob corrente de 50 V por 60 min. A coloração do gel foi feita por 30 min, no escuro, com brometo de etídio 0,5 μg/mL.

### 3.4.2 - Extração de Plasmídeos

O DNA plasmideal foi isolado pelo método da lise alcalina (Birnboim e Doly, 1979). Para mini-preparações, as células de 2 mL de cultura foram centrifugadas a 18.000 g por 5 min e ressuspendidas em 100  $\mu$ L de 25 mM Tris-HCl pH 8,0; 10 mM EDTA e 50 mM glicose. As células foram lisadas com 200  $\mu$ L de uma solução de 1% SDS e 0,2 M NaOH. A suspensão foi neutralizada com 150  $\mu$ L de acetato de amônio 7,5 M. O agregado foi separado por centrifugação (10 min a 18.000 g) e o DNA presente no sobrenadante precipitado com 400  $\mu$ L de isopropanol. Outra centrifugação foi feita a 18.000 g por 10 min e o sedimento contendo o DNA foi lavado com etanol gelado a 70%, seco e ressuspendido em 100  $\mu$ L de TE (10 mM Tris-HCl pH 8,0 e 1 mM EDTA).

Preparações em larga escala de DNA plasmideal foram feitas utilizando-se o sistema *Plasmid Midiprep Kit* fornecido pela empresa Qiagen. As células de culturas de 100 mL foram coletadas por centrifugação a 3.000 g por 15 min, ressuspendidas em 4 mL de tampão contendo 50 mM Tris-HCl pH 8,0; 10 mM EDTA e 100 µg/mL RNase. Quatro mL de uma solução de lise (200 mM NaOH e 1% SDS) foram adicionados e a solução foi neutralizada com 4 mL de acetato de potássio 3 M pH 5,5. Após incubação no gelo por 15 min, a suspensão foi centrifugada a 4°C, por 45 min a 16.000 g. O sobrenadante foi aplicado em uma coluna previamente equilibrada com 4 mL de tampão contendo 750 mM NaCl; 50 mM MOPS pH 7,0; 15% isopropanol; 0,15% de Triton X100. A coluna foi lavada com 20 mL de 1 M NaCl; 50 mM MOPS pH 7,0 e 15% isopropanol. O DNA foi eluído em 5 mL de tampão composto por 1,25 M NaCl; 50 mM Tris-HCl pH 8,5; 15% isopropanol e posteriormente precipitado em 3,5 mL de isopropanol. A amostra foi centrifugada a 10.000 g por 30 min a 4°C. O sedimento contendo o DNA foi lavado com etanol gelado a 70%, seco e ressuspendido em TE.

# **3.4.3 - Purificação de Fragmentos de DNA e de Produtos de PCR Separados por** Eletroforese em Gel de Agarose

Fragmentos de DNA foram isolados de gel de agarose utilizando o sistema *Concert Matrix Gel Extraction System* (Life Technologies). O gel foi dissolvido em tampão L1 contendo iodeto de potássio e o DNA isolado por ligação à resina de sílica. Após incubação (50°C por 15 min), a resina foi sedimentada por centrifugação (30 seg, 10.000 g), lavada com tampão L1 e subseqüentemente com tampão L2 (receita não fornecida pelo fabricante). O DNA foi eluído em tampão TE.

Os produtos de PCR foram purificados com o auxílio do kit *QIAquick PCR Purification* (Qiagen), onde o DNA foi ligado a uma membrana de sílica-gel e, após centrifugação a 18.000 g por 1 min, foram descartadas as impurezas. A coluna foi então lavada com tampão contendo etanol e o DNA foi eluído em 10 mM Tris-HCl pH 8,0.

#### 3.4.4 - Digestão e Ligação de DNA

Tipicamente as reações de digestão de DNA foram realizadas numa proporção de 1-2 µg de DNA por unidade de enzima de restrição em volumes de 30-50 µL, com tampão indicado para cada enzima fornecido pelo fabricante. As enzimas de restrição foram obtidas das empresas Promega, Gibco e Biolabs.

As reações de ligação de fragmentos de DNA foram feitas com 100-200 ng de DNA do vetor, 200-300 ng de DNA do inserto em tampão contendo 50 mM Tris-HCl pH 7,6; 10 mM MgCl<sub>2</sub>; 1 mM ATP; 1 mM DTT e 5% PEG 8000, 1 unidade de T4 DNA Ligase (Life Technologies) num volume total de 20  $\mu$ L. Controles de ligação foram feitos como descrito acima, sem a inclusão de DNA do inserto. As ligações foram incubadas por 16 h a 16°C.

#### 3.4.5 - Seqüenciamento de DNA

O DNA dos insertos foi seqüenciado utilizando-se o sistema *Big Dye* e o seqüenciador automático ABI PRISM 377 *Genetic Analyser* (Applied Biosystems), seguindo as instruções de reação indicadas pelo fabricante.

#### 3.5 - Ensaios de Expressão de Proteínas em E. coli

#### 3.5.1 - Procedimento para Indução da Expressão e Preparo dos Extratos Celulares

Os ensaios de indução foram realizados em meio LB com o devido antibiótico e as células incubadas a 30°C sob agitação constante de 200 rpm até atingir densidade óptica a 600 nm entre 0,8 e 1,0. Nesta condição, acrescentou-se 0,5 mM de IPTG nas culturas, as quais foram incubadas por mais 4 h. As células foram coletadas por centrifugação a 3.500 g por 10 min a 4°C e posteriormente ressuspendidas em tampão A (50 mM tampão fosfato de sódio pH 7,2; 5% glicerol; 100 mM NaCl). Adicionou-se lisozima às suspensões que foram incubadas por 30 min no gelo. As suspensões foram sonicadas durante 20 segundos por oito vezes em um sonicador Branson 450 (Branson Ultrasonics Corporation), seguido de centrifugação a 20.000 g por 10 min a 4°C. Os extratos foram isolados e os sedimentos (nos quais estavam os corpos de inclusão), ressuspendidos em tampão B contendo 50 mM Tris-HCl pH 8,0; 100 mM EDTA; 25% sacarose e 2% Triton X100. Após o mesmo ciclo de sonicação descrito acima, os corpos de inclusão foram centrifugados (20.000 g, 10 min, 4°C) e ressuspendidos em tampão A. As amostras contendo a fração solúvel e insolúvel foram analisadas por eletroforese.

# 3.5.2 - Análise de Proteínas em Gel de SDS-Poliacrilamida

Todas as análises de proteína foram feitas por eletroforese em géis de poliacrilamida a 10% ou 12% (SDS-PAGE, Laemmli, 1970). O gel de separação é composto por 375 mM Tris-HCl pH 8,8; 10% ou 12% acrilamida; 0,1% SDS; 0,1% de APS e Temed (0,08%). O gel de alinhamento é composto por 126 mM Tris-HCl pH 6,8; 5% acrilamida; 0,1% SDS; 0,1% APS e Temed (0,05%). As amostras de 10  $\mu$ L foram preparadas com tampão na proporção 1:1 (1 volume de amostra para 1 volume do tampão de amostra), contendo 250 mM Tris-HCl pH 6,8; 10% SDS; 0,25% azul de bromofenol e 50% glicerol, juntamente com 2% de β-mercapto-etanol. As amostras foram incubadas a 95°C por 5 min, aplicadas no gel em uma cuba modelo Mini Protean III (BioRad) com tampão de corrida contendo 25 mM Tris base; 0,25 M de glicina e 0,1% SDS. A coloração do gel foi feita com solução Coomassie 2% e para a posterior descoloração usou-se solução descorante para Coomassie (metanol 30% e ácido acético 7,5%).

# 3.6 - Purificação de Proteínas

#### 3.6.1 - Cromatografia de Afinidade

Proteínas solúveis contendo cauda de histidinas foram isoladas de extratos celulares por cromatografia de afinidade a níquel imobilizado, utilizando-se resina Ni-NTA Superflow (Qiagen). Extratos preparados a partir de 1 L de cultura das cepas BL21-H34, BL21-H35 e BL21-H35N induzidas com IPTG, conforme descrito acima, foram incubados com a resina (1 mL para HisTif34p; 0,7 mL para HisTif35p e HisTif35N), previamente tratada com água e equilibrada no mesmo tampão da amostra, por 1 h a 4°C em um *shaker* rotatório (Glas-Col RD4512l). As

proteínas foram eluídas com tampão A (50 mM tampão fosfato de sódio pH 7,2; 5% glicerol e 100 mM NaCl) contendo concentrações crescentes de imidazol (de 5 a 200 mM).

Os complexos Tif34p/HisTif35p e Tif34p/HisTif35N foram isolados através da interação da HisTif35p ou HisTif35N, as quais possuem poli-histidina, com a resina Ni-NTA. Nestes experimentos, culturas de cepas BL21-34H35 e BL21-34H35N foram induzidas com IPTG. O preparo dos extratos e a purificação por afinidade seguiram o mesmo protocolo descrito acima.

Para o isolamento do complexo Tif34p/HisTif35N reconstituído, a Tif34p e a HisTif35N foram submetidas a ensaios de co-sedimentação. As proteínas purificadas separadamente foram reunidas na proporção aproximada de 1:1, incubadas por 30 min no gelo e co-sedimentadas por afinidade. Após incubação das proteínas com a resina Ni-NTA, esta foi lavada 3 vezes com tampão A descrito anteriormente, seguido de eluição com tampão A contendo 200 mM imidazol. Paralelamente, foram feitos controles com as proteínas Tif34p e HisTif35N separadamente.

#### 3.6.2 - Cromatografia de Troca-Iônica

Após a purificação por afinidade, as frações com maior concentração das proteínas de interesse, foram dialisadas e submetidas à cromatografia líquida de troca iônica utilizando o sistema AKTA-FPLC (Amersham-Biosciences). A HisTif34p foi dialisada, centrifugada (15 min, 4°C, 20.000 g) e fracionada em coluna Q-Sepharose HP de 5 mL (Amersham-Biosciences) utilizando-se tampão A (20 mM tampão fosfato de sódio pH 7,2; 50 mM NaCl; 0,5 mM EDTA; 0,5 mM DTT; 0,05% *tween20*) e tampão B (20 mM tampão fosfato de sódio pH 7,2; 1 M NaCl; 0,5 mM EDTA; 0,5 mM DTT e 0,05% *tween20*). Já as proteínas HisTif35p e HisTif35N foram submetidas à cromatografia por troca-iônica em coluna Mono-Q de 1 mL (Amersham-Biosciences), fazendo uso do tampão A composto por 10 mM tampão fosfato de sódio pH 7,2; 20 mM NaCl; 0,5 mM DTT; 5% glicerol e do tampão B, composto por 10 mM tampão fosfato de

sódio pH 7,2; 1 M NaCl; 0,5 mM DTT e 5% glicerol, sendo este responsável pelo gradiente iônico. Todas as amostras de proteína foram mantidas no gelo e a 4°C.

A Tif34p expressa na forma nativa teve seu isolamento parcial do extrato celular através de duas etapas de purificação por troca-iônica. Primeiramente, o extrato bruto foi filtrado em membrana estéril Durapore (Millipore) de 0,45 µm e fracionado em coluna de 45 cm<sup>3</sup> de DEAE-Sepharose FF (Amersham-Biosciences). Para esta etapa, foi utilizado tampão A contendo 10 mM Tris-HCl pH 8,0; 20 mM NaCl; 0,5 mM DTT; 5% glicerol e tampão B contendo 10 mM Tris-HCl pH 8,0; 1 M NaCl; 0,5 mM DTT e 5% glicerol. As frações contendo a Tif34p, foram reunidas e fracionadas em coluna Q-Sepharose HP de 5 mL, utilizando-se o tampão A (10 mM de tampão fosfato de sódio pH 7,2; 1 M NaCl; 0,5 mM DTT; 5% glicerol) e tampão B (10 mM tampão fosfato de sódio pH 7,2; 1 M NaCl; 0,5 mM DTT; 5% glicerol).

As frações mais concentradas contendo os complexos Tif34p/HisTif35p e Tif34p/HisTif35N obtidos por cromatografia de afinidade foram submetidas à cromatografia por troca iônica em coluna Q-Sepharose HP de 5 mL, utilizando-se tampão A contendo 20 mM tampão fosfato de sódio pH 7,2; 50 mM NaCl; 5% glicerol e tampão B contendo 20 mM tampão fosfato de sódio pH 7,2; 1 M NaCl e 5% glicerol. Todas as trocas de tampão das proteínas foram realizadas através de diálise a 4°C por 16 h.

## 3.6.3 - Cromatografia de Gel Filtração

Para as medidas de espalhamento de raios-X a baixo ângulo (SAXS), as frações mais puras da HisTif34p obtidas na cromatografia por troca-iônica foram reunidas, concentradas até o volume final de 250  $\mu$ L e aplicadas em uma coluna Superdex 75 HR 10/30 (Amersham-Biosciences), onde foram eliminados possíveis agregados, sob fluxo de 0,3 mL/min, utilizando-se tampão contendo 20 mM tampão fosfato de sódio pH 7,2; 100 mM NaCl e 0,5 mM DTT.

Para o isolamento dos complexos Tif34p/HisTif35p e Tif34p/HisTif35N, frações mais puras obtidas por troca-iônica foram reunidas, concentradas até o volume final de 250  $\mu$ L e aplicadas em coluna Superdex 75 HR 10/30, utilizando-se tampão contendo 10 mM tampão fosfato de sódio pH 7,2 e 10 mM NaCl, sob fluxo de 0,3 mL/min.

# 3.7 - Digestão da Poli-Histidina com Trombina

A proteína HisTif34p purificada por afinidade foi dialisada a 4°C, por 16 h em 1000 mL de tampão de clivagem contendo 20 mM Tris-HCl pH 8,4; 120 mM NaCl e 2,5 mM CaCl<sub>2</sub>. Após este período, foram adicionadas à amostra 0,5 unidade da protease trombina (Amersham-Biosciences) por mg de proteína. A reação foi incubada por 16 h a 4°C, sempre sob agitação. Subseqüentemente, a reação foi filtrada em membrana Durapore (Millipore) de 0,45 μm e submetida a fracionamento por cromatografia de afinidade a níquel imobilizado em colunas BioSpin para a remoção da poli-histidina e da fração de HisTif34p que não foi digerida pela trombina. A amostra contendo a Tif34p clivada, isto é, sem a cauda de histidinas, foi dialisada, concentrada e submetida a testes de cristalização.

# 3.8 - Concentração e Quantificação de Proteínas

Para concentrar proteínas foi utilizado o concentrador UltraFree4 (Millipore) submetido à centrifugação a 3.000 g, 4°C. A quantificação de proteínas purificadas foi realizada utilizando-se espectrofotômetro modelo DU640 (Beckman Coulter) a 280 nm. Os coeficientes de extinção molar  $\varepsilon(\lambda)$  foram calculados utilizando-se o programa ProtParam Tool (http://bo.expasy.org/tools/protparam.html) (Gill e von Hippel, 1989). Os valores de  $\varepsilon(\lambda)$  obtidos

(1,24 para HisTif34p; 0,77 para HisTif35p; 0,80 para HisTif35N) foram utilizados na equação de Beer-Lambert (A =  $\varepsilon$  x I x C), onde A é a absorbância a 280 nm, I é o comprimento do caminho óptico em cm e C, a concentração molar (mol/L).

# 3.9 - Ensaios de Proteólise Limitada

Foram realizados experimentos de proteólise limitada da proteína HisTif34p e do complexo Tif34p/HisTif35N co-purificado. A HisTif34p diluída em 10 mM de tampão fosfato de sódio pH 7,2 e 10 mM NaCl, numa concentração de 0,3 mg/mL, foi tratada com 1,35 µL de tripsina (0,1 mg/mL) (Sigma) em um volume final de reação igual a 55 µL, a 30°C. Para a análise, foram coletadas alíquotas de 10 µL antes da reação e após 10, 30, 45 e 60 min. A reação foi bloqueada com a adição de 3 mM de benzamidina (Sigma), seguida por adição de tampão de amostra Laemmli (Laemmli, 1970) e aquecida a 95°C por 5 min. As amostras foram analisadas por SDS-PAGE. Paralelamente, foram realizados experimentos de clivagem complexo do Tif34/HisTif35N co-purificado. Para esta reação, utilizou-se o complexo co-purificado isolado por gel filtração a 1 mg/mL, diluído no mesmo tampão citado acima. A amostra foi tratada com 5,6  $\mu$ L de tripsina (0,1 mg/mL) em um volume final de reação igual a 50  $\mu$ L, a 30°C. Esta reação seguiu os mesmos procedimentos descritos anteriormente. Um fragmento produto da clivagem com tripsina, de aproximadamente 30 kDa, foi utilizado para seqüenciamento do seu N-terminal. Para isso, a reação de proteólise parcial foi fracionada em gel de poliacrilamida e o fragmento transferido para uma membrana PVDF (BioRad) 9,5 x 6,5 cm através do sistema de transferência semi-seco Multiphor II (Pharmacia Biotech). A transferência foi feita em tampão contendo 39 mM glicina; 48 mM Tris base; 0,037% SDS; 20% metanol por 1 h a 10 V. Transcorrido este tempo, a membrana foi corada em Ponceau S 10X (30% ácido tricloroacético; 30% ácido

sulfosalicílico; 2% Ponceau S) por 20 min. A região correspondente à banda de interesse foi cortada e enviada para seqüenciamento do N-terminal pelo método de degradação de Edman no Laboratório de Bioquímica e Biologia Molecular do Departamento de Ciências Fisiológicas da Universidade Federal de São Carlos.

O produto da reação de proteólise limitada foi submetido à purificação por afinidade, com o objetivo de se verificar se houve a perda da poli-histidina dos fragmentos resistentes à tripsina. A HisTif34p foi submetida à digestão com tripsina conforme descrito acima. Após 30 min, adicionou-se 3 mM de benzamidina à reação, a qual foi incubada com 100 μL de resina Ni-NTA por 30 min a 4°C. A resina foi lavada 4 vezes com tampão contendo 50 mM tampão fosfato de sódio pH 7,2; 100 mM NaCl; 5% glicerol e posteriormente 4 vezes com o mesmo tampão contendo 200 mM imidazol.

### 3.10 - Espectroscopia de Dicroísmo Circular (CD)

Os experimentos de CD foram realizados num espectropolarímetro Jasco J-180 (JASCO) cuja temperatura da câmara da cubeta é controlada por sistema tipo *Peltier* PFD 425S. Os dados foram coletados utilizando-se varredura de 100 nm/min e intervalo de dados de 1 nm. Os experimentos foram realizados sob temperatura constante de 4°C, em cubeta de 0,1 cm, no comprimento de onda de 200 a 260 nm (UV-distante), obtendo-se 20 coletas de cada amostra. O programa utilizado para registro e tratamento dos dados foi o Spectra Manager (JASCO), com o auxílio do programa Origin 6.1 (OriginLab Corporation).

Os valores obtidos na leitura de CD em mgraus foram convertidos para elipticidade residual molar [ $\theta$ ] de acordo com a equação [ $\theta$ ] = ( $\theta$  x PM) / (c x 0,1 x n); onde  $\theta$  é a elipticidade em graus, PM é o peso molecular da proteína (g/mol), c é a concentração da proteína (mg/mL),

0,1 é o comprimento do caminho ótico (cm) e n o número de resíduos de aminoácidos da proteína. Através do programa CDNN (Bohm et al, 1992) foram realizadas as desconvoluções do espectro, o qual compara o espectro de CD da proteína em questão com 13, 23 ou 33 espectros de um banco de dados de proteínas que já possuem a estrutura cristalográfica determinada.

A proteína HisTif35N foi analisada na concentração de 12 µmol/L diluída em 10 mM de tampão fosfato de sódio pH 7,2 contendo 10 mM NaCl. HisTif34p e HisTif35p foram analisadas nas concentrações de 4,8 e 12 µmol/L, respectivamente, diluídas em 20 mM de tampão fosfato de sódio pH 7,2 e 50 mM NaCl. Para analisar o complexo reconstituído, as proteínas HisTif34p e HisTif35p tiveram suas concentrações igualadas em 5 µmol/L, sendo subseqüentemente reunidas e incubadas por 30 min no gelo. As medidas de CD do complexo reconstituído foram ajustadas e comparadas com as medidas obtidas para as proteínas separadamente.

Foram realizadas medidas de CD do complexo co-purificado Tif34p/HisTif35p e do complexo reconstituído HisTif34p/HisTif35p. O complexo Tif34p/HisTif35p co-purificado por afinidade e isolado por gel filtração, foi diluído em 10 mM tampão fosfato de sódio pH 7,2 e 10 mM NaCl numa concentração final de 5,6 µmol/L. As proteínas HisTif34p e HisTif35p (purificadas separadamente) numa concentração de 5,2 e 4,99 µmol/L, respectivamente, diluídas no mesmo tampão descrito acima, foram reunidas e incubadas por 30 min no gelo. As medidas de CD do complexo reconstituído foram ajustadas e comparadas com as medidas obtidas do complexo co-purificado.

#### 3.11 - Espalhamento de Luz Dinâmico (DLS)

Foram realizados ensaios de espalhamento de luz da HisTif34p e do complexo Tif34p/HisTif35N nas concentrações de 3 e 2 mg/mL, respectivamente, num volume final de 60 μL. As amostras foram previamente centrifugadas a 20.000 g, a 4°C por 15 min e analisadas a 20°C. Os experimentos foram realizados no equipamento DynaPRO-MS/X (Protein Solutions) com temperatura controlada e os dados foram obtidos através do programa DYNAMICS versão 6.2.

# 3.12 - Ensaios de Cristalização de Proteínas

As sequintes amostras foram submetidas a testes de cristalização: HisTif34p nas concentrações de 5,2 e 9,3 mg/mL, diluídas em 20 mM Tris-HCl pH 7,4 ou 10mM tampão fosfato de sódio pH 7,2 e 20 mM NaCl; Tif34p clivada com trombina na concentração de 5 mg/mL diluída nos tampões descritos anteriormente; complexos Tif34p/HisTif35p e Tif34p/HisTif35N nas concentrações de 4,3 e 5,1 mg/mL, respectivamente, diluídos em 20 mM tampão fosfato de sódio pH 7,2 e 20 mM NaCl. Para a HisTif34p e Tif34p clivada com trombina, foram utilizados os kits *Crystal Screen I* e *II* (Hampton Research), *JB Screen* de 1 a 10 (Jena Bioscience), *Grid Screen* (Hampton Research), *Quik Screen* (Hampton Research), *Wizard I* e *II* (Emerold BioStructures) através da técnica *hanging-drop*. Já para os complexos, foram utilizados os kits *Crystal Screen I* e *II* (Hampton Research). A temperatura de cristalização testada foi 21°C.

# 3.13 - Modelagem Molecular

O modelo tridimensional da HisTif34p foi construído por homologia, baseado nas coordenadas da estrutura cristalográfica da proteína Cdc4 de *S. cerevisiae* (código de acesso no PDB: 1NEX-B) (Orlicky et al, 2003). A Cdc4, possui 30% de identidade de aminoácidos com a

Tif34p, pertence à família *WD-repeat* e é formada por oito domínios WD, os quais são responsáveis pela formação das folhas-β antiparalelas.

A construção e análise dos modelos foram realizadas numa estação de trabalho Compaq AlphaServer ES 40 com a colaboração do Dr. Sérgio Oyama Júnior (LNLS). A visualização molecular e os refinamentos foram realizados através do programa InsightII (Accelrys Inc.) na estação de trabalho Octane2 (Silicon Graphes Inc.). Os modelos por homologia da HisTif34p foram gerados pelo programa MODELLER (Šali e Blundell, 1993). O programa PROCHECK (Laskowski et al, 1993) foi utilizado para obter a qualidade estereoquímica dos modelos finais. O alinhamento das seqüências primária e secundária foi ajustado manualmente até os modelos gerados atingirem uma boa qualidade estereoquímica. A predição de estrutura secundária da HisTif34p foi obtida através do programa PSIPRED (http://bioinf.cs.ucl.ac.uk/psiform.html).

# 3.14 - Espalhamento de Raios-X a Baixo Ângulo (SAXS)

Os experimentos foram realizados na linha SAXS do Laboratório Nacional de Luz Síncrotron, com a colaboração da Profa. Dr. Íris C. L. Torriani e Cristiano L. P. de Oliveira (LNLS e Instituto de Física/UNICAMP). As medidas foram feitas com feixe de raios-X monocromáticos no comprimento de onda  $\lambda$ =1,488Å e distâncias detector-amostra iguais a 934,5 e 334,3 mm. Foram coletados dados de espalhamento da proteína HisTif34p nas concentrações 1,13; 3,58; 6,88 e 10,1 mg/mL e do complexo Tif34p/HisTif35N nas concentrações 1,38; 3,24 e 4,99 mg/mL, em tampão contendo 20 mM tampão fosfato de sódio pH 7,2, 100 mM NaCl e 0,5 mM DTT. A intensidade produzida pelo tampão foi subtraída e para a determinação do peso molecular da HisTif34p e do complexo, foi usada uma amostra padrão de albumina bovina (de 66 kDa) a 5 mg/mL diluída no mesmo tampão das proteínas estudadas. Durante as exposições, as amostras foram mantidas em um porta-amostras capilar à temperatura constante de 4°C.

Os programas utilizados para a obtenção e análise dos dados foram o GNOM (Svergun, 1992) e o Crysol (Svergun et al, 1995). Para a construção dos modelos *ab initio* foram utilizados os programas CREDO (Petoukhov et al, 2002) e DAMAVER (Svergun et al, 2001) onde foi calculada a média de 10 modelos independentes.

# **4. RESULTADOS**

#### 4.1 - Vetores de Expressão

A proteína expressa a partir do vetor pET28a+HisTIF34 (Tabela 3.1) contém uma seqüência N-terminal de 32 aminoácidos, contendo seis histidinas (poli-histidina), um sítio de clivagem da trombina e alguns resíduos do vetor. O vetor pET23a+TIF34 (Tabela 3.1) codifica a Tif34p na sua forma nativa, isto é, sem a seqüência adicional de 32 resíduos.

Asano e colaboradores (1998) identificaram a formação de um complexo entre as proteínas Tif34p e Tif35p e caracterizaram a porção N-terminal da Tif35p com sendo a responsável pela interação. Com o objetivo de se obter o complexo Tif34p/Tif35p para estudos estruturais, foi construído o vetor pET28a+HisTIF35 (Tabela 3.1) para a expressão da HisTif35p. Devido ao fato de a Tif34p interagir especificamente com a porção N-terminal da proteína Tif35p e que o RRM presente na porção C-terminal desta proteína pode ser instável e flexível, foi realizada outra construção onde apenas os primeiros 456 nucleotídeos do gene *TIF35* foram inseridos no pET28a, dando origem ao vetor pET28a+HisTIF35N (Tabela 3.1). Este vetor expressa a proteína HisTif35N. As proteínas expressas a partir destes vetores contém uma seqüência de 32 resíduos adicionais na extremidade N-terminal, incluindo seis histidinas e um sítio de clivagem da trombina.

#### 4.2 - Expressão e Purificação das Proteínas HisTif34p e Tif34p

A expressão das proteínas HisTif34p e Tif34p foi feita em cepas BL21-H34 e BL21-34 (Tabela 3.2), respectivamente e induzidas de acordo com os procedimentos descritos em materiais e métodos. Como se pode observar na Figura 4.1, as proteínas HisTif34p, de aproximadamente 42 kDa e Tif34p, de aproximadamente 39 kDa, são expressas em boa

quantidade sendo encontradas tanto na forma solúvel com na forma de corpos de inclusão. A Tif34p obtida através do vetor pET23a possui expressão basal, sendo observada mesmo sem a adição de IPTG, já a HisTif34p obtida através do pET28a, é observada apenas após a adição do indutor. O controle da expressão a partir do vetor pET28a é mais eficiente, não havendo transcrição basal, pelo fato dele possuir uma cópia do gene *lac*I, que codifica o repressor do operon *lac*, sob o qual está o controle da RNA polimerase do fago T7 na cepa BL21(DE3) (Studier e Moffatt, 1986). A seqüência de aminoácidos destas proteínas está detalhada no Anexo II.



**Figura 4.1** - Análise da expressão das proteínas HisTif34p e Tif34p em SDS-PAGE a 10%, obtidas a partir de vetores diferentes (pET28a e pET23a) de *E. coli*. Indução de 250 mL de cultura feita por 4 h. M: marcador de peso molecular (kDa). A) Extrato: cepa BL21-T34. B) Extrato: cepa BL21-HT34. \* Extrato controle: cepa BL21(DE3)-pET23a. C) Corpos de inclusão: cepa BL21-T34. D) Corpos de inclusão: cepa BL21-HT34.

A HisTif34p foi purificada a partir da fração solúvel através de uma cromatografia de afinidade a metal imobilizado, seguida de uma purificação por troca-iônica. O extrato incubado com a resina de Ni-NTA foi eluído utilizando-se tampão fosfato de sódio contendo concentrações crescentes de imidazol. A análise da eficiência da purificação foi feita por eletroforese (Fig.

4.2A). As frações de maior pureza e concentração da HisTif34p foram submetidas à cromatografia de troca iônica em coluna Q-Sepharose a qual foi eluída com gradiente de NaCl. Nesta etapa, a proteína foi eluída com aproximadamente 200 mM NaCl apresentando pureza suficiente para os ensaios estruturais (Fig. 4.2B - canaleta 11). O detergente *Tween* 20 na concentração de 0,05% foi adicionado ao tampão para que a amostra não agregasse depois de concentrada. Para as medidas de SAXS, a proteína HisTif34p foi submetida à cromatografia de gel filtração, para eliminar os possíveis agregados. A HisTif34p foi utilizada nos ensaios de espalhamento de luz dinâmico (DLS), dicroísmo circular (CD), testes de cristalização, espalhamento de raios-X a baixo-ângulo (SAXS) e proteólise limitada. A fração solúvel da Tif34p teve sua purificação parcial realizada através de duas etapas de cromatografia de troca-iônica. A concentração e a pureza das amostras foram suficientes para os ensaios de co-sedimentação.



**4.2** - Análise da purificação da proteína HisTif34p em SDS-PAGE a 10%. Indução de 3 litros de cultura da cepa BL21-HT34. **A**: Purificação por cromatografia de afinidade com resina Ni-NTA. Lavagens/Eluições: 1-2: tampão com 20 mM de imidazol. 3-5: 50 mM de imidazol. 6: 100 mM de imidazol. 7-8: 200 mM de imidazol. M: marcador de peso molecular (kDa). **B**: Purificação por cromatografia de troca-iônica em coluna Q-Sepharose. S: amostra injetada na coluna. 8-14: frações eluídas com gradiente de NaCl.

#### 4.3 - Clivagem da Poli-Histidina com Trombina

Amostras da HisTif34p foram submetidas à clivagem com trombina com o objetivo de se eliminar a poli-histidina presente na sua extremidade N-terminal. Em seguida, o produto da reação foi purificado por afinidade em resina Ni-NTA, conforme descrito em materiais e métodos. O rendimento da clivagem não foi bom, onde se perdeu muita proteína. O protocolo foi reavaliado, onde algumas condições, como tempo de incubação e quantidade de protease, foram alteradas, mas o baixo rendimento continuou. O tampão de clivagem da trombina possui Tris-HCl pH 8,4 e esta condição não favoreceu a estabilidade da HisTif34p, levando à sua agregação.

#### 4.4 - Expressão e Purificação das Proteínas HisTif35p e HisTif35N

A expressão das proteínas HisTif35p e HisTif35N foi feita em cepas derivadas da BL21(DE3)*slyD*<sup>-</sup>. As cepas BL21-H35 e BL21-H35N (Tabela 3.2) foram induzidas de acordo com os procedimentos descritos em materiais e métodos. Pode-se observar na Figura 4.3 que a proteína HisTif35p, de aproximadamente 34 kDa, é parcialmente solúvel. Com o objetivo de se obter a HisTif35p para experimentos de interação com a Tif34p, a fração solúvel da HisTif35p foi purificada por afinidade com resina Ni-NTA, ocorrendo sua eluição a partir de 20 mM de imidazol (Fig. 4.4A). Foi necessária mais uma etapa de purificação, a qual foi realizada em coluna Mono-Q (Fig. 4.4B - canaletas de 13 a 15) para que alguns contaminantes fossem eliminados. A eluição da HisTif35p ocorreu com aproximadamente 200 mM de NaCl. A HisTif35p foi utilizada nas medidas de dicroísmo circular.



**Figura 4.3** - SDS-PAGE a 10% para análise da expressão da HisTif35p. Extrato da cepa BL21-HT35 induzida com IPTG por 4 horas. M: marcador de peso molecular (kDa).



**Figura 4.4** - Análise por SDS-PAGE a 12% da purificação da HisTif35p. Extrato da cepa BL21-HT35 induzido com IPTG. **A**: Purificação por cromatografia de afinidade com resina Ni-NTA. M: marcador de peso molecular (kDa). S: extrato obtido no teste de indução. Lavagens/Eluições: 1-2: tampão sem imidazol. 3-4: 5 mM de imidazol. 5-6: 20 mM de imidazol. 7-8: 50 mM de imidazol. 9: 100 M de imidazol. 10: 200 mM imidazol. **B**: Purificação por cromatografia de troca-iônica em coluna Mono-Q. S: amostra injetada na coluna. 11 - 19: frações eluídas com gradiente de NaCl.

A proteína HisTif35N de aproximadamente 20 kDa corresponde à porção N-terminal (aminoácidos 1 a 152) da proteína Tif35p e contém 32 aminoácidos adicionais incluindo seis histidinas na sua porção N-terminal. Esta proteína possui 184 aminoácidos no total e é parcialmente solúvel (Fig. 4.5). A HisTif35N foi primeiramente fracionada por cromatografia de afinidade, sendo eluída a partir de 20 mM de imidazol (Fig. 4.6A). Para aumentar a pureza da amostra, as frações mais concentradas foram fracionadas em coluna Q-Sepharose. Nesta etapa, a proteína foi eluída com aproximadamente 200 mM de NaCl (Fig. 4.6B - canaleta 11). A HisTif35N foi utilizada nos ensaios de co-sedimentação. A seqüência de aminoácidos da HisTif35p e HisTif35N está representada no Anexo III.



**Figura 4.5** - SDS-PAGE a 12% para análise da expressão da proteína HisTif35N na cepa BL21(DE3)*slyD*, induzida com IPTG por 4 horas. M: marcador de peso molecular (kDa).



**Figura 4.6** - SDS-PAGE a 12% para análise da purificação da proteína HisTif35N, obtida a partir da cepa BL21-HT35N. **A**: Purificação por afinidade com resina Ni-NTA. M: marcador de peso molecular (kDa). E: extrato obtido no teste de indução. F: extrato após passagem pela coluna. Lavagens/Eluições: 1-2: tampão sem imidazol. 3-4: 5 mM imidazol. 5-6: 20 mM imidazol. 7-8: 50 mM imidazol. 9: 100 mM imidazol. 10: 200 mM imidazol. **B**: Purificação por cromatografia de troca-iônica em coluna Mono-Q. S: amostra injetada na coluna. 11 - 14: frações eluídas com gradiente de NaCl.

# 4.5 - Co-expressão e Isolamento dos Complexos

Culturas contendo as cepas BL21-34H35 e BL21-34H35N (Tabela 3.2) foram induzidas e os extratos preparados de acordo como descrito em materiais e métodos. A co-purificação foi uma das alternativas para se obter o complexo. Extratos das cepas descritas acima foram submetidos à cromatografia de afinidade em resina Ni-NTA obtendo-se já nesta etapa um grau de pureza relativamente alto (Fig. 4.7). A proteína Tif34p, que não possui a cauda de histidinas, ficou retida em uma coluna de afinidade, devido a sua interação com a HisTif35p ou HisTif35N, as quais possuem a poli-histidina. Subseqüentemente, os complexos Tif34p/HisTif35p e Tif34p/HisTif35N, foram fracionados em coluna Q-Sepharose para a eliminação dos contaminantes restantes.



**Figura 4.7 -** Análise por SDS-PAGE a 12% da purificação por afinidade dos complexos a partir das cepas BL21-T34HT35 e BL21-T34HT35N. **A**: Tif34p/HisTif35p e **B**: Tif34p/HisTif35N. M: marcador de peso molecular (kDa). F: extrato após passagem pela coluna. 1 e 2: tampão sem imidazol. 3 e 4: 5 mM imidazol. 5 e 6: 20 mM imidazol. 7 e 8: 50 mM imidazol. 9: 100 mM imidazol. 10-11: 200 mM imidazol.

Para isolar os complexos, as frações da cromatografia em coluna Q-Sepharose foram submetidas à cromatografia por gel filtração em coluna Superdex 75. Nota-se nos cromatogramas (Fig. 4.8) que o tempo de retenção do complexo Tif34p/HisTif35N é de 30,45 minutos e do complexo Tif34p/HisTif35p é de 29,95 minutos. A Figura 4.9 traz a pureza das amostras e indica que nos complexos, as proteínas constituintes estão aparentemente na proporção 1:1. A concentração e pureza das amostras de ambos os complexos são suficientes para os ensaios estruturais. O complexo Tif34p/HisTif35p foi utilizado em medidas de CD e o complexo Tif34p/HisTif35N foi utilizado em medidas de SAXS e DLS. Além destes experimentos, os complexos foram utilizados nos testes de cristalização.



**Figura 4.8** - Perfil cromatográfico dos complexos fracionados em coluna Superdex 75 HR 10/30. Os valores acima dos picos são correspondentes aos tempos de retenção na coluna em minutos. **A**: Complexo Tif34p/HisTif35p. **B**: Tif34p/HisTif35N.



**Figura 4.9** - SDS-PAGE a 12% para análise das frações obtidas no fracionamento em coluna Superdex 75. S: amostra injetada na coluna. 9, 10 e 11: frações obtidas e relacionadas nos cromatogramas anteriores.

Para verificar a hipótese de que as proteínas Tif34p e HisTif35N interagem quando são purificadas separadamente e depois reunidas, foram realizados ensaios de co-sedimentação. Neste experimento foi utilizada a proteína HisTif35N purificada e a Tif34p sem poli-histidina parcialmente purificada, conforme descrito em materiais e métodos. As proteínas foram incubadas com resina Ni-NTA e eluídas juntas com 200 mM de imidazol (Fig. 4.10). A Tif34p ficou retida na coluna devido sua interação com a HisTif35N. Nos experimentos controle contendo apenas a Tif34p (Fig. 4.10B), a proteína não se ligou na resina, sendo eluída nas lavagens sem imidazol; já a HisTif35N foi eluída apenas na presença de imidazol.



**Figura 4.10** - SDS-PAGE a 12 % para análise da co-sedimentação. A: Tif34p/HisTif35N. B: Controle Tif34p. M: marcador de peso molecular (kDa). S: amostra aplicada na coluna. F: amostra que não interagiu com a resina. L1-L3: lavagens com tampão sem imidazol. E1-E3: eluição com tampão contendo 200 mM imidazol.

# 4.6 - Caracterização por Espalhamento de Luz Dinâmico

Foram realizados ensaios de espalhamento de luz dinâmico, com o objetivo de se identificar e caracterizar a formação do complexo co-purificado. O complexo Tif34p/HisTif35N e a proteína HisTif34p foram submetidos aos ensaios de espalhamento, os quais resultaram na obtenção do raio hidrodinâmico e peso molecular. O complexo Tif34p/HisTif35N (Fig. 4.11A) apresentou um raio hidrodinâmico de 34 Å, peso molecular de 59 kDa e monodispersão da amostra. Já a proteína HisTif34p (Fig. 4.11B) apresentou um raio hidrodinâmico de 29 Å, peso molecular de 41 kDa e também monodispersão da amostra.



**Figura 4.11 -** Gráficos característicos de monodispersão, expressos em %Massa x Raio (nm) obtidos nas medidas de DLS. **A**: Tif34p/HisTif35N - apresentou raio de giro de 34 Å, peso molecular de 59 kDa e monodispersão da amostra. **B**: HisTif34p - apresentou raio de giro de 29 Å, peso molecular de 41 kDa e monodispersão.

# 4.7 - Caracterização por Proteólise Limitada

Amostras da proteína HisTif34p foram submetidas a digestão parcial com a protease tripsina. O objetivo foi determinar se a HisTif34p possui uma estrutura compacta como as de outras proteínas descritas da família *WD-repeat*. Como se pode observar na Figura 4.12A, um fragmento de aproximadamente 30 kDa e outro de 39 kDa (indicados por setas) se formaram já em 10 minutos e persistiram até 1 hora de reação. Com o objetivo de verificar se a interação com a HisTif35N interfere de alguma forma no padrão de proteólise da Tif34p, realizou-se paralelamente digestão com tripsina do complexo Tif34p/HisTif35N co-purificado. De acordo com a Figura 4.12B, não há interferência da interação no padrão de digestão da Tif34p, onde persistiram os fragmentos de aproximadamente 30 e 39 kDa. A HisTif35N foi totalmente digerida logo nos primeiros 10 minutos de reação.



**Figura 4.12 -** SDS-PAGE a 12% para análise das reações de proteólise limitada utilizando tripsina. As setas verdes indicam os principais fragmentos formados. **A**: Proteína HisTif34p. **B**: Complexo copurificado Tif34p/HisTif35N. M: marcador de peso molecular (kDa). T<sub>0</sub>: amostra antes da reação. T<sub>10</sub>: amostra após 10 minutos de reação. T<sub>30</sub>: após 30 minutos. T<sub>45</sub>: após 45 minutos. T<sub>60</sub>: após 60 minutos.

Através da cromatografia de afinidade, pode-se observar que o fragmento de aproximadamente 30 kDa não possui a poli-histidina, pois não foi retido pela resina Ni-NTA

(Fig. 4.13). Já o fragmento de aproximadamente 39 kDa, foi eluído parcialmente, sendo encontrado tanto nas primeiras lavagens, quanto nas lavagens com imidazol.

Foram feitas duas tentativas para seqüenciar o N-terminal do fragmento de aproximadamente 30 kDa com o objetivo de se identificar o local da digestão por tripsina. O melhor resultado obtido foi a seqüência de aminoácidos A I F L V (Alanina, Isoleucina, Fenilalanina, Leucina, Valina). Infelizmente, esta seqüência não é idêntica a nenhuma região da HisTif34p, o que indica que o seqüenciamento não funcionou, provavelmente devido à baixa concentração das amostras.



**Figura 4.13** - SDS-PAGE a 12 % para análise da purificação por afinidade em resina Ni-NTA da HisTif34p submetida à proteólise limitada com tripsina. M: marcador de peso molecular (kDa). A: amostra antes da digestão. F: fração que não interagiu com a resina. L1-L4: lavagens com tampão sem imidazol. E1-E4: eluições com tampão contendo 200 mM imidazol.

# 4.8 - Caracterização Estrutural por Dicroísmo Circular

As medidas de CD foram realizadas com dois objetivos: caracterização do conteúdo de estrutura secundária de cada uma das proteínas separadamente e verificação de alguma alteração de estrutura secundária quando as proteínas Tif34p e Tif35p interagem. Optou-se por medir a quantidade de estrutura secundária da HisTif34p, pois algumas proteínas da família WD não são

estáveis quando isoladas. No caso foi testado se a interação com a HisTif35p ajudaria a estabilizar a HisTif34p, e se isso resultaria na mudança da quantidade total de estruturas secundárias. Para a caracterização de estrutura secundária de cada proteína foram realizadas medidas de CD no UV-distante em comprimentos de onda de 200 a 260 nm. Observando a Figura 4.14, pode-se notar que o espectro de CD da HisTif34p apresenta uma banda negativa em aproximadamente 212 nm e uma banda positiva incomum entre 230-235 nm. Uma estimativa do conteúdo de estrutura secundária utilizando o programa CDNN (Bohm et al, 1992) indica que a HisTif34p é formada predominantemente por folhas β-antiparalelas (aproximadamente 40%), dobras-β (aproximadamente 20%) e estrutura randômica (aproximadamente 35%). Estes dados são compatíveis com o esperado para proteínas da família *WD-repeat*, cuja estrutura tridimensional representa uma β-propeller, formada por folhas-β. A banda positiva observada entre 230-235 nm se deve à presença de vários aminoácidos aromáticos ou dobras-β, componentes do domínio WD. A predição de estrutura secundária da HisTif34p está representada no Anexo IV.



**Figura 4.14** - Espectro de CD da proteína HisTif34p a 4,8 μmol/L em 20 mM tampão fosfato de sódio pH 7,2 e 50 mM NaCl a 4°C.

A desconvolução do espectro de CD no UV-distante da proteína HisTif35p (Fig. 4.15) indica uma grande quantidade de estrutura randômica (aproximadamente 40%), além de 20% de dobras- $\beta$  e em torno de 32% de folhas- $\beta$ . Esta grande quantidade de estrutura randômica pode se dever ao fato de a HisTif35p possuir um domínio de interação com RNA na sua extremidade Cterminal. Estes domínios, em geral quando não ligados ao RNA apresentam uma estrutura flexível. Quando comparado com o espectro da HisTif35p, o espectro da HisTif35N (Fig. 4.15) apresenta algumas diferenças como a migração da banda negativa de aproximadamente 206 para 204 nm. Além disso, não há mudança na banda próxima a 222 nm. A principal diferença entre as duas curvas é uma pequena diminuição (de 40% para 35%) do conteúdo de estrutura randômica no espectro da HisTif35N dada pelo CDNN, o que é compatível com a ausência do domínio RRM do C-terminal. A predição de estrutura secundária da HisTif35p e HisTif35N está representada no Anexo V.



**Figura 4.15** - Espectros de CD das proteínas HisTif35p e HisTif35N, a 4°C. **Proteína HisTif35p** - 12 μmol/L diluída em 20 mM tampão fosfato de sódio pH 7,2 e 50 mM NaCl. **Proteína HisTif35N** - 12 μmol/L diluída em 10 mM tampão fosfato de sódio pH 7,2 e 10 mM NaCl.
Com o objetivo de verificar se há alguma alteração de estrutura secundária quando as proteínas HisTif34p e HisTif35p interagem, foram realizadas medidas de CD no UV-distante Uma alteração estrutural é observada se o resultado experimental diferir consideravelmente do resultado teórico. Foram coletados espectros da HisTif34p e HisTif35p separadamente e a soma dos espectros (teórica) foi utilizada para estimar o espectro de CD das duas proteínas juntas. Estes dados foram sobrepostos aos dados obtidos para o complexo HisTif34p/HisTif35p reconstituído. Nota-se na Figura 4.16, que houve uma diferença pouco significativa de - 0,4 unidades de mgraus a 208 nm quando os gráficos teórico e experimental foram sobrepostos.



**Figura 4.16** - Sobreposição dos espectros de CD das proteínas HisTif34p, HisTif35p, do complexo reconstituído HisTif34p/HisTif35p (experimental) com o espectro calculado pelo programa Spectra Manager (teórico). Medidas no UV-distante a 4°C em 20 mM tampão fosfato de sódio pH 7,2 e 50 mM NaCl. Concentração de cada proteína: 5 µmol/L. A diferença entre os sinais a 208 nm é igual a -0,4 unidades de mgraus.

A fim de verificar se as proteínas são capazes de interagir após serem purificadas separadamente e também demonstrar se há mudanças de estrutura secundária na formação do complexo, foram realizadas medidas de dicroísmo circular do complexo co-purificado Tif34p/HisTif35p e do complexo reconstituído HisTif34p/HisTif35p, onde as proteínas foram purificadas separadamente e depois reunidas. Com as concentrações igualadas em µmol/L, foi possível fazer as comparações dos sinais obtidos. Os espectros obtidos estão sobrepostos na Figura 4.17 onde se nota que há uma pequena variação de sinal de aproximadamente - 1 unidade de mgrau nos picos a 208 nm. Os espectros não indicam uma mudança significativa na quantidade de estrutura secundária das duas proteínas após a interação. As duas proteínas purificadas separadamente interagem, conforme demonstrado por co-sedimentação (Fig. 4.10).



**Figura 4.17 -** Sobreposição dos espectros de CD do complexo co-purificado Tif34p/HisTif35p e do complexo reconstituído HisTif34p/HisTif35p com ajustes de concentração. Medidas realizadas a 4°C em 10 mM tampão fosfato de sódio pH 7,2 e 10 mM NaCl. A diferença entre os sinais a 208 nm é de aproximadamente -1 unidade de mgrau.

#### 4.9 - Ensaios de Cristalização

Foram realizados inúmeros testes de cristalização da HisTif34p e dos complexos Tif34p/HisTif35p ou Tif34p/HisTif35N, utilizando-se as condições descritas em materiais e métodos, mas nenhuma apresentou a formação de cristais. Não houve a formação de nenhum cristal de proteína, apenas de sal. Devido ao fato de não se conseguir cristais para o estudo da estrutura tridimensional, os estudos foram centrados na caracterização da estrutura terciária através de espalhamento de raios-X a baixo ângulo (SAXS) e modelagem molecular.

#### 4.10 - Modelagem Molecular

O modelo molecular da HisTif34p foi obtido através da homologia de estrutura primária e secundária com a proteína Cdc4 (Fig. 4.18). A predição de estrutura secundária da HisTif34p obtida do programa PSIPRED está representada no Anexo IV. Os traços dos carbonos  $\alpha$  do modelo molecular da HisTif34p gerado pelo programa MODELLER está representado na Figura 4.19. Como se pode ver, a poli-histidina não tem estrutura definida e aparece como uma cauda ligada à estrutura  $\beta$ -propeller. A Figura 4.20 mostra a representação esquemática da Tif34p, sem a cauda adicional de 32 aminoácidos. Nota-se a formação das folhas- $\beta$  antiparalelas circularmente arranjadas as quais formam a estrutura  $\beta$ -propeller. O modelo traz a formação de oito conjuntos de quatro fitas- $\beta$  correspondentes aos prováveis domínios WD presentes em sua estrutura primária.

TIF34P 1NEX_B	MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGR
TIF34P	MKAIK <mark>LTGHERPLTQVKYN</mark> KEGDLLFSCSKDSSAS <mark>VWYSLN</mark> GERLGTLDGHTGTIWS
1NEX_B	PQRTTLR <mark>GH</mark> MTSVITCLQFEDNYVITGA-DDKMIRVYD <mark>SIN</mark> KKFLLQLSGHDGGVWA
TIF34P	IDVDCFTKYC <mark>VTGSAD</mark> YSIKLWDVSN <mark>GQCVATWKSPV</mark> P-VKR <mark>VE</mark> FSPCGNYFLAIL-DNV
1NEX_B	LKYA-HGGIL <mark>V</mark> SGSTDRTVRVWDIKKGCCTHVFEGHN <mark>STV</mark> RCLDI <mark>VE</mark> YK-NIK <mark>Y</mark> IVTGSRDNT
TIF34P	MKNPGSINIYEIERDSATHELTKVS <mark>EE</mark> PIHKIITHE <mark>G</mark> LDAA <mark>TV</mark> AGWSTK <mark>G</mark> KYIIAGHKDGK
1NEX_B	LHVWKLPK/DYPLVFHTP <mark>EE</mark> NPYFVGVLR <mark>G</mark> HMASVR <mark>TV</mark> SGHGNIVVSGSYDNT
TIF34P	ISKY <mark>DV</mark> SNNYEYVDSIDL <mark>H</mark> EKS <mark>I</mark> SDMQFSPDLTYFITS <mark>SRDTNSFLVD</mark> VSTLQVLKKYETDCP
1NEX_B	LIVW <mark>DV</mark> AQ-MKCLYILSG <mark>H</mark> TDR <mark>I</mark> YSTIYDHERKRC <mark>I</mark> SA <mark>SMDT</mark> TIRIWDLENGELMYTLQGHT-
TIF34P	LNTAVITPLKE <mark>F IILGGGQEAKDVTTTSAN</mark> EGKFEARF <mark>YH</mark> KIFEEE <mark>IGRVQ</mark> GHF
1NEX_B	ALVGLLRLSDK <mark>F</mark> LVSAAADGSIRGWD <mark>AN</mark> DY-SRKFS <mark>YH</mark> HTNLSA <mark>I</mark> TTFYVSD
TIF34P	GP <mark>L</mark> NTVAISP <mark>Q</mark> GTS <mark>X</mark> ASG- <mark>G</mark> EDGFIRLHHFEKSYFDFKYD <mark>VEK</mark> AAEAKEHMQEAN
1NEX_B	NILVSGSE-NQFNIYNLRS <mark>G</mark> KLVHANILKDADQIWSVNFKGKTLVAAVEKDGQSFLEILDFS

**Figura 4.18** - Alinhamento das seqüências de aminoácidos das proteínas Tif34p e Cdc4 (código de acesso no PDB: 1NEX\_B). **Amarelo**: aminoácidos idênticos. **Azul**: estrutura secundária tipo folhas- $\beta$ . **Preto**: estrutura randômica. **Vermelho**:  $\alpha$ -hélices. --- espaços sem alinhamento.



**Figura 4.19** - Traços dos carbonos  $\alpha$  do modelo molecular obtido para a proteína HisTif34p. Representação de 5 modelos sobrepostos.



**Figura 4.20** - A: Modelo molecular obtido para a proteína Tif34p. Nota-se que sem a poli-histidina, a proteína apresenta apenas a estrutura  $\beta$ -propeller. B: Modelo girado de aproximadamente 90° em relação ao eixo Z.

#### 4.11 - Caracterização Estrutural por Espalhamento de Raios-X a Baixo Ângulo

Experimentos de SAXS foram realizados com o objetivo de se obter dados estruturais da proteína HisTif34p e do complexo Tif34p/HisTif35N, tais como envelope externo, raio de giro, peso molecular e possíveis diferenças estruturais entre a HisTif34p sozinha e o complexo. Optouse por este tipo de análise para caracterizar a estrutura da HisTif34p e do complexo Tif34p/HisTif35N, devido ao fato destas proteínas não cristalizarem. A primeira opção, no caso da caracterização estrutural é sempre a cristalografia, visto que a resolução pode chegar em torno de 2 Å, enquanto que no caso de SAXS, a resolução é de apenas 15-20 Å. As medidas de SAXS estão resumidas na Tabela 4.1 e as amostras utilizadas estão representadas na Figura 4.21. Das medidas de longa distância obteve-se para a HisTif34p o peso molecular de 46±1 kDa e para o complexo, o peso de 62±2 kDa. As curvas de espalhamento obtidas em diferentes concentrações podem ser visualizadas na Figura 4.22.

AMOSTRA	CONCENTRAÇÃO	DISTÂNCIA
	(mg/ml)	(mm)
BSA	5,00	934,5
HisTif34p (HT133)	1,33	934,5
HisTif34p (HT358)	3,58	934,5
HisTif34p (HT688)	6,88	934,5
Tif34p/HisTif35N (COMP138)	1,38	934,5
Tif34p/HisTif35N (COMP324)	3,24	934,5
Tif34p/HisTif35N (COMP499)	4,99	934,5
BSA	5,00	334,3
HisTif34p (HT101)	10,1	334,3
Tif34p/HisTif35N (COMP499)	4,99	334,3

**Tabela 4.1** - Amostras utilizadas nas medidas de SAXS, com as respectivas concentrações e distâncias.



**Figura 4.21** - Gel SDS-PAGE a 12% das amostras utilizadas nas medidas de SAXS, as quais foram obtidas por cromatografia de gel filtração. **A**: Proteína HisTif34p. 1- amostra na concentração de 6,88 mg/mL. 2 - 10,1 mg/mL. 3 - 3,58 mg/mL. 4 - 1,13 mg/mL. **B**: Complexo co-purificado Tif34p/HisTif35N. 5 - amostra na concentração de 1,38 mg/mL. 6 - 4,99 mg/mL. 7 - 3,24 mg/mL.



**Figura 4.22** - Curvas de espalhamento de Intensidade (u.a.) *versus* q (Å<sup>-1</sup>) obtidas para as amostras em diferentes concentrações. **A**: Proteína HisTif34p a 1,33, 3,58 e 6,88 mg/mL. **B**: Complexo Tif34p/HisTif35N a 1,38, 3,24 e 4,99 mg/mL.

Unindo-se as curvas de curta e longa distância, têm-se curvas de espalhamento para um grande intervalo experimental. Utilizando-se estas curvas para o cálculo da função de distribuição de distância [p(r)], têm-se os gráficos das Figuras 4.23 e 4.24. Estes resultados revelam que a proteína HisTif34p e o complexo Tif34p/HisTif35N possuem raio de giro de aproximadamente 26 Å e 36 Å, respectivamente. Vê-se um ombro na p(r) (Fig. 4.24B - seta vermelha), indicando a existência de duas regiões bem determinadas.

Sabe-se que a HisTif34p e o complexo possuem uma grande região de sobreposição conformacional, isto é, uma região idêntica que corresponde justamente ao espaço ocupado pela Tif34p. Desta forma, é de se esperar que possuam uma identidade conformacional ao menos nesta região. Para uma melhor descrição desta identidade, partiu-se para os cálculos de modelos *ab initio* tendo como base os parâmetros obtidos através da modelagem molecular citada anteriormente. Inicialmente foi verificado se os dados referentes ao  $\beta$ -propeller eram capazes de se ajustar aos dados experimentais de SAXS. O ajuste utilizando-se o programa Crysol (Svergun

et al, 1995) está representado na Figura 4.25A. Como o ajuste é ruim, tem-se a informação de que a proteína possui conformação diferente do modelo molecular. Na realidade, o modelo utilizado para se fazer estes ajustes, não possui a seqüência adicional contendo a poli-histidina (de aproximadamente 32 aminoácidos), a qual não apresenta estrutura secundária definida e isso pode gerar as alterações da curva de SAXS. Para simular a proteína com a seqüência adicional, utilizou-se o programam CREDO (Petoukhov et al, 2002) adicionando-se uma *dummy sequence* aos dados do modelo. O programa, no processo de otimização, encontra a melhor configuração desta seqüência de modo a ter o melhor ajuste dos dados experimentais. Adicionando-se o número exato de resíduos faltantes (32 aminoácidos), o programa não conseguiu ajustar os dados experimentais. Um bom ajuste com baixo  $\chi^2$  somente foi obtido adicionando-se uma seqüência de 100 resíduos. A Figura 4.25B traz este novo ajuste da curva de SAXS aos dados da modelagem utilizando o programa Crysol.



**Figura 4.23** - Dados de SAXS da proteína HisTif34p. **A**: Curva de espalhamento ajustada pelo programa Gnom. Intensidade (u.a.) *versus* q (Å<sup>-1</sup>). Rg: raio de giro em Å. **B**: Cálculo da função de distribuição de distância. p(r) *versus* raio (Å).



**Figura 4.24** - Dados de SAXS do complexo Tif34p/HisTif35N. A: Curva de espalhamento ajustada pelo programa Gnom. Intensidade (u.a.) *versus* q (Å<sup>-1</sup>). Rg: raio de giro em Å. **B**: Cálculo da função de distribuição de distância. p(r) *versus* raio (Å).



**Figura 4.25** - Curvas de espalhamento da HisTif34p ajustadas pelo programa Crysol. Intensidade (u.a.) *versus* q (Å<sup>-1</sup>). **A**: Dados de SAXS ajustados aos dados de modelagem molecular sem a adição dos aminoácidos correspondentes à seqüência contendo a poli-histidina. **B**: Dados de SAXS ajustados aos dados de modelagem molecular com a adição dos 100 aminoácidos.

De um total de 10 modelos independentes utilizando-se o programa CREDO, obtém-se a estrutura da HisTif34p mais provável utilizando-se o programa DAMAVER (Svergun et al, 2001) (Fig. 4.26). Na Figura 4.27, está representado o modelo *ab initio* obtido por SAXS sobreposto ao modelo molecular obtido através do programa MODELLER.



**Figura 4.26** - Envelope externo de baixa resolução da HisTif34p obtido pelo programa DAMAVER. A: Modelo *ab initio* representado por esferas harmônicas idênticas. **B**: Modelo girado de aproximadamente 90° em relação ao eixo Z.



**Figura 4.27** - **A**: Sobreposição do modelo *ab initio* da HisTif34p (cinza claro) com a representação esquemática do modelo molecular (azul). **B**: Modelo gerado de aproximadamente 90° em relação ao eixo Z.

Para a modelagem do complexo Tif34p/HisTif35N utilizou-se o modelo que fornece o menor  $\chi^2$  no ajuste dos dados de SAXS e repetiu-se o procedimento descrito anteriormente, adicionando-se agora o número de resíduos igual ao da parcela ligada à Tif34p (HisTif35N). O ajuste utilizando-se o programa Crysol (Svergun et al, 1995) está representado na Figura 4.28. De um total de 10 modelos independentes obteve-se a estrutura do complexo Tif34p/HisTif35N mais provável representada na Figura 4.29. A Figura 4.30 representa o modelo *ab initio* obtido por SAXS para o complexo sobreposto ao modelo *ab initio* da HisTif34p e a estrutura  $\beta$ -propeller.



**Figura 4.28** - Curva de espalhamento do complexo Tif34p/HisTif35N ajustada pelo programa Crysol. Intensidade (u.a.) *versus* q ( $Å^{-1}$ ).



**Figura 4.29** - Envelope externo de baixa resolução do complexo Tif34p/HisTif35N obtido pelo programa DAMAVER. A: Modelo *ab initio* representado por esferas harmônicas idênticas. **B**: Modelo girado de aproximadamente 90° em relação ao eixo Z.



**Figura 4.30** - **A**: Sobreposição do modelo *ab initio* do complexo Tif34p/HisTif35N (cinza claro) com o modelo *ab initio* e modelo molecular da HisTif34p (azul). **B**: Modelos girados de aproximadamente 90° em relação ao eixo Z.

Os dados de SAXS permitiram concluir que o envelope conformacional obtido para a HisTif34p é coerente com a estrutura  $\beta$ -propeller característica das proteínas da família WDrepeat, as quais possuem domínios que se estruturam formando um anel com uma cavidade central. No caso do complexo Tif34p/HisTif35N, foi possível a identificação de uma região de grande similaridade entre os modelos, indicando que a parte da cadeia de aminoácidos que corresponde à Tif34p, mantém a mesma forma na estrutura tridimensional. Ajustando-se os dados de SAXS aos dados de modelagem molecular, foi possível identificar o provável posicionamento da  $\beta$ -propeller.

## 5. DISCUSSÃO

As proteínas recombinantes usadas neste estudo foram produzidas na cepa de *E. coli* BL21 (DE3) (Studier e Moffatt, 1986) e sua derivada BL21(DE3)*slyD*<sup>-</sup> (Yan et al, 2001) em conjunto com os vetores da série pET (Novagen). Este sistema de expressão é baseado na clonagem da seqüência codificadora da proteína de interesse sob o promotor da RNA polimerase do fago T7. O gene desta RNA polimerase, por sua vez, está integrado no genoma da cepa sob o promotor *lac*UV5, induzido por IPTG.

A expressão da Tif34p e da HisTif34p a partir dos vetores pET23a e pET28a, respectivamente, foi bastante eficiente. Devido a dificuldades com a purificação da proteína Tif34p nativa, o trabalho de caracterização foi feito utilizando-se a proteína contendo uma seqüência adicional de 32 aminoácidos na extremidade N-terminal, incluindo nesta seqüência seis histidinas, um sítio de digestão para trombina, além de alguns resíduos do vetor pET28a. A presença desta seqüência adicional pareceu não afetar a solubilidade da HisTif34p (Fig. 4.1). Apenas a cromatografia por afinidade não foi suficiente para a sua purificação, sendo necessária uma etapa de purificação por troca-iônica. Estas duas etapas mostraram-se eficientes em relação à pureza e quantidade de proteína obtida. A HisTif34p foi utilizada nos ensaios de proteólise limitada, dicroísmo circular, DLS, testes de cristalização e medidas de SAXS. Inúmeros tipos e etapas de cromatografia e também diversos tampões de corrida foram testados, entretanto não foi possível obter a Tif34p com alto grau de pureza para seu estudo estrutural. A proteína nesta forma é altamente instável e precipita rapidamente, além de não interagir eficientemente com nenhuma coluna específica. A Tif34p na sua forma nativa, parcialmente purificada por trocaiônica em colunas DEAE e Q-Sepharose foi utilizada em ensaios de co-sedimentação.

Asano e colaboradores (1998) identificaram a formação do complexo entre as proteínas Tif34p e Tif35p. Para tentar reconstituir este complexo para estudos estruturais, foi construído o vetor pET28a+HisTIF35, o qual expressa a proteína HisTif35p contendo uma seqüência de seis histidinas, além de um sítio de digestão para trombina, no seu N-terminal. Devido ao fato de a Tif34p interagir especificamente com a porção N-terminal da proteína Tif35p e que o RRM presente na porção C-terminal desta proteína pode ser instável e flexível, prejudicando estudos posteriores, foi construído um vetor de expressão para a proteína HisTif35p truncada, possuindo apenas os primeiros 184 aminoácidos incluindo as seis histidinas e o sítio de trombina no Nterminal (HisTif35N).

Em cepas BL21(DE3) e BL21-SI, as proteínas HisTif35p e HisTif35N tiveram uma expressão muito baixa e mesmo a purificação por afinidade foi pouco eficiente. As cromatografias de afinidade com extratos provenientes destas cepas sempre resultaram na copurificação de uma proteína de peso molecular aparente em SDS-PAGE de 25 kDa. Esta proteína corresponde ao produto do gene *slyD*, sendo uma peptidil-prolil-cis-isomerase que contém quinze histidinas no seu C-terminal (Roof et al, 1994). Este problema foi contornado com a obtenção da cepa BL21(DE3)*slyD*<sup>-</sup> (do Dr. Ryland F. Young da *Texas A&M University*), a qual possui uma deleção no gene não essencial *slyD* (Wülfing et al, 1994). As proteínas HisTif35p e HisTif35N são parcialmente solúveis e foram purificadas a partir de extratos da cepa BL21(DE3)*slyD*<sup>-</sup> por cromatografia de afinidade seguida por troca-iônica. As proteínas HisTif35p e a HisTif35N foram utilizadas em experimentos de CD e co-sedimentação, respectivamente.

Antes de serem iniciados os experimentos envolvendo os complexos, foram realizados alguns testes para confirmar a interação entre Tif34p e Tif35p. A co-expressão/co-purificação foi uma alternativa para se obter esta confirmação, pois a HisTif35p e a HisTif35N possuem cauda de histidina, a qual interage com o níquel da resina Ni-NTA. Foram realizados dois experimentos de co-purificação, um com a Tif34p e a HisTif35p e outro com a Tif34p e a HisTif35N. Nos dois

casos a Tif34p que não possui a poli-histidina, co-purificou com a proteína que a contém. Os complexos Tif34p/HisTif35p e Tif34p/HisTif35N co-purificados por afinidade, foram subseqüentemente fracionados por troca-iônica e por gel filtração. Após as três etapas de fracionamento, foram isoladas somente as proteínas complexadas, praticamente puras e em alta concentração. Os tempos de retenção obtidos na cromatografia de gel filtração para os complexos são coerentes, devido à diferença nos pesos moleculares existentes entre eles. O complexo Tif34p/HisTif35N possui um peso molecular de aproximadamente 59 kDa (39 kDa da Tif34p e 20 kDa da HisTif35N) e o complexo Tif34p/HisTif35D, possui aproximadamente 72 kDa (39 kDa da Tif34p e 34 kDa da HisTif35p). Pela análise em SDS-PAGE (Fig. 4.9), há evidências de que no complexo as proteínas estão na proporção 1:1 e que estes são estáveis, pois mesmo após três etapas de purificação, utilizando-se tampões e condições diferentes, as proteínas permaneceram associadas. Para testar se as proteínas interagem após terem sido purificadas separadamente, foram realizados experimentos de co-sedimentação da Tif34p e HisTif35N, onde foi possível verificar a formação do complexo entre as duas proteínas.

Ensaios de espalhamento de luz dinâmico, com a HisTif34p e o complexo Tif34p/HisTif35N co-purificado revelaram um peso molecular de 41 kDa para a HisTif34p, o qual é coerente com o peso molecular calculado para esta proteína que é de aproximadamente 42 kDa. Já para o complexo Tif34p/HisTif35N, obteve-se um peso molecular de 59 kDa, consistente com a soma dos pesos moleculares da Tif34p (39 kDa) e da HisTif35N (20 kDa). As duas amostras analisadas apresentaram monodispersão, condição necessária para ensaios estruturais.

A Tif34p pertence à família de proteínas *WD-repeat*, cujos membros estão envolvidos em uma variedade de funções celulares (Neer et al, 1994). Proteínas desta família possuem cópias de um domínio altamente conservado, denominado WD. Domínios WD foram preditos como formadores de folhas- $\beta$  e alças (Neer et al, 1994), os quais foram confirmados primeiramente através da estrutura cristalográfica da subunidade  $\beta$  da proteína G e caracterizados como formadores de uma estrutura  $\beta$ -propeller (Wall et al, 1995; Lambright et al, 1996, Sondek et al, 1996). A subunidade  $\beta$  da proteína G possui trinta e dois potenciais sítios de clivagem da tripsina, uma serino-protease que cliva preferencialmente no C-terminal de resíduos de arginina (R) e lisina (K). Apesar dos trinta e dois sítios, apenas um, localizado na superfície de uma alça é acessível à tripsina na proteína nativa (Fong et al, 1986). Todas as outras lisinas e argininas da subunidade  $\beta$  estão localizadas nas regiões formadoras de folhas- $\beta$ , onde estão protegidas por uma estrutura compacta. Garcia-Higuera e colaboradores (1996) afirmaram que se as proteínas da família WD formam uma estrutura similar a da G $\beta$ , suas propriedades hidrodinâmicas deveriam ser de proteínas globulares e compactas e deveriam ser resistentes à clivagem com tripsina. De fato, além da subunidade  $\beta$  da proteína G, outras proteínas desta família são resistentes à clivagem com tripsina mesmo possuindo em sua estrutura primária dezenas de potenciais sítios de clivagem para esta protease. Como exemplo, pode-se citar a RACK1, Sec13, LIS1, coronina e Tup1 (Garcia-Higuera et al, 1996).

A seqüência primária da Tif34p possui cinco domínios WD conservados, entretanto, há espaços entre estes domínios suficientes para a formação de outras folhas- $\beta$ . A proteína HisTif34p possui trinta e oito potenciais sítios de clivagem para protease tripsina. No experimento de proteólise limitada da HisTif34p, fragmentos de aproximadamente 30 e 39 kDa mostraram ser relativamente resistentes à proteólise. Estes dados indicam que a HisTif34p também apresenta uma estrutura compacta, sendo consistentes com a estrutura  $\beta$ -propeller. Tentativas de seqüenciamento do N-terminal da banda de aproximadamente 30 kDa não geraram dados conclusivos, pois as amostras estavam em baixa concentração, gerando picos pequenos e numerosos no seqüenciamento. Os resíduos obtidos não são idênticos à nenhuma região da estrutura primária da HisTif34p. Experimentos de proteólise parcial do complexo

Tif34/HisTif35N co-purificado revelaram que a interação entre Tif34p e HisTif35N não interfere no padrão de clivagem da Tif34p, persistindo os fragmentos de aproximadamente 30 e 39 kDa. Já a HisTif35N foi totalmente digerida (os sítios de clivagem da HisTif35N estão relacionados no Anexo V).

A análise por CD no UV-distante da proteína HisTif34p sugere uma composição de estrutura secundária predominantemente formada por folhas- $\beta$  (40%). Estes dados obtidos através do programa CDNN, são coerentes com o modelo desenvolvido para proteínas da família *WD*-*repeat*, cujos membros têm sua estrutura secundária formada predominantemente por folhas- $\beta$ , sendo estas correspondentes aos domínios WD. Observando-se a Figura 4.14, pode-se notar que este espectro possui um pico incomum por volta de 230 nm a qual tem sido atribuída à presença de pontes dissulfeto (Hider et al, 1988), interações entre resíduos aromáticos (como triptofanos, fenilalaninas e tirosinas) (Woody, 1978) ou dobras- $\beta$  (Crisma et al, 1984). HisTif34p possui cinco triptofanos, dezesseis fenilalaninas e dezesseis tirosinas em sua estrutura primária, e além disso, possui muitas dobras- $\beta$  que conectam os domínios WD, sendo estes, portanto, os prováveis responsáveis pela característica do gráfico obtido. Este perfil é semelhante ao obtido por Saxena e colaboradores (1996) na análise das propriedades físicas da proteína Sec13, a qual também pertence à família *WD-repeat*.

O espectro de UV-distante da proteína HisTif35p indica que sua estrutura secundária é formada por grande quantidade de estrutura randômica, apesar de possuir  $\alpha$ -hélices e folhas- $\beta$ . Esta proteína não faz parte de uma família de proteínas específica com a qual se possa fazer comparações estruturais. Sabe-se que HisTif35p possui um RRM na porção C-terminal e este tipo de domínio normalmente apresenta uma certa flexibilidade estrutural quando não ligado ao RNA. Isto pode explicar o relativo alto conteúdo de estrutura randômica. A proteína HisTif35N também apresentou grande quantidade de estrutura randômica, a qual pode estar relacionada com o fato de a proteína estar na forma truncada.

Caso a Tif34p perca estrutura secundária quando não associada ao complexo, isto poderia ser detectado através de medidas de dicroísmo circular. Para testar esta hipótese foram comparadas medidas de dicroísmo circular das proteínas HisTif34p e HisTif35p purificadas separadamente com medidas do complexo reconstituído utilizando-se UV-distante. Uma alteração estrutural é observada se o resultado experimental diferir consideravelmente do resultado teórico dado pelo programa Spectra Manager (Jasco). Os dados mostram que não há alteração significativa na estrutura secundária das proteínas, pois a variação de sinal existente entre o resultado teórico e o experimental corresponde a no máximo - 0,4 unidades de mgraus (Fig. 4.16) a qual está dentro da margem de erro do experimento.

A comparação das medidas de dicroísmo circular do complexo co-purificado Tif34p/HisTif35p com as do complexo reconstituído HisTif34p/HisTif35p (Fig. 4.17) revelou que os espectros apresentam uma diferença de sinal de -1 unidade de mgrau nos picos a 208 nm. Esta variação também não é muito significativa e pode ser conseqüência da variação experimental decorrente de erros de pipetagem; diferenças na pureza das amostras; presença dos 32 aminoácidos adicionas no N-terminal da HisTif34p do complexo reconstituído que não estavam presentes na Tif34p do complexo co-purificado. Estes dados mostram que não há alteração significativa na estrutura secundária das proteínas após a formação do complexo e que as proteínas são capazes de interagir após serem purificadas separadamente. O experimento de co-sedimentação excluiu a hipótese de que as duas proteínas poderiam não estar interagindo.

O modelo molecular obtido para a HisTif34p de acordo com sua identidade (30%) com a proteína Cdc4, também pertencente à família *WD-repeat*, revelou que a HisTif34p forma uma estrutura  $\beta$ -propeller, composta por oito domínios WD. Cada domínio se estrutura em folhas- $\beta$  antiparalelas compostas por quatro fitas, sendo que o primeiro e o último domínio WD interagem

fechando a estrutura circular. A análise de estrutura primária revela a presença de apenas cinco domínios WD conservados, entretanto há espaços entre estes domínios onde podem se formar outras folhas- $\beta$ , as quais foram preditas pelo programa PSIPRED (Anexo IV) e confirmadas no modelo molecular obtido.

Ensaios de cristalização da HisTif34p não resultaram na formação de cristais de proteína. As amostras utilizadas nos ensaios não apresentavam agregados, conforme estabelecido por espalhamento de luz dinâmico. Para eliminar a hipótese de que a cauda de histidinas, que normalmente não possui uma estrutura rígida, interferiu no processo de cristalização, removeu-se a poli-histidina por digestão com trombina. A digestão com trombina não foi eficiente, restando pouca quantidade de proteína clivada para os testes de cristalização. No tampão de clivagem desta protease tem Tris-HCl pH 8,4 e esta condição deixa a HisTif34p instável, favorecendo sua precipitação. A proteína Tif34p clivada com trombina também não cristalizou. Na análise da seqüência primária da Tif34p digerida, nota-se que ainda restam após a digestão, 13 aminoácidos além da estrutura  $\beta$ -propeller. Estes resíduos talvez ainda ficaram instáveis, prejudicando a formação dos cristais.

As proteínas  $\beta$ -propeller formadas por domínios WD, em geral, atuam como adaptadoras formando parte de complexos ou então formam parte de proteínas multi-domínios (Neer et al, 1994). Naranda e colaboradores (1997) observaram que células de levedura deficientes em Tif34p apresentavam uma redução em todas as subunidades do eIF3, o que levou a estes autores sugerir que a Tif34p teria um papel central na manutenção da estrutura do eIF3. Posteriormente demonstrou-se que a Tif34p forma parte de um sub-complexo do fator de tradução eIF3 que inclui a Tif35p e a Prt1p (Asano et al, 1998). A interação da Tif34p com o N-terminal da Tif35p é forte o suficiente para permitir se isolar o complexo binário a partir da co-expressão das duas proteínas em *E. coli*. Esperava-se com a co-expressão e co-purificação do complexo dar uma

estabilidade maior à Tif34p e favorecer a cristalização. A deleção do C-terminal da Tif35p, que contém uma região de reconhecimento de RNA, foi proposital para evitar que esta região também pudesse interferir com o processo de cristalização. Entretanto, os ensaios de cristalização dos complexos Tif34p/HisTif35p e Tif34p/HisTif35N não deram resultados positivos. Uma hipótese provável é a de que a formação do complexo apenas da Tif34p com a Tif35p ainda não é o suficiente, haja vista que na célula estas duas proteínas fazem parte do complexo central do eIF3 formado por mais três proteínas. A poli-histidina da HisTif35p e HisTif35N presente nos complexos co-purificados também pode ter atrapalhado na formação dos cristais. A hipótese mais provável é a de que não foi encontrada a concentração correta nem os tampões de diluição e cristalização ideais para a cristalização da Tif34p e dos complexos.

Devido ao fato de não se ter conseguido nenhum cristal para estudos da estrutura tridimensional por difração de raios-X, os experimentos foram centrados na caracterização da estrutura secundária das proteínas recombinantes, assim como no estudo de seu estado conformacional por espalhamento de raios-X a baixo ângulo (SAXS) da proteína HisTif34p e do complexo Tif34p/HisTif35N. A técnica de SAXS fornece dados estruturais de baixa resolução, que quando combinados com dados adicionais conhecidos, como modelo molecular teórico e interação entre proteínas, podem revelar importantes detalhes sobre a relação estrutura-função de uma proteína.

A técnica de SAXS permite o estudo de sistemas monodispersos, onde é possível determinar a forma e o tamanho das partículas observadas (Svergun e Koch, 2002). Para se ter uma boa região a baixos ângulos, foi realizada uma série de medidas a longa distância com diferentes concentrações de amostra, de modo a verificar possíveis efeitos de agregação ou concentração, e também determinar o peso molecular da proteína e do complexo. Nas medidas a curta distância foram utilizadas amostras mais concentradas, buscando-se uma melhor estatística na região a altos ângulos. Das medidas de longa distância obteve-se para a HisTif34p o peso

molecular de 46±1 kDa e para o complexo o valor de 62±2 kDa. Os pesos moleculares destas proteínas são respectivamente 41,79 e 58,62 kDa, estando bem próximos dos valores experimentais face as limitações da técnica. Unindo-se as curvas de curta e longa distância, têm-se curvas de espalhamento para um grande intervalo experimental. Os resultados referentes ao cálculo da função de distribuição de distâncias revelaram que a proteína HisTif34p e o complexo Tif34p/HisTif35N possuem raio de giro de aproximadamente 26 e 36 Å, respectivamente.

Para uma melhor caracterização do envelope externo da HisTif34p e do complexo Tif34p/HisTif35N, partiu-se para os cálculos de modelos ab initio tendo como base os parâmetros obtidos através da modelagem molecular discutida acima. Inicialmente foi verificado se os dados referentes ao  $\beta$ -propeller (isto é, sem a presença da seqüência adicional de trinta e dois aminoácidos) eram capazes de se ajustar aos dados experimentais de SAXS. O primeiro ajuste utilizando-se o programa Crysol ficou muito ruim (Fig. 4.25A), dando a informação de que a HisTif34p analisada por SAXS possui conformação diferente do modelo molecular. É importante ressaltar que o modelo molecular utilizado para se fazer estes ajustes, não possuía a seqüência de 32 resíduos, sendo este o provável motivo das diferenças existentes entre as curvas de SAXS e os dados da modelagem molecular. Para simular a proteína com a seqüência adicional, utilizou-se o programa CREDO adicionando-se uma dummy sequence aos dados do modelo molecular. O programa, no processo de otimização, encontra a melhor configuração desta sequência de modo a ter o melhor ajuste dos dados experimentais. Adicionando-se o número exato de resíduos faltantes (os 32 aminoácidos), o programa não conseguiu ajustar os dados experimentais. Um bom ajuste somente foi obtido adicionando-se uma seqüência de 100 resíduos (Fig. 4.25B). Devido ao fato de a seqüência adicional ao  $\beta$ -propeller possuir grande flexibilidade e levando-se em consideração a limitação da técnica de SAXS, pode acontecer que no modelo *ab initio* ele ocupe um volume médio maior do que o esperado pela sua següência. Isto explicaria a necessidade de na modelagem *ab initio* (onde não se tem a capacidade de simular um domínio flexível) se ter que aumentar o número de resíduos da parcela adicionada.

Devido à redundância intrínseca das medidas de SAXS, deve-se fazer uma série de modelos independentes e posteriormente tirar uma média destes modelos de modo a buscar o envelope conformacional mais provável. Para a proteína HisTif34p, a média para o cálculo de 10 modelos independentes, utilizando o programa DAMAVER, é mostrada na Figura 4.26. Neste modelo *ab initio*, a parte inferior que não faz parte da estrutura  $\beta$ -propeller, corresponde ao espaço ocupado pela seqüência de 32 aminoácidos além da seqüência nativa. O modelo molecular obtido para a HisTif34p com base em sua similaridade com proteína Cdc4 pode ser perfeitamente sobreposto ao anel do modelo ab initio construído (Fig. 4.27). A localização da estrutura  $\beta$ -propeller foi obtida e o efeito da falta de estrutura da seqüência de 32 aminoácidos adicionais pode ser observado. A HisTif34p e o complexo possuem uma região de sobreposição conformacional, isto é, uma região idêntica que corresponde ao espaço ocupado pela Tif34p. Desta forma, é de se esperar que possuam uma identidade conformacional ao menos nesta região. Realizando-se o mesmo processo de geração de modelo para o complexo Tif34p/HisTif35N, de um total de 10 modelos independentes, obteve-se o modelo ab initio da Figura 4.29. Estes modelos obtidos pelo programa DAMAVER foram construídos de acordo com os ajustes feitos pelo programa Crysol, o qual levou em consideração os parâmetros obtidos na modelagem molecular por similaridade entre Cdc4 e Tif34p. No modelo ab initio obtido para o complexo, vêse uma similaridade entre a sua parte inferior e o modelo ab initio obtido para a HisTif34p (Fig. 4.30). Nota-se que a região que se sobrepõe entre os dois modelos é a que corresponde a  $\beta$ propeller, a qual pode ser melhor visualizada com a sobreposição do modelo molecular obtido para a Tif34p. A parcela adicionada a  $\beta$ -propeller no modelo ab initio obtido para o complexo, corresponde à HisTif35N. Comparando-se as Figuras 4.26 e 4.29, nota-se uma diferença na parte

inferior do modelo *ab initio* gerado para a HisTif34p em relação ao modelo *ab initio* gerado para o complexo. Esta diferença provavelmente deve-se ao fato de que no complexo foi usada a Tif34p nativa, sem os 32 aminoácidos adicionais do N-terminal.

Aliando-se os resultados obtidos por dicroísmo circular, espalhamento de luz dinâmico e SAXS, obteve-se informações que nos possibilitaram definir melhor a proteína Tif34p e o subcomplexo Tif34p/HisTif35N. Os dados de CD mostraram que a HisTif34p é composta predominantemente por folhas-β, o que é consistente com a estrutura secundária dos membros da família WD-repeat. Com esta técnica, foi possível avaliar também que o complexo pode ser reconstituído após as duas proteínas serem purificadas separadamente e, além disso, deu a informação de que não há alteração significativa na estrutura secundária após a formação do complexo entre a HisTif34p e HisTif35p. Os dados de espalhamento de luz dinâmico e SAXS deram informações sobre o raio hidrodinâmico, peso molecular e estado conformacional da Tif34p e do complexo Tif34p/HisTif35N. A qualidade dos dados de SAXS foi muito boa tanto pela estatística das curvas experimentais quanto pelas medidas de controle que foram realizadas (peso molecular). Com os dados de SAXS foi possível também determinar a conformação da HisTif35N no complexo com a Tif34p. E finalmente, pelos dados de SAXS e modelagem molecular, pode-se inferir o posicionamento mais provável da  $\beta$ -propeller. Uma melhor caracterização do complexo Tif34p/Tif35p poderia ser obtida mapeando-se a região responsável pela interação entre as duas proteínas através do método do duplo-híbrido, o que complementaria os dados obtidos por SAXS. Novas medidas de SAXS utilizando-se a proteína Tif34p, sem a poli-histidina, seriam convenientes para confirmar a hipótese de que a parcela adicional a  $\beta$ propeller do modelo ab initio obtido para a HisTif34p é correspondente à sequência adicional de 32 resíduos incluindo a poli-histidina e, além disso, por espectrometria de massas seria possível identificar com precisão a seqüência de aminoácidos correspondente ao fragmento de 30 kDa resistente a digestão com tripsina.

Sabe-se que o sub-complexo Tif34p/Tif35p faz parte de um complexo central maior, que contém ainda as subunidades Tif32p, Nip1p e Prt1p, formadoras do eIF3 (Asano et al, 1998). Cada uma das três subunidades maiores (Tif32p, Nip1p e Prt1p) tem um sítio de ligação para as outras duas subunidades e a extremidade C-terminal da Prt1p adicionalmente interage com as subunidades Tif34p e Tif35p (Verlhac et al, 1997; Asano et al, 1998, 1999, 2000; Phan et al, 1998; Vernlocher et al, 1999; Valásek et al, 2001, 2002). Juntas e estáveis, estas subunidades formam um complexo que se liga à subunidade ribossomal 40S via Tif32p e Nip1p (Valásek et al, 2003). A estabilidade destas interações é estritamente necessária para que ocorra a síntese de proteínas, sendo então muito importante não só a caracterização funcional de cada subunidade, mas também sua caracterização estrutural, relacionando cada subunidade, com sua função e como elas se relacionam dentro da célula no processo da tradução.

A proteína Tif34p, como já dito, foi caracterizada como sendo responsável pela integridade, estabilidade e manutenção do fator de iniciação da tradução eIF3 (Naranda et al, 1997). Não existem relatos sobre a participação dos domínios WD na interação com as outras duas subunidades do complexo, porém estes devem ter alguma participação, baseando-se nas informações obtidas sobre a função dos domínios WD de outras proteínas da família *WD-repeat* (van Nocker e Ludwig, 2003). De acordo com o modelo molecular, o qual se sobrepõe ao envelope conformacional dado por SAXS, a Tif34p apresenta oito domínios WD, os quais estão arranjados de forma a possibilitar sua interação, com as outras proteínas. Esta característica, aliada ao fato de a Tif34p ser uma subunidade importante na manutenção e integridade do eIF3, a enquadra no segundo potencial papel das proteínas *WD-repeat* discutido anteriormente, que é de servir como um componente integral de complexos de proteínas. Com os dados obtidos neste trabalho, foi possível traçar algumas características do sub-complexo Tif34p/Tif35p,

acrescentando não só alguns dados estruturais sobre estas subunidades do eIF3, mas também acrescentando dados sobre proteínas da família *WD-repeat*.

Assim como a Tif34p, a Tif35p foi caracterizada como tendo um possível papel na organização estrutural e integridade funcional do complexo eIF3 de *S. cerevisiae* (Verlhac et al, 1997). O RRM presente na sua extremidade C-terminal não possui atividade de ligação específica a RNA e, além disto, esta extremidade não é necessária para o crescimento celular (Hinachi et al, 1999). Provavelmente, a Tif34p e a Tif35p interagem para servir como base para as outras subunidades do eIF3 interagirem entre si e interagirem com os fatores eIF5, eIF2 $\alpha$ , eIF1 e com a subunidade 40S ribossomal. Desta maneira as células podem prosseguir na correta síntese de proteínas.

Em estudos que caracterizam as interações existentes entre as subunidades do eIF3 humano, a eIF3i/TRIP-1 foi identificada como uma proteína que interage com as subunidades eIF3g e eIF3b (Block et al, 1998), as quais são ortólogas às subunidades Tif35p e Prt1p de levedura, respectivamente. Além de interagirem com subunidades ortólogas em ambos os organismos, estas proteínas possuem os domínios WD altamente similares (Asano et al, 1997). Este é mais um indício de que os domínios WD podem estar envolvidos na interação com as outras subunidades do eIF3. Estes dados indicam que provavelmente a Tif34p e a TRIP-1 estão relacionadas igualmente com a iniciação da tradução, servindo como proteínas adaptadoras, capazes de interagir com duas proteínas ao mesmo tempo, servindo como base para todo o complexo eIF3 de iniciação da tradução.

### 6. CONSIDERAÇÕES FINAIS

A caracterização estrutural da proteína Tif34p e do sub-complexo Tif34p/Tif35p acrescentou alguns dados à família de proteínas WD-repeat e ao complexo eIF3. Através de medidas de dicroísmo circular, foi possível estabelecer que não há mudanças na estrutura secundária das proteínas HisTif34p e HisTif35p quando estão complexadas e, além disso, foi possível verificar que elas são capazes de interagir mesmo após terem sido purificadas separadamente. A técnica de espalhamento de luz dinâmico permitiu a caracterização do complexo com sendo formado na proporção 1:1 e também indicou o peso molecular que é de aproximadamente 59 kDa, o qual é coerente com a soma dos pesos moleculares das duas proteínas. Para a HisTif34p foi estabelecido o peso de 41 kDa. Através de reação de proteólise limitada, foi constatado que a HisTif34p é formada por uma estrutura compacta e resistente à digestão com tripsina, mesmo quando está ligada à HisTif35N. Esta estrutura compacta denominada  $\beta$ -propeller é característica das proteínas da família WD-repeat, as quais estão envolvidas em diferentes funções celulares. Outra indicação de que a HisTif34p é formada por este tipo de estrutura foi através da modelagem molecular. O modelo gerado para a HisTif34p com base na sua homologia estrutural com a proteína Cdc4, apontou a estrutura  $\beta$ -propeller, a qual é composta por oito folhas-β antiparalelas compostas por quatro fitas cada. Adicionada a esta estrutura, está uma seqüência de 32 resíduos contendo uma cauda de seis histidinas, que não tem estrutura secundária definida, ficando oscilando na molécula em solução, conforme sugerido pela análise por SAXS. Com as medidas de SAXS, pode-se obter o envelope conformacional da HisTif34p e do complexo Tif34p/HisTif35N. Aliando-se os dados de SAXS e modelagem molecular, foi possível localizar e caracterizar a  $\beta$ -propeller e, além disso, verificar a identidade estrutural existente entre a HisTif34p e o complexo.

Tendo em vista os dados funcionais sobre a Tif34p e de outras proteínas da família *WD*repeat, pode-se inferir que os domínios WD formadores da estrutura  $\beta$ -propeller, devem estar envolvidos na interação com as outras subunidades do eIF3. Estes dados indicam que provavelmente a Tif34p está relacionada com a iniciação da tradução, servindo como proteína adaptadora, capaz de interagir não só com a Tif35p, mas também com duas proteínas ao mesmo tempo, servindo como base para todo o complexo eIF3 de iniciação da tradução. Para complementar este trabalho, seria interessante mapear a região responsável pela interação Tif34p/Tif35p e verificar se os domínios WD estão realmente envolvidos na sua função. E, além disso, fazer novas medidas de SAXS da Tif34p com sua seqüência nativa, para certificar se a parcela adicionada à estrutura  $\beta$ -propeller, obtida no modelo *ab initio*, é de fato referente à seqüência adicional contendo a poli-histidina. Por espectrometria de massas seria possível identificar com precisão a seqüência de aminoácidos correspondente ao fragmento de 30 kDa resistente à digestão com tripsina.

### 7. CONCLUSÕES

- As proteínas HisTif34p, HisTif35p e HisTif35N são parcialmente solúveis quando expressas em *E. coli*. Através de duas etapas de purificação, obteve-se as proteínas puras o suficiente para os estudos estruturais.

- Os complexos Tif34p/HisTif35p e Tif34p/HisTif35N foram obtidos por co-purificação e, além disso, as duas proteínas purificadas separadamente são capazes de interagir.

- Tif34p e HisTif35N formam complexos na proporção 1:1, conforme demonstrado por espalhamento de luz dinâmico.

- Ensaios de proteólise limitada indicam que a HisTif34p possui uma estrutura compacta, característica de proteínas *WD-repeat*. A interação com a HisTif35N não altera a estrutura da Tif34p, a qual permanece resistente à clivagem com tripsina.

- A HisTif34p é formada predominantemente por folhas-β, característica de proteínas *WDrepeat*. As proteínas HisTif35p e HisTif35N apresentaram grande quantidade de estrutura randômica. Não houve alteração significativa de estrutura secundária das proteínas após a interação.

- A modelagem molecular da HisTif34p mostrou a formação de uma estrutura  $\beta$ -propeller, composta por oito folhas- $\beta$  antiparalelas, formadas por quatro fitas cada. A seqüência adicional de trinta e dois resíduos ligada à estrutura  $\beta$ -propeller não possui estrutura definida.

- Os pesos moleculares da HisTif34p e do complexo Tif34p/HisTif35N, conforme análises de SAXS, são de 46 e 62 kDa, respectivamente. O raio de giro (26 Å para a HisTif34p e 36 Å para o complexo) e o envelope conformacional também foram determinados. No modelo *ab initio* 

da HisTif34p foi possível estabelecer o volume ocupado pela seqüência adicional de trinta e dois aminoácidos e, além disso, este modelo apresenta identidade com parte do modelo *ab initio* gerado para o complexo. O modelo molecular da HisTif34p pode ser sobreposto aos modelos *ab initio* gerados, indicando boa semelhança entre os mesmos.

- Para completar este trabalho, seria interessante fazer espectrometria de massas do fragmento de 30 kDa resistente à digestão com tripsina; realizar medidas de SAXS da Tif34p (sem a seqüência adicional de trinta e dois aminoácidos); e, além disso, fazer um ensaio de duplohíbrido entre a Tif35p e mutantes WD da Tif34p, para identificar qual dos domínios é responsável pela interação com a Tif35p.

# 8. REFERÊNCIAS BIBLIOGRÁFICAS

Ang, L. H., Chattopadhyay, S., Wei, N., Oyama, T., Okada, K., Batschauer, A., Deng, X. W. (1998) Molecular interactions between COP1 and HY5 defines a regulatory switch for light control of *Arabidopsis* development. *Molecular Cell*, **1**: 213-222.

Asano, K., Clayton, J., Shalev, A., Hinnebusch, A. G. (2000) A multifactor complex of eukaryotic initiation factors eIF1, eF2, eIF3, eIF5 and initiator tRNA<sub>met</sub> is an important translation initiation intermediate *in vivo*. *Genes e Development*, **14**: 2534-2546.

Asano, K., Krishnamoorthy, T., Phan, L., Pavitt, G. D., Hinnebusch, A. G. (1999) Conserved bipartite motifs in yeast eIF5 and eIF5Bɛ, GTPase-activating and GDP-GTP exchange factors in translation initiation, mediate binging to their common substrate eIF2. *EMBO Journal*, **18**: 1673-1688.

Asano, K., Kruzi, T. G., Merrick, W. C., Hershey, J. W. B. (1997) Conservation and diversity of eukaryotic translation initiation factor eIF3. *The Journal of Biological Chemistry*, **272**: 1101-1109.

Asano, K., Phan, L., Anderson, J., Hinnebusch, A. G. (1998) Complex formation by all five homologues of mammalian translation initiation factor 3 subunits from yeast *Saccharomyces cerevisiae*. *The Journal of Biological Chemistry*, **273**: 18573-18585.

Bandyopadhyay, A., Maitra, U. (1999) Cloning and characterization of the p42 subunit of mammalian translation initiation factor 3 (eIF3): demonstration that eIF3 interacts with eIF5 in mammalian cells. *Nucleic Acids Research*, **27**:1331-1337.

Bene, R., Hershey, J. W. B. (1976) Purification and characterization of initiation factor IF-E3 from rabbit reticulocytes. *Proceedings of the National Academy of Sciences USA*, **73**: 3005-3009.

Birnboim, H. C., Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research*, **7**: 1323-1513.

Block, K. L., Vornlocher, H. P., Hershey, J. W. B. (1998) Characterization of cDNAs Encoding the p44 and p35 Subunits of Human Translation Initiation Factor eIF3. *The Journal of Biological Chemistry*, **273**: 31901-31908.

Boelens, R., Gualerzi, C. O. (2002) Structure and function of bacterial initiation factors. *Current Protein Peptide Science*, **3**: 107-119.

Boesen, T., Mohammad, S. S., Pavitt, G. D., Andersen, G. R. (2004) Structure of the catalytic fragment of translation initiation factor 2B and identification of a critically important catalytic residue. *Journal of Biology Chemistry*, **279**: 10584-10592.

Bohm, G., Muhr, R., Jaenicke, R. (1992) Quantitative analysis of protein far UV circular dichroism spectra by neural networks. *Protein Engineering Design and Selection*, **5**: 191-195.

Browning, K. S., Gallie, D. R., Hershey, J. W. B., Hinnebusch, A. G., Maitra, U., Merrick, W. C., Norbury, C. (2001) Unified nomenclature for the subunits of eukaryotic initiation factor 3. *Trends in Biochemical Sciences*, **26**: 284-285.

Burke, D. J., Church, D. (1991) Protein synthesis requirements for nuclear division, cytokinesis and cell separation in *Saccharomyces cereveisiae*. *Molecular and Cellular Biology*, **11**: 3691-3698.

Carr, S., Penfold, C. N., Bamford, V., James, R., Hemmings, A. M. (2000) The Structure of TolB, an essential component of the tol-dependent tronslocation system, and its protein-protein interaction with translocation domain of colicin E9. *Structure*, **8**: 57-66.

Chaudhuri, J., Chowdhury, D., Maitra, U. (1999) Distinct functions of eukaryotic translation initiation factors eIF1A and eIF3 in the formation of the 40 S ribosomal preinitiation complex. *The Journal of Biological Chemistry*, **274**: 17975-17980.

Chen, R-H, Miettinem, P. J., Maruoka, E. M., Choy, L., Derynck, R. (1995) A WD-domain protein that is associated with and phosphorylated by the type II TGF- $\beta$  receptor. *Nature*, **377**: 548-552.

Chothia, C., Hubbard, T., Brenner, S., Barns, H., Merzin, A. (1997) Protein folds in the all-beta and all-alpha classes. *Annual Review of Biophysics and Biomolecular Structure*, **26**:597-627.

Choy, L., Derynck, R. (1998) The type II transforming growth factor (TGF- $\beta$ ) receptorinteracting protein TRIP-1 acts as a modulator of the TGF- $\beta$  response. *The Journal of Biological Chemistry*, **273**: 31455-31462.

Crisma, M., Fasmam, G. D., Balaram, H., Balaram, P. (1984) Peptide models for beta-turns. A circular dichroism study. *International Journal of Peptide and Protein Research*, **23**: 411-419.

Danaie, P., Wittmer, B., Altmann, M., Trachsel, H. (1995) Isolation of a protein complex containing translation initiation factor PRT1 from *Saccharomyces cerevisiae*. *The Journal of Biological Chemistry*, **270**: 4288-4292.

Fasman, G. D. (1996) Circular Dichroism and the Conformational Analysis of Biomoleculaes. New York: Plenum Press.

Feigin, L. A, Svergun, D. I. (1987) Structure Analysis by Small Angle X-Ray and Neutron Scattering. New York: Plenum Press.

Fletcher, C. M., Pestova, T. V., Hellen, C. U. T., Wagner, G. (1999) Structure and interactions of the translation initiation factor eIF1. *EMBO Journal*, **18**: 2631-2639.

Fong, H. K. W., Hurley, J. B., Hopkins, R. S., Miake-Lye, R., Johnson, M. S., Doolitle, R. F., Simon, M. I. (1986) Repetitive segmental structure of the transducin B subunit: homology with the *CDC4* gene and identification of related mRNAs. *Proceedings of the National Academy of Sciences USA*, **83**: 2162-2166.

Fraser, C. F., Lee, J. Y., Mayeur, G. L., Bushell, M., Doudna, J. A., Hershey, J. W. B. (2003) The j-subunit of human translations initiation factor eIF3 is requires for the stable binding of eIF3 and its subcomplexes to 40S ribosomal subunits in vitro. *The Journal of Biological Chemistry*, **279**: 8946-56.

Garcia-Higuera, I., Fenoglio, J. Li, Y., Lewis, C., Panchenko, M. P., Reiner, O., Smith, T. F., Neer, E. J. (1996) Folding of proteins with WD-repeats: comparison of six members of the WD-repeat superfamily to the G protein beta subunit. *Biochemistry*, **35**: 13985-13994.

Gill, S. C., von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Analytical Biochemestry*, **182**: 319-326.

Glatter, O., Kratky, O. (1982) Small Angle X-Ray Scattering. Londres: Academic Press.

Goss, D. J., Rounds, D., Harrigan, T., Woodley, C. L., Wahba, A. J. (1988) Effects of eucaryotic initiation factor 3 on eucaryotic ribosomal subunit equilibrium and kinetics. *Biochemistry*, **27**: 1489-1494.

Guo, J., Hui, D. J., Meriick, W. C., Sen G. C. (2000) A new pathway of translational regulation mediated by eukaryotic initiation factor 3. *EMBO Journal*, **19**: 6891-6899.

Hanahan, D. (1983) Studies on transformation of Escherichia coli with plasmids. *Journal of Molecular Biology*, **166**: 557-580.

Henning, K. A., Li, L., Iyer, N., MacDaniel, L. D., Reagan, M. S., Legerski, R. (1995) The cockayne syndrome group A gene encodes a WD repeat protein that interacts with CSB protein and subunit of RNA polymerase II TFIIH. *Cell*, **82**: 555-564.

Hershey, J. W. B., Merrick, W. C. (2000) Pathway and mechanism of initiation of protein synthesis. In *Translational Control of Gene Expression*, N. Sonenberg, J. W. B. Hershey, and M. B. Mathews, eds. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Hider, R. C., Kupryszewski, G., Rekowski, P., Lammek, B. (1988) Origin of the positive 225-230 nm circular dichroism band in proteins. Its application to conformational analysis. *Biophysical Chemistry*, **31**: 45-51.

Hinachi, P., Hershey, J. W. B., Vornlocher, H-P. (1999) Characterization of the p33 subunit of Eukaryotic Translation Initiation Factor-3 from *Saccharomyces cerevisiae*. *The Journal of Biological Chemistry*, **274**: 8546-8553.

Hinnebusch, A. G. (2000) Mechanism and regulation of initiator methionyl-tRNA binding to ribosomes. In *Translational Control of Gene Expression*, N. Sonenberg, J. W. B. Hershey, and M. B. Mathews, eds. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Hoecker, U., Quail, P. H. (2001) The phytochrome A-specific signalling intermediate SPA1 interacts directly with COP1, a constitutive repressor of light signaling in *Arabidopsis*. *The Journal of Biological Chemistry*, **276**: 38173-38178.

Jang, S. K., Krusslich, H. G., Nicklin, M. J., Duke, G. M., Palmenberg, A. C., Wimmer, E. (1988) A segment of 5'nontranslated region of encephalomyocarditis virus RNA directs intenal entry of ribosomes during *in vitro* translation. *Journal of Virology*, **62**: 2636-2643.

Kelly, S. M., Price, N. C. (1997) The application of circular dichroism to studies of protein folding and unfolding. *Biochimica et Biophysica Acta*, **1338**: 161-185.

Kelly, S. M., Price, N. C. (2000) The use of circular dichroism in the investigation of protein structure and function. *Current Protein and Peptide Science*, **1**: 349-384.

Kozak, M. (1989) The scanning model for translation: an update. *Journal Cell Biology*, **108**: 229-241.

Laemmli, U. K. 1970. Cleavage of Strucutral Proteins During the Assembly of the Head of Bacteriophage T4. *Nature*, **277**:680-685.

Lambright, D. G., Sondek, J., Bohm, A., Skiba, N. P., Hamm, H. E., Singler, P. B. (1996) The 2.0 Å crystal structure of a heterotrimeric G protein. *Nature*, **379**: 311-319.

Lamphear, B. J., Kirchweger, R., Skern, T., Rhoads, R. E. (1995) Mapping of functional domains in eukaryotic protein synthesis initiation factor 4G (eIF4G) with picornaviral proteases: implications for cap-dependent and cap-independent translational initiation. *Journal of Biological Chemistry*, **270**: 21975-21983.

Laskowski, R. A.; MacArthur, M. W.; Moss, D. S., Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, **26**: 283-291.

Li, D., Roberts, R. (2001) WD-repeat proteins: structure characteristics, biological function and their involvement in human diseases. *Cellular and Molecular Life Sciences*, **58**: 2085-2097.

Lin, L., Holbro, T., Alonso, G., Gerosa, D., Burger, M. M. (2001) Molecular interaction between human tumor marker protein p150, the largest subunit of eIF3, and intermediate filament protein K7. *Journal Cell Biochemistry*, **80**: 483-490.

Linder, P., Zemb, T. (1991) Neutron, X-Ray and Light Scattering: Introduction to an Investigative Tool for Colloidal and Polymeric Systems. New York: Elsevier Science Publiser.

Madrona, A. Y. e Wilson, D. K. (2004) The structure of Sky8p, aprotein regulating mRNA degradation: Implications for WD protein structure. *Protein Science*, **13**: 1557-1565.

Marchetti, A., Buttitta, F., Pellegrini, S., Bertacca, G. e Callahan, R. (2001) Reduced expression of Int-6/eIF3-p48 in human tumors. *International Journal of Oncology*, **18**: 175-179.

Marcotrigiano, J., Gingras, A. C., Sonenberg, N., Burley, S. K. (1997) Cocrystal structure of the Messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. *Cell*, **89**: 951-961.

Massague, J. (1998) TGF-β signal transduction. Annual Review of Biochemistry, 67: 753-791.
Massague, J. (2000) How cells read TGF-β signal. *Molecular and Cellularl Biology*, **1**: 169-178.

Méthot, N., Song, M. S., Sonenberg, N. (1996) A region rich in aspartic acid, arginine, tyrosine, and glycine (DRYG) mediates eukaryotic initiation factor 4B (eIF4B) self-association and interaction with eIF3. *Molecular and Cellularl Biology*, **16**: 5328-5334.

Minich, W. B., Balasta, M. L., Goss, D. J., Rhoads, R. E. (1994) Chromatographic resolution of *in vivo* phosphorylated and nonphosphorylated eukaryotic translation initiation factor eIF4E: increased cap affinity of the phosphorylated form. *Proceedings of the National Academy of Sciences USA*, **91**: 7668-7672.

Morris-Desbois, C., Retry, S., Ferro, M., Garin, J., Jalinot, P. (2001) The human protein HSPC021 interacts with Int-6 and is associated with eukaryotic translation initiation factor 3. *Journal of Biological Chemistry*, **276**: 45988-45995.

Murphy, R. M. (1997) Static and dynamic light scattering of biological macromolecules: what can we learn? *Current Opinion in Biotechnology*, **8**: 25-30.

Naranda, T., Kainuma, M., MacMillan, S., Hershey, J. W. B. (1997) The 39-kilodalton subunit of eukaryotic translation initiation factor 3 is essential for the complex's integrity and for cell viability in *S. cerevisiae*. *Molecular and Cell Biology*, **17**: 145-153.

Neer, E. J., Schmidt, C. J., Nambudripad, R., Smith, T. F. (1994) The ancient regulatory-protein family of WD-repeat proteins. *Nature*, **371**: 297-300.

Neer, E. J., Smith, T. F. (1996) G protein heterodimers: new structures propel new questions. *Cell*, **84**: 1047-1058.

Nonato, M. C., Widom, J., Clardy, J. (2002) Crystal structure of the N-terminal segment of human eukaryotic translation initiation factor 2α. *Journal of Biological Chemistry*, **277**: 17057-17061.

Nupponem, N. N., Porkka, K., Kakkola, L., Tanner, M., Persson, K., Borg, A., Isola, T., Visakorp, T. (1999) Amplification and overexpression of p40 subunit of eukaryotic translation initiation factor 3 in breast and prostate cancer. *American Journal of Pathology*, **154**: 1777-1783.

Orlicky, S., Tang, X., Willems, A., Tyers, M., Sicheri, F. (2003) Structural Basis for Phosphodependent Substrate Selection and Orientation by the SCFCdc4 Ubitiquin Ligase. *Cell*, **112**: 243-256.

Papish, A. L., Tari, L. W., Vogel, H. J. (2002) Dynamic Light Scattering Study of Calmodulin-Target Peptide Complexes. *Biophysical Journal*, **83**: 1455-1464.

Pecora, R. (1985) Dynamic Light Scattering - Aplications of Photon Correlation Spectroscopy. New York: Plenum Press.

Pelletier, J., Sonenberg, N. (1988) Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature*, **334**: 230-325.

Pelton, J. T., McLean, L. R. (2000) Spectroscopic Methods for Analysis of Protein Secondary Structure. *Analytical Biochemistry*, **277**: 167-176.

Pestova, T. V., Kolupaeva, V. G., Lomakin, I. B., Pilipenko, E. V., Shatsky, I. N., Agol, V. I., Hellen, C. U. (2001) Molecular mechanisms of translation initiation in eukaryotes. *Proceedings of the National Academy of Sciences USA*, **98**: 7029-7036.

Petoukhov, M. V., Eady, N. A. J., Brown, K. A., Svergun, D. I. (2002) Addition of Missing Loops and Domains to Protein Models by X-Ray Solution Scattering. *Biophysical Journal*, **83**: 2113-3125.

Phan, L., Zhang, X., Asano, K., Anderson J., Vornlocher, H. P., Greenberg, J. R., Qin, J. and Hinnebusch, A. G. (1998) Identification of a translation initiation factor 3 (eIF3) core complex, conserved in yeast and mammals, that interacts with eIF5. *Molecular and Cellular Biology*, **18**: 4935-4946.

Pickels, L., Roe, S. M, Hemingway, E. J., Stifani, S. Pearl, E. L. H. (2002) Crystal Strucutre of the C-terminal WD40 Repeat Domin of the Human Groucho/TLE1 Transcriptional Corepressor. *Structure*, **10**: 751-761.

Pincheira, R., Chen, Q., Zhang, J. T. (2001) Identification of a 170-kDa protein over-expressed in lung cancers. *British Journal of Cancer*, **84**: 1520-1527.

Provencher, S. W., Glöckner, J. (1981) Estimation of globular protein secondary structure from circular dichroism. *Biochemistry*, **20**: 33-37.

Raught, B. e Gingras, A-C. (1999) eIF4E activity is regulated at multiple levels. *International Journal of Biochemistry and Cell Biology*, **31**:43-57.

Robinson, R. C., Turbrdsky, K., Kaiser, D. A., Marchand, J-B., Higgs, H. N., Choe, S., Pollard,T. D. (2001) Crystal structure of Arp2/3 complex. *Science*, **294**: 1679-1684.

Roll-Mecak, A., Cao, C., Dever, T. E., Burley, S. K. (2000) X-ray structures of the universal translation initiation factor IF2/eIF5B. Conformational changes on GDP and GTP binding. *Cell*, **103**: 781-792.

Roll-Mecak, A., Shin, B. S., Dever, T. E., Burkley, S. K. (2001) Engaging the ribosome: universal IFs of translation. *Trends Biochemistry Science*, **26**: 705-709.

Roof, W. D., Hornes, S. M., Young, K. D., Young R. (1994) *slyD*, a Host Gene Required for  $\phi$ X174 Lysis, Is Related to the FK506-binding Protein Family of Peptidyl-prolyl *cis-trans*-Isomerases. *Journal of Biological Chemistry*, **269**: 2902-2910.

Rothe, M., Ko, Y., Alberts, P., Wernert, N. (2000) Eukaryotic initiation factor 3 p110 mRNA is overexpressed in testicular seminomas. *American Journal of Pathology*, **157**: 1597-1604.

Ruiz-Garcia, A. B., Sendra, R., Galiana, M., Pambranco, M., Perez-Ortin, J. E., Tordera, V. (1998) HAT1 and HAT2 proteins are components of a yeast nuclear histone acethyltransfease enzyme specific for free histone H4. *Journal of Biological Chemistry*, **273**: 12599-12605.

Safer, B., Adams, S. L., Kemper, W. K., Berry, K. W., Lloyd, M., Merrick, W. C. (1976) Purification and characterization of two initiation factors required for maximal activity of a highly fractionated globin mRNA translation system. *Proceedings of the National Academy of Sciences USA*, **73**: 2584-2588

Šali, A., Blundell, T. L. (1993) Comparative protein modeling by satisfaction of spatial restraints. *Journal of Molecular Biology*, **234**: 779-815.

Sambrook, J., Fritsch, E. F., Maniats, T. (1989) Molecular Cloning: a laboratory manual. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Samuel, C. E. (1993) The eIF-2 $\alpha$  protein kinases, regulators of translation in eukaryotes from yeast to humans. *Journal of Biological Chemistry*, **268**: 7603-7606.

Saxena, K., Gaitatzes, C., Walsh, M. T., Eck, M., Neer, E., Smith, T. (1996) Analysis of the physical properties and molecular modeling of Sec13: a WD repeat protein involved in vesicular traffic. *Biochemistry*, **35**: 15215-15221.

Schreier, M. H., Erni, B., Staehelin, T. (1977) Initiation of mammalian protein synthesis: Purification and characterization of seven initiation factors. *Journal of Molecular Biology*, **116**: 727-753.

Sheu, T-J., Schwarz, E. M., Martinez, D. A., O'Keefe, R. J., Zuscik, M. J. e Puzas, J. E. (2003) A phage display technique identifies a novel regulator of cell differentiation. *Journal of Biological Chemistry*, **278**: 438-443.

Shilo, B., Riddle, V. G. H., Pardee, A. B. (1979) Protein turn-over and cell-cycle initiation in yeast. *Experimental Cell Research*, **123**: 221-227.

Shin, B-S., Maag, D., Roll-Mecak, A., Arefin, M. S., Burley, S. K., Lorsch, J. R., Dever, T. E. (2002) Uncoupling of initiation factor eIF5B/IF2 GTPase and translational activities by mutations that lower ribosome affinity. *Cell*, **111**: 1015-1025.

Smith, T. F., Gaitatzes, C., Saxena, K., Neer, E. J. (1999) The WD repeat: a common architecture for diverse function. *Trends in Biochemical Sciences*, **24**: 181-185.

Sondek, J., Bohm, A., Lambright, D. G., Hamm, H. E., Sigler, P. B. (1996) Crystal structure of a G-protein βγ dimmer at 2.1Å resolution. *Nature*, **379**: 369-374.

Sonenberg, N., Dever, T. E. (2003) Eukaryotic translation initiation factors and regulators. *Current Opinion in Structural Biology*, **13**: 56-63.

Sprague, E. R., Redd, M. J., Johnson, A. D., Wolberger, C. (2000). Structure of the C-teminal domain Tup1, a corepressor of transcription in yeast. *EMBO Journal*, **19**: 3016-3027.

Studier, F., Moffatt, B. A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high level expression of cloned genes. *Journal of Molecular Biology*, **189**: 113-130.

Svergun, D. I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *Journal of Appllied Crystallography*, **25**: 495-503.

Svergun, D. I. (1999) Restoring Low Resolution Structure of Biological Macromolecules from Solution Scattering Using Simulated Annealing. *Biophysical Journal*, **75**: 2879-2886.

Svergun, D. I., Koch, M. H. J. (2002) Advances in structure analysis using samll-angle scattering in solution. *Current Opinion in Structural Biology*, **12**: 654-660.

Svergun, D. I., Koch, M. H. J. (2003) Small-angle scattering studies of biological macromolecules in solution. *Progress in Physics*, **66**: 1735-1782.

Svergun, D. I., Petoukhov, M. V. and Koch, M. H. J. (1995) CRYSOL - A program to evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *Journal of Appllied Crystallography*, **28**: 768-773.

Svergun, D. I., Petoukhov, M. V. and Koch, M. H. J. (2001) Determination of Domain Structure of Proteins from X-ray Solution Scattering. *Biophysical Journal*, **80**: 2946-2953.

Thompson, H. A., Sadnik, I., Scheinbuks, J., Moldave, K. (1977) Studies on native ribosomal subunits from rat liver. Purification and characterization of a ribosome dissociation factor. *Biochemistry*, **16**: 2221-2230.

Tomoo, K., Shen, X., Okabe, K., Nozoe, Y., Fukuhara, S., Morino, S., Ishida, T., Taniguchi, T., Hasegawa, H., Terashima, A. (2002) Crystal structures of 7-methylquanosine 5'triphosphate (m(7)GTP)- and P(1)-7-methylguanosine-P(3)-adenosine-5, 5'-triphosphate (1(7)GpppA)-bound human full-length eukaryotic initiation factor 4E: biological importance of the C-terminal flexible region. *Biochemical Journal*, **362**: 539-544.

Tsukiyama-Kohara, K., Iizuka, N., Kohara, M., Nomoto, A. (1992) Internal ribosome entry site within hepatitis C virus RNA. *Journal of Virology*, **66**: 1476-1483.

Tullio-Pelet, A., Salomon, R., Hadj-Rabia, S., Mugnier, C., de Laet, M. H., Chaouachi, B.,
Bakiri, F., Brottier, P., Cattolico, L., Penet, C., Begeot, M., Naville, D., Nicolino, M., Chaussain,
J. L., Weissenbach, J., Munnich, A., Lyonnet, S. (2000) Mutant WD-repeat protein in triple-A syndrome. *Nature Genetics*, 26: 332-335.

Valásek, L. Mathew, A. A., Shin, B-S, Nielsen, K. H., Szamecz, B., Hinnebusch, A. G. (2003) The yeast eIF3 subunits TIF32/a, NIP1/c, and eIF5 make critical connections with the 40S rivosome *in vivo*. *Genes e Development*, **17**: 786-799.

Valásek, L., Nielsen, K. H., Hinnebusch, A. G. (2002) Direct eIF2-eIF3 contact in the multfactor complex is important for translation initiation *in vivo*. *EMBO Journal*, **21**: 5886-5898.

Valásek, L., Phan, L., Schoenfeld, L. W., Valaskova, V., Hinnebusch, A. G. (2001) Related eIF3 subunits TIF32 and HCR1 interact with an RNA recognition motif in PRT1 required for eIF3 integrity and ribosome binding. *EMBO Journal*, **20**: 891-904.

van der Knaap, M. S., Leegwater, P. A., Konst, A. A., Visser, A., Naidu, S., Oudejans, C. B., Schutgens, R. B., Pronk, J. C. (2002) Mutations in each of the five subunits of translation initiation factor eIF2B can cause leukoencephalopathy with vanishing white matter. *Annals of Neurology*, **51**: 264–270

van der Voorn, L., Ploegh, H. L. (1992) The WD-40 repeat. FEBS Letters, 307:131-134.

van Nocker, S., Ludwig, P. (2003) The WD-repeat protein superfamily in *Arabidopsis*: conservation and divergence in structure and function. *BMC Genomics*, **4**: 50.

Verlhac, M. H., Chen, R. H., Hinachi, P. Hershey, J. W. B., Derynk, R. (1997) Identification of partners of TIF34, a component of yeast eIF3 complex, required for cell proliferation and translation initiation. *EMBO Journal*, **16**: 6812-6822.

Vernlocher, H. P., Hanachi, P., Ribeiro, S., Hershey, J. W. B. (1999) A 110-kilodalton subunit of translation initiation factor eIF3 and an associated 135-kilodalton protein are encoded by the *Saccharomyces cereveisiae TIF32* and *TIF31* genes. *Journal of Biological Chemistry*, **274**: 16802-16812.

Voegtli, W. C., Madrona, Y. and Wilson, D. K. (2003). The Structure of Aip1p, a WD Repeat Protein that Regulates Cofilin-Mediated Actin Depolymerization. *Journal of Biological Chemistry*, **278**: 34373-34379.

Wall, M. A., Coleman, D. E., Lee, E., Iniguez-Lluhi, J. A., Posner, B. A., Gilman, A. G., Sprang, S. R. (1995) The Structure of the G Protein Heterotrimer  $G_{1\alpha 1}\beta_1\gamma_2$ . *Cell*, **83**: 1047-1058.

Woody, R. W. (1978) Aromatic side chain contributions to the far ultraviolet circular dichroism of peptides and proteins. *Biopolymers*, **17**: 1451-1467.

Wulfing, C., Lombardero, J., Pluckthun, A. (1994) An *Escherichia coli* Protein Consisting of a Domain Homologous to FK506-binding Proteins (FKBP) and a New Metal Binding Motif. *Journal of Biological Chemistry*, **269**: 2895-2901.

Yan, S-Z, Beeler, J. A., Chen, Y., Shelton, R. K., Tang, W-J. (2001) The Regulation of Type 7 Adenylyl Cyclase by Its C1b Region and *Escherichia coli* Peptidylprolyl Isomerase, SlyD. *Journal of Biological Chemistry*, **276**: 8500-8506. Zhao, J., Hyman, L., Moore, C. (1999) Formation of mRNA 3'-ends in eukaryotes: mechanism, regulation and interrelationships with other steps in mRNA synthesis. *Microbiology and Molecular Biology Reviews*, **63**: 405-445.

## 9. ANEXOS

**Anexo I -** Mapas dos vetores de expressão de *E. coli*. **A**: Vetor pET28a onde foi inserida a seqüência codificadora das proteínas HisTif34p, HisTif35p e HisTif35N. **B**: Vetor pET23a onde foi inserida a seqüência codificadora da proteína Tif34p. (www.novagen.com).



**Anexo II** - Seqüência de aminoácidos da proteína HisTif34p. Os 32 resíduos adicionais à sua estrutura nativa estão sublinhados. Os domínios WD estão coloridos. Total: 347 aminoácidos.

MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRMKAIKLTGHERPLTQVKYNKEGDLL FSCSKDSSASVWYSLNGERLGTLDGHTGTIWSIDVDCFTKYCVTGSADYSIKLWDVSNG QCVATWKSPVPVKRVEFSPCGNYFLAILDNVMKNPGSINIYEIERDSATHELTKVSEEPIH KIITHEGLDAATVAGWSTKGKYIIAGHKDGKISKYDVSNNYEYVDSIDLHEKSISDMQFS PDLTYFITSSRDTNSFLVDVSTLQVLKKYETDCPLNTAVITPLKEFIILGGGQEAKDVTTTS ANEGKFEARFYHKIFEEEIGRVQGHFGPLNTVAISPQGTSYASGGEDGFIRLHHFEKSYFD FKYDVEKAAEAKEHMQEAN

**Anexo III** - Seqüência de aminoácidos da proteína HisTif35p. Os 32 resíduos adicionais à sua estrutura nativa estão sublinhados. Em vermelho: domínio de ligação a RNA. Total: 274 aminoácidos. Em azul, seqüência de aminoácidos da proteína HisTif35N. Total: 184 aminoácidos.

MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRMSEVAPEEIIENADGSRSIITYKIEDG VKYKITQKVKEVKVLEKVHKSVAERKNWHKYGSEKGSPAGPSAVTARLGEEVELRLSR NWKQAEEERIQKEKASLTKTGLQCRLCGNDHMTMNCPFKTILSELSALEDPATNEGGVE AASEEKAGQVGGAGSIPGQYVPPSRRAGARDPSSDAYRDSRERDDMCTLKIMQVNENA DENSLREELLFPFAPIPRVSVVRNKETGKSRGLAFVTFSSEEVAEQALRFLDGRGYMNLIL RVEWSKPKVKE **Anexo IV** - Predição de estrutura secundária da proteína HisTif34p. H:  $\alpha$ -hélice; E: folha- $\beta$ ; C: estrutura randômica; Conf: confiança da predição de 0 a 9; Pred: estrutura secundária predita; AA: seqüência alvo. Os sítios de clivagem da tripsina estão sublinhados (R e K). Em vermelho estão destacados os domínios WD característicos. Predição obtida através do programa PSIPRED (http://bioinf.cs.ucl.ac.uk/psiform.html).

## 

- AA: TWKSPVPVKRVEFSPCGNYFLAILDNVMKNPGSINIYEIERDSATHELTKVSEEPIHKII
- Conf: 36765428999997999999973698699998778965889983788828799994899
- AA: THEGLDAATVAGWSTKGKYIIAGHKDGKISKYDVSNNYEYVDSIDLHEKSISDMQFSPDL
- AA: TYFITSSRDTNSFLVDVSTLQVLKKYETDCPLNTAVITPLKEFIILGGGQEAKDVTTTSA
- Conf: 9991364022002100269
- Pred: EEEECCCHHHHHCCCCCCCCC
- AA: FKYDVEKAAEAKEHMQEAN

**Anexo V** - Predição de estrutura secundária da proteína HisTif35p. H:  $\alpha$ -hélice; E: folha- $\beta$ ; C: estrutura randômica; Conf: confiança da predição de 0 a 9; Pred: estrutura secundária predita; AA: seqüência alvo. Em vermelho, estão os aminoáciodos correspondentes a sequência da HisTif35N. Os sítios de clivagem da tripsina estão sublinhados (R e K). Predição obtida através do programa PSIPRED (http://bioinf.cs.ucl.ac.uk/psiform.html).

## Conf: 952011222222112366410000140231000278864245489872789999973795

- AA: MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRMSEVAPEEIIENADGSRSIITYKIEDGV
- Conf: 899998899998530058457621013226754467888854564123417888537763
- Conf: 124444445554213751257871786236217887507899987415566641123020
- AA: QAEEERIQKEKASLTKTGLQCRLCGNDHMTMNCPFKTILSELSALEDPATNEGGVEAASE
- AA: EKAGQVGGAGSIPGQYVPPSRRAGARDPSSDAYRDSRERDDMCTLKIMQVNENADENSLR
- AA: EELLFPFAPIPRVSVVRNKETGKSRGLAFVTFSSEEVAEQALRFLDGRGYMNLILRVEWS
- Conf: 787889
- Pred: CCCCCC
- AA: KPKVKE