

UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE BIOLOGIA



MARCELO JUN MURAI

“EXPRESSÃO E PURIFICAÇÃO DE PROTEÍNAS

RELACIONADAS À EPILEPSIA”

Este exemplar corresponde à redação final
da tese defendida pelo(a) candidato (a)

Marcelo Jun Murai
[Signature]
e aprovada pela Comissão Julgadora.

Tese apresentada ao Instituto de
Biologia para obtenção do Título de
Doutor em Genética e Biologia
Molecular, na área de Genética Humana
e Médica.

Orientadora: Profa. Dra. Iscia Lopes-Cendes

Campinas
2007

**FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DO INSTITUTO DE BIOLOGIA – UNICAMP**

M931e

Murai, Marcelo Jun
Expressão e purificação de proteínas relacionadas à
epilepsia / Marcelo Jun Murai. – Campinas, SP: [s.n.], 2007.

Orientadora: Iscia Teresinha Lopes-Cendes.
Tese (doutorado) – Universidade Estadual de
Campinas, Instituto de Biologia.

1. Epilepsias parciais. 3. Epilepsia mioclônica juvenil.
I. Lopes-Cendes, Iscia Terezinha. II. Universidade
Estadual de Campinas. Instituto de Biologia. III. Título.

(rcdt/ib)

Título em inglês: Expression and purification of epilepsy related proteins.

Palavras-chave em inglês: Partial epilepsies; Juvenile myoclonic epilepsy.

Área de concentração: Genética Animal e Evolução.

Titulação: Doutor em Genética e Biologia Molecular.

Banca examinadora: Iscia Teresinha Lopes-Cendes, João Alexandre Ribeiro Gonçalves Barbosa, Paula Regina Kuser Falcão, Ivan de Godoy Maia, Francisco Benedito Teixeira Pessine.

Data da defesa: 30/08/2007.

Programa de Pós-Graduação: Genética e Biologia Molecular.

Campinas, 30 de Agosto de 2007.

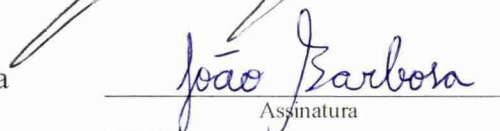
BANCA EXAMINADORA

Profa. Dra. Iscia Lopes-Cendes (Orientadora)



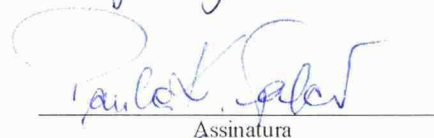
Assinatura

Prof. Dr. João Alexandre Ribeiro Gonçalves Barbosa



Assinatura

Dra. Paula Regina Kuser Falcão



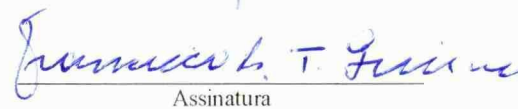
Assinatura

Prof. Dr. Ivan de Godoy Maia



Assinatura

Prof. Dr. Francisco Benedito Teixeira Pessine



Assinatura

Prof. Dra. Ljubica Tasic

Assinatura

Prof. Dr. José Camillo Novello

Assinatura

Prof. Dr. Gonçalo Amarante Guimarães Pereira

Assinatura

“De que serve ao homem ganhar o mundo inteiro, se vier a perder a sua alma?”

(S. Marcos VIII,36)

S. Inácio de Loyola a S. Francisco Xavier, antes da conversão deste último.

À minha querida esposa Luciana, e a cada um de nossos filhos,
Gabriel, André, Isabel, Maria Teresa e Maria Luísa.

AGRADECIMENTOS

A Deus Nosso Senhor Jesus Cristo, pois o fim último do homem é este: conhecer, amar e servir a Deus nesta vida e salvar, assim, sua alma.

A meus pais Akio e Nanci.

À minha mulher, por me ajudar a cumprir fielmente meu dever de estado.

Aos meus caros amigos flüelitas e montfortianos.

À professora Iscia, pela orientação, confiança e apoio.

Aos meus colegas e às técnicas do laboratório.

À professora Anete, pelo apoio nos momentos difíceis da pós-graduação.

Ao professor Ricardo Aparício.

À CAPES pelo auxílio financeiro concedido.

ÍNDICE

Resumo	xii
Abstract	xiv
1. Introdução	01
1.1 A proteína LGI1	02
1.2 A proteína EFHC1	06
2. Objetivos	12
3. A proteína LGI1	13
3.1 Metodologia	13
3.2 Resultados e discussão	30
3.3 Análises adicionais finais	44
3.4 Conclusões	49
3.5 Perspectivas	54
4. A proteína EFHC1	56
4.1 A porção N-terminal da proteína humana EFHC1	56
4.1.1 Metodologia	56
4.1.2 Resultados e discussão	61
4.2 Manuscrito “ <i>The C-terminal half of juvenile myoclonic epilepsy human EFHC1: preliminary structural analysis and dimerization evidence</i> ”	65
4.3 Resultados complementares	88
4.4 Conclusão	93
4.5 Perspectivas	94

5. Referências bibliográficas	96
6. Endereços eletrônicos	108
7. Anexos	109
Anexo 1: Códon raros da LGI1	110
Anexo 2: Predição de estrutura secundária (<i>PSIPRED</i>) da LGI1	111
Anexo 3: Tabela das proteínas estudadas, com informações básicas	112
Anexo 4: Códon raros da EFHC1N	113
Anexo 5: Predição de estrutura secundária (<i>PSIPRED</i>) da EFHC1N	114
Anexo 6: Códon raros da EFHC1C	115
Anexo 7: Predição de estrutura secundária (<i>PSIPRED</i>) da EFHC1C	116
Artigo publicado “ <i>Identification and characterization of a proteolysis resistant fragment containing the PCI domain in the Arabidopsis thaliana INT6/eIF3e translation factor</i> ”	117
Declaração de bioética	127

ÍNDICE DE FIGURAS E TABELAS

Figura 1: Esquema da proteína LGI1 humana	05
Figura 2: Organização estrutural da proteína mRib72/Efhc1 murina e EFHC1 humana	10
Figura 3: Gel SDS-PAGE do teste de purificação da proteína Trx-LGI1	32
Figura 4: Gel SDS-PAGE do teste de expressão dos clones fusionados a GST e Trx	32
Figura 5: Testes de solubilidade com as proteínas SUMO-LGI1 e SUMO-LGI1C	36
Figura 6: Teste de solubilidade com a proteína NusA-LGI1C	37
Figura 7: Teste de solubilidade com a proteína MBP-LGI1C	37
Figura 8: Teste de solubilidade com a proteína MBP-LGI1C, co-expressando GroEL	38
Figura 9: Purificação por amilose de MBP-LGI1C	40
Figura 10: Filtração em gel de MBP-LGI1C/GroEL	41
Figura 11: Amostras fracionadas em cromatografia de troca iônica	41
Figura 12: <i>Western blot</i> anti-LGI1 contra NusA-LGI1C e contra MBP-LGI1C	43
Figura 13: Gel desnaturante reduzido e não reduzido de MBP-LGI1C	43
Figura 14: Análise de DLS de MBP-LGI1C em diversos tampões	46
Figura 15: Esquema inicialmente adotado dos clones truncados de LGI1	50
Figura 16: Teste de expressão a 30°C da construção SUMO-EFHC1N	61
Figura 17: Gel SDS-PAGE da purificação da proteína SUMO-EFHC1N	62
Figura 18: Gel SDS-PAGE do teste de clivagem da proteína SUMO-EFHC1N	63
Figura 19: Curvas de SAXS das proteínas EFHC1N e EFHC1C	89
Figura 20: Envelope reconstituído da proteína de fusão SUMO-EFHC1N	90
Figura 21: Modelo <i>ab initio</i> da proteína EFHC1C	91

Tabela 1: Dados básicos das proteínas selecionadas	16
Tabela 2: Oligos sintetizados para as reações de PCR	17
Tabela 3: Construções e cepas utilizadas para GST e Trx	22
Tabela 4: Construções cepas utilizadas para SUMO, NusA e MBP.....	23
Tabela 5: Tampões para teste de solubilidade	24
Tabela 6: Amostras analisadas por DLS	44

LISTA DE ABREVIACÕES

λ	Comprimento de onda
Amp^R	Resistência à ampicilina
BCIP/NBT	(5-bromo-4-cloro-3-indolil-fosfato/nitroblue tetrazólio)
bp	Pares de base
BSA	Albumina bovina
CD	Dicroísmo circular
cv	Volumes de coluna
DLS	Espalhamento de luz dinâmico (<i>Dynamic Light Scattering</i>)
DTT	Ditiotreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ácido etilenodiaminotetracético
EMJ	Epilepsia mioclônica juvenil
EPADSA	Epilepsia parcial autossômica dominante com sintomas auditivos
FPLC	Cromatografia líquida de proteínas de rápida performance
IMAC	Cromatografia de afinidade em metal imobilizado (<i>Immobilized Metal Affinity Chromatography</i>)
IPTG	Isopropil- β -D-tiogalactopiranosídeo
Kan^R	Resistência à canamicina
kDa	10 ³ Dalton
LB	Meio Luria Bertani
MBP	Proteína ligadora de maltose, de <i>E. coli</i> (<i>Maltose-Binding Protein</i>)
Ni-NTA	Ácido Níquel-NitriloTriAcético
NusA	Substância A de utilização N, de <i>E. coli</i> (<i>N-utilization substance protein A</i>)
ORF	Seqüência codificadora (<i>Open Reading Frame</i>)
PCR	Reação em cadeia da polimerase (<i>Polimerase Chain Reaction</i>)
PDB	<i>Protein Data Bank</i>
PMSF	Fenil-metil-sulfonil-fluoreto
SAXS	Espalhamento de raios X a baixo ângulo (<i>Small Angle X-ray Scattering</i>)
SDS-PAGE	Eletroforese em gel de poliacrilamida em presença de dodecil-sulfato de sódio (<i>sodium dodecyl-sulphate poliacrilamide gel electrophoresis</i>)
SUMO	Pequeno modificador semelhante à ubiquitina, de <i>Homo sapiens</i> (<i>Small Ubiquitin-like Modifier</i>)
TEV	<i>Tobacco Etch Virus</i>
T_m	Temperatura média de transição (<i>Temperature of melting</i>)
Ulp1	<i>Ubiquitin-like-specific protease 1</i>
X-gal	5-bromo-4-cloro- β -D-galactopiranosídeo

RESUMO

Os genes *LGII* e *EFHC1* não codificam canais iônicos; entretanto, afetam indiretamente a corrente nesses canais em síndromes epiléticas determinadas geneticamente. O gene *LGII* (do inglês *Leucine-rich, glioma inactivated gene 1*) está relacionado à epilepsia parcial autossômica dominante com sintomas auditivos. Recentemente, observou-se a interação de LGI1 com ADAM22 em um complexo que possivelmente regula a transmissão sináptica. A forma mutante da LGI1 é incapaz de se ligar à ADAM22, o que fortalece a hipótese de um mecanismo relacionado à perda de função nesta síndrome epilética. O gene *EFHC1* (do inglês *EF-hand domain C-terminal containing 1*) foi encontrado mutado em algumas famílias com epilepsia mioclônica juvenil. Células transfectadas com EFHC1 apresentam maior taxa de apoptose, o que indica uma possível participação em morte celular programada. Clonamos e expressamos as proteínas humanas LGI1 em sua forma inteira (AA 1-557) e a porção C-terminal (AA 224-557; LGI1C), também chamada de domínio epitempina; as formas N-terminal (AA 78-364; EFHC1N), que compreende dois domínios DM10, e C-terminal (AA 403-640; EFHC1C) de EFHC1, formada por um domínio DM10 e um EF-hand putativo. Diversas fusões foram testadas com a proteína LGI1, seja inteira (GST, Trx, SUMO), seja a porção C-terminal (GST, Trx, NusA, MBP e SUMO), em diferentes cepas de *Escherichia coli*. Obteve-se proteína solúvel com NusA-LGI1C e MBP-LGI1C; entretanto, só foi possível a captura por cromatografia de afinidade à amilose com a fusão à MBP, na presença ou não da chaperonina GroEL. Apesar do protocolo de purificação de MBP-LGI1C ter sido estabelecido, a clivagem da cauda não foi eficiente, além de apresentar baixo rendimento. Espalhamento de luz dinâmico e SAXS mostraram que a proteína fusionada apresentava-se em um estado de forte agregação. As porções N- e C-terminal de

EFHC1 foram clonadas com a fusão SUMO, apresentando alta solubilidade em bactéria. Essas construções foram capturadas em cromatografia de afinidade a níquel e submetidas à clivagem da cauda: obteve-se rendimento próximo de 100% com EFHC1C, enquanto que com a EFHC1N não houve reação. EFHC1C clivada foi separada da cauda e passou por mais um passo de purificação, para polimento. Espectroscopia de dicroísmo circular mostrou que esta porção é composta principalmente por α -hélices e a temperatura de transição foi determinada em 54,5°C, na presença ou ausência de agente redutor DTT, indicando alta estabilidade. Espalhamento de luz dinâmico e SAXS mostraram que a proteína está monodispersa e modelagem computacional reconstituindo o envelope da EFHC1C por SAXS indica uma forma prolata. Em relação à EFHC1N, observou-se a presença de duas populações distintas de proteínas por SAXS e por filtração em gel.

ABSTRACT

Non-ion channel genes, as *LGII* and *EFHC1*, have been shown to indirectly affect ion channel currents in genetically determined epilepsy syndromes. *LGII* (Leucine-rich, glioma inactivated gene 1) is linked to a rare form of partial epilepsy (autosomal dominant partial epilepsy with auditory features, ADPEAF). Recently, LGI1 protein was associated with ADAM22 in a complex that regulates synaptic transmission. The mutated form of LGI1 is incapable of binding to ADAM22, leading to a loss of function mechanism causing ADPEAF. *EFHC1* is mutated in some families with juvenile myoclonic epilepsy (JME). It has been observed that EFHC1 transfected cells have a higher rate of apoptosis; therefore, it seems that EFHC1 protein could be involved in programmed cell death. We have successfully cloned and expressed the full form (AA 1-557) and C-terminal epitope domain of human LGI1 (AA 224-557); and the N-terminus (AA 78-364; named EFHC1N), comprising two DM10 domains in tandem, and the C-terminus portion of human EFHC1 (AA 403-640; named EFHC1C), comprising one DM10 domain and the EF-hand motif. Several fusion constructs were tested with full LGI1 (GST, Trx and SUMO) and C-terminal half (GST, Trx, NusA, MBP and SUMO) in different *Escherichia coli* strains. Soluble protein was obtained with NusA-LGI1C and MBP-LGI1C, but only MBP-LGI1C was captured by affinity amilose in the presence or absence of chaperonine GroEL. Despite the fact that we successfully established a purification protocol for MBP-LGI1C, tag cleavage presented low yield. Dynamic light scattering and SAXS showed that LGI1C fused to MBP was strongly aggregated. Furthermore, the N- and C-terminal of EFHC1 were cloned with SUMO fusion and showed high solubility in bacteria. These constructions were captured by nickel affinity chromatography and submitted to cleavage reaction. Near complete cleavage was achieved with

EFHC1C, but no cleavage was obtained with EFHC1N. SUMO tag was separated from EFHC1C and a final purification step was performed. Circular dichroism spectroscopy showed that EFHC1C is composed mainly by α -helices and transition temperature in the presence or absence of reducing agent DTT was 54.5°C, indicating high stability. Dynamic light scattering and SAXS showed that EFHC1C is in a monodisperse state and presents a prolate shape. EFHC1N presents two different populations of proteins, as determined by SAXS and size exclusion chromatography.

1. INTRODUÇÃO

As epilepsias formam um grupo de síndromes neurológicas crônicas, decorrentes de alterações das funções cerebrais, associadas ou não a outras condições patológicas. Possuem grande variedade de manifestações clínicas, etiologias, graus variados de morbidade e prognóstico. No entanto, a característica comum a todas as síndromes epiléticas é a ocorrência de crises. A crise epilética é causada por descargas neuronais anormais que ocorrem de forma passageira, sincrônica e desorganizada (Zielinski, 1988).

As epilepsias são classificadas em generalizadas, quando a descarga epilética envolve a totalidade do córtex cerebral em ambos os hemisférios, e em parciais, quando há inicialmente o envolvimento de uma área cortical restrita (*International League Against Epilepsy*, ILAE, 1989). As epilepsias generalizadas e as parciais são divididas em três categorias de acordo com a provável etiologia: idiopáticas, sintomáticas e criptogênicas (ILAE, 1989).

O termo epilepsia se refere não apenas a uma doença específica, mas a um grupo heterogêneo de desordens crônicas, caracterizado por uma tendência a ataques recorrentes. Estima-se que afete aproximadamente 1% da população mundial, portanto, é um problema de saúde pública, de acordo com a Organização Mundial de Saúde (OMS) (Sander & Shorvon, 1996). Nessa síndrome, ocorre alteração na atividade elétrica do cérebro, temporária e reversível, que produz manifestações motoras, sensitivas, sensoriais ou psíquicas. Sendo uma síndrome, ou seja, um conjunto de sinais e sintomas de diferentes etiologias e prognósticos, apresenta como característica comum a ocorrência de crises epiléticas. Além disso, pode ser adquirida ou ter causa genética em 40% dos casos.

A maior parte dos genes envolvidos com epilepsia codifica subunidades de canais iônicos, enquanto a epilepsia adquirida inclui traumas, lesões, tumores e infecções cerebrais. Dentre os raros genes não relacionados a canal iônico estão *LGII* e *EFHCI*. Descreveremos em seguida cada um deles mais detalhadamente.

1.1 A PROTEÍNA LGII

Mutações no gene humano *LGII* (*Leucine-rich Glioma Inactivated 1*) foram identificadas em indivíduos com epilepsia parcial autossômica dominante com sintomas auditivos, EPADSA (Hirose *et al.*, 2005). O primeiro relato familiar de EPADSA foi publicado em 1995, descrevendo uma grande família com epilepsia parcial (Ottman *et al.*, 1995). Este tipo de epilepsia é clinicamente caracterizado por crises parciais complexas com sintomas psíquicos, visuais e auditivos (Ottman *et al.*, 1995; Poza *et al.*, 1999; Winawer *et al.*, 2000). Algumas famílias descritas com EPADSA apresentaram ligação genética na região 10q24 (Ottman *et al.*, 1995; Pozza *et al.*, 1999; Winawer *et al.*, 2000).

Inicialmente o gene *LGII*, também chamado *Epitempina*, foi identificado rearranjado em glioblastoma, sendo sua expressão completamente ausente nesta linhagem celular (Chernova *et al.*, 1998). Devido a este fato, o gene *LGII* era um forte candidato a gene supressor de tumor. Todavia, não é considerado atualmente como tendo papel importante em tumores cerebrais (Gu *et al.*, 2005; Turnbull *et al.*, 2005; Piepoli *et al.*, 2006), pois não foram encontradas mutações na região promotora ou na região codificadora de *LGII* na linhagem utilizada (Krex *et al.*, 2002); não existem rearranjos do locus *LGII* nesta linhagem (Chernova *et al.*, 1998) e o seqüenciamento direto de glioblastomas não acusou nenhuma mutação (Gu *et al.*, 2005).

O gene *LGII* codifica duas isoformas diferentemente expressas no cérebro humano, uma curta, de aproximadamente 60 kDa, e outra longa, de 65 kDa (Furlan *et al.*, 2006), com forte expressão regional em neurônios do lobo temporal (Senechal *et al.*, 2005). Estas duas proteínas localizam-se em diferentes compartimentos subcelulares, em quantidades variáveis, sendo seu maior nível de expressão no córtex temporal lateral (Senechal *et al.*, 2005; Furlan *et al.*, 2006). Sabe-se, ainda, que a isoforma longa é secretada e que mutações impedem que a mesma seja exportada para fora da célula (Senechal *et al.*, 2005; Sirerol-Piquer *et al.*, 2006). Outros membros da família LGI (LGI2B – isoforma de LGI2 – LGI3 e LGI4) também são secretados (Senechal *et al.*, 2005) e as mutações identificadas no gene *LGII*, além de impedir a secreção da proteína, tornam a mesma instável, indicando que as bases patológicas da EPADSA estão relacionadas com perda de função da LGI1 (Senechal *et al.*, 2005).

Na porção N-terminal há quatro repetições ricas em leucina (LRR, *leucine-rich repeat*), flanqueadas por domínios ricos em cisteínas. As repetições LRRs, que consistem de α -hélices e folhas- β conectadas por alças, são encontradas em proteínas citoplasmáticas e de membrana com diferentes funções e podem servir como arcabouço estrutural para a formação de interações proteína-proteína (Kobe & Deisenhofer, 1994; 1995a; Gu *et al.*, 2005; Ko & Kim, 2007). As proteínas da família LRR estão relacionadas a diversos processos, como transdução de sinal, adesão celular, desenvolvimento celular, reparo de DNA e processamento de RNA (Kobe & Deisenhofer, 1995b).

Na porção C-terminal há sete cópias de uma repetição interna de 40-43 resíduos associada à epilepsia (EARs, *epilepsy-associated repeats*), também chamada repetição epitempina (EPTP), que se enovela provavelmente em uma estrutura do tipo β -*propeller* (Scheel *et al.*, 2002; Staub *et al.*, 2002). Este arranjo estrutural define a família gênica LGI e as repetições EPTP podem estar relacionadas a interações proteína-proteína (Steinlein, 2004). Além

destes genes que possuem similaridade estrutural com o *LGII*, outros dois genes, *VLGR1* e *TNEP1*, também codificam proteínas que contêm números variáveis de repetições EPTP (Staub *et al.*, 2002). A proteína MASS1/VLGR1 apresenta-se mutada em epilepsia audiogênica de camundongo (Skradski *et al.*, 2001; Scheel *et al.*, 2002; Staub *et al.*, 2002) e mutações no gene correspondente também são causa rara de crises febris em humanos (Nakayama *et al.*, 2002). A presença destas repetições tanto na *LGII* quanto na MASS1/VLGR1 sugere um envolvimento direto na epileptogênese, apesar da natureza de tal envolvimento permanecer apenas alusiva (Steinlein, 2003; Gu *et al.*, 2005).

Adicionalmente, a *LGII* parece ter papel importante na supressão da produção das metaloproteinases (MMPs) 1 e 3, através da via de fosfatidil-inositol 3-kinase/ERK (Kunapuli *et al.*, 2004). As MMPs estão relacionadas ao controle da proliferação e neovascularização (Crawford & Matrisian, 1994; Nabeshima *et al.*, 2002). Todavia, por que as mutações na *LGII* afetam o córtex temporal permanece completamente desconhecido, em particular sua relevância para o córtex auditório.

A maioria das mutações encontradas altera enormemente a estrutura primária da proteína *LGII* (Gu *et al.*, 2005; Ayerdi-Izquierdo *et al.*, 2006). Das dezenove mutações publicadas na literatura, oito tornam a proteína truncada, seja alterando a pauta de leitura do RNAm pela existência de inserções ou deleções, seja por mutação pontual que altera o consenso do *splice* (referências em Ayerdi-Izquierdo *et al.*, 2006). As mutações restantes são trocas de aminoácidos nos domínios da proteína (figura 1).

Jordi Pérez-Tur e seu grupo demonstraram que a *LGII* é N-glicosilada em três sítios (N192, N277 e N422) e que a isoforma curta é retida dentro da célula (Sirerol-Piquer *et al.*, 2006). Também mostraram que os domínios LRRs não são necessários para a secreção da

isoforma longa, mas que os sete domínios EPTP e pelo menos uma N-glicosilação da LGI1 são necessários para a proteína sair da célula.

Recentemente, Fukata *et al.* (2006) identificaram as subunidades de um complexo de proteínas localizado nas sinapses neurais de rato. Dentre os componentes, a LGI1 funcionaria como ligante, a ADAM22 como receptor e a PSD-95 como uma âncora. A LGI1 controlaria a intensidade das sinapses excitatórias. Curiosamente, além da LGI1, outras duas proteínas que se associam à PSD-95 estão relacionadas à epilepsia. A ADAM22 é uma proteína transmembrana neural, da família das metaloproteases, porém cataliticamente inativa (Novak, 2004), e quando mutada causa a morte por crise em camundongos (Sagane *et al.*, 2005). A estargazina, uma proteína ligadora à PSD-95, é uma proteína reguladora do receptor transmembrana AMPA (ácido α -amino-3-hidroxi-5-metil-4-isoxazol propiônico) na densidade pós-sináptica (Chen *et al.*, 2000; Tomita *et al.*, 2005; Nicoll *et al.*, 2006). A interação entre LGI1 e ADAM22 ocorre através da ligação das repetições EPTP da primeira com o domínio desintegrina da segunda (Fukata *et al.*, 2006). Este estudo mostra que, apesar dos canais iônicos serem o principal foco nas investigações sobre epilepsia, a interação ligante-receptor nos neurônios pode ser importante para a patogênese da doença.

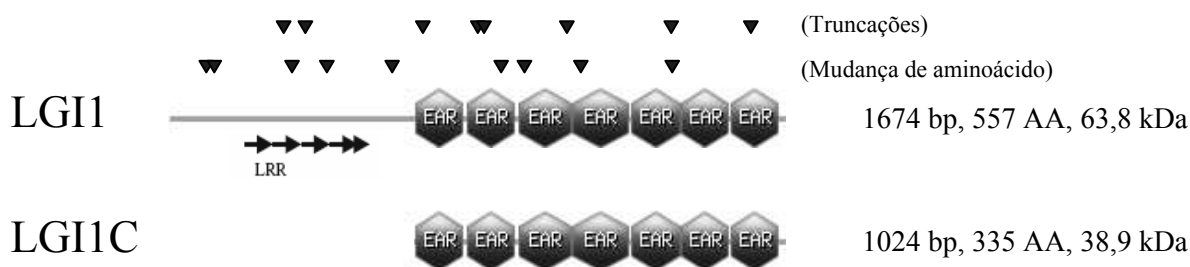


Figura 1: Esquema das proteínas LGI1 humana em sua forma completa (AA 1-557) e a porção C-terminal (LGI1C; AA 224-557) estudados nesta tese. Na parte superior estão indicadas as mutações, descritas na literatura, que causam truncações e mudanças de aminoácidos (referências em Gu *et al.*, 2005).

Foi relatado, ainda, que a LGI1 se associa à Kv1.1, uma subunidade do canal de potássio pré-sináptico voltagem-dependente (Schulte *et al.*, 2006). A LGI1 inibe a inativação do canal através de uma proteína citoplasmática reguladora, a Kvβ1.3 (Schulte *et al.*, 2006). Intrigantemente, também foi identificada a proteína ADAM22 nos fracionamentos celulares realizados por este grupo. Entretanto, como a LGI1 é uma proteína secretada, mais estudos são necessários para tornar seus ligantes/receptores conhecidos ou, ainda, investigar se sua forma truncada ou completa tem de fato função de modulador do canal de potássio citosólico.

Todavia, pouco se sabe sobre a função individual de cada domínio da LGI1 na transmissão sináptica ou na regulação do influxo de potássio da célula. Em seu possível papel regulador da proteína Kv1.1, não foi identificado o domínio de interação da LGI1 com a subunidade Kvβ1.3 (Schulte *et al.*, 2006). Enquanto as repetições EPTP na porção C-terminal são responsáveis pela interação com a proteína ADAM22 (Fukata *et al.*, 2006), não se sabe a função do domínio LRR na transmissão sináptica. Outro ponto não esclarecido é a função da isoforma de menor tamanho que não é secretada da célula (Senechal *et al.*, 2005; Sirerol-Piquer *et al.*, 2006); e de maneira global, se as mutações que afetam a estrutura primária da LGI1 estão relacionadas à EPADSA.

1.2 A PROTEÍNA EFHC1

Foi identificado em algumas famílias com epilepsia mioclônica juvenil (EMJ) mutação no gene *EFHC1* (*EF-hand domain C-terminal containing 1*) (Suzuki *et al.*, 2004). Segundo classificação da *International League Against Epilepsy* (ILAE, 1989), este tipo de epilepsia é classificada como idiopática generalizada, apresentando prevalência de 7-9% entre adultos e adolescentes epiléticos, de ambos os sexos (Genton & Gelisse, 2001). Seu início se dá no

período pré- ou pós-puberdade, sendo que 85% dos casos ocorre entre os 13 e 20 anos de idade (Janz, 1989; Wolf, 1992).

Esta doença é caracterizada por abalos mioclônicos, principalmente ao despertar, acompanhada por crises tônico-clônicas e/ou clônico-tônico-clônicas generalizadas (90-95% dos pacientes) e crises de ausência (até 39,5% dos casos) (Delgado-Escueta & Enrile-Bacsal, 1984). Adicionalmente, os pacientes são freqüentemente fotossensíveis. Assim como em outras síndromes idiopáticas generalizadas, observa-se o envolvimento de ambos os hemisférios cerebrais desde o princípio da crise, o que define a EMJ (ILAE, 1989).

Esta síndrome apresenta predisposição genética marcante, perfazendo 39-59% de todas as epilepsias (ILAE, 1989; Hauser & Hesdorffer, 1990), o que a torna a forma mais comum de epilepsia idiopática generalizada (Gardiner, 2005). Por esta razão, a EMJ tem sido uma das formas de epilepsia mais amplamente estudadas do ponto de vista molecular. A região cromossomal 6p21 foi identificada como o primeiro *locus* associado a esta doença (Greenberg *et al.*, 1988), através de estudos de ligação. Foram encontradas ainda evidências de que o gene *BRD2*, localizado no cromossomo 6p21, poderia estar relacionado à EMJ (Pal *et al.*, 2003). Além disso, diversos SNPs (*single nucleotide polymorphisms*) foram associados a esse gene, mas não se conhece a relevância desses polimorfismos.

Foi identificada uma mutação no gene *GABRA1* em indivíduos afetados segregando uma forma autossômica dominante de EMJ (Cossette *et al.*, 2002). Este gene está localizado na região 5q34, e codifica a subunidade $\alpha 1$ do receptor GABA_A, principal transmissor inibitório do sistema nervoso central. Esses autores demonstraram que os receptores GABA_A que continham a subunidade mutada apresentavam diminuição da amplitude de correntes ativadas por GABA, indicando que as crises poderiam estar relacionadas à perda de função desses receptores.

O gene *EFHC1* foi identificado na região 6p12.3 pelo grupo de Yamakawa (Suzuki *et al.*, 2004). Os autores investigaram a função *in vitro* deste gene e de seus mutantes e a associação dos mesmos com a EMJ. Neurônios de hipocampo de camundongo mantidos em cultura foram transfectados com um cassete de expressão do gene *EFHC1*. Observou-se, após 48h de transfecção, que as células mostravam sinais de neurodegeneração e morte celular. Análises de interação da proteína EFHC1 com canais de cálcio voltagem-dependente, cujas anormalidades foram descritas em epilepsias humanas e murinas (Fletcher *et al.*, 1996; Jouvenceau *et al.*, 2001), revelaram que a taxa de sobrevivência dos neurônios transfectados com EFHC1 foi aumentada de maneira específica quando estas células foram tratadas com um antagonista de canal de cálcio voltagem-dependente do tipo R (Ca_v2.3). Este grupo concluiu que a proteína EFHC1 se associa a este canal e aumenta o influxo de cálcio pelo mesmo, acarretando, dessa forma, morte celular programada.

Neste mesmo estudo, foram identificadas cinco mutações no gene *EFHC1* que ocasionavam troca de aminoácidos na proteína. Essas mutações co-segregavam com o quadro epiléptico em membros de seis famílias não relacionadas acometidos por EMJ, estando ausentes em 382 indivíduos não-afetados por esta doença. Apesar das proteínas EFHC1 mutantes se associarem ao canal Ca_v2.3, todas apresentavam queda significativa na indução de apoptose. Segundo os autores, esses resultados sugerem que as mutações no gene *EFHC1* causam EMJ porque comprometem a atividade apoptótica da proteína selvagem, evitando a eliminação, durante o desenvolvimento do sistema nervoso central, de neurônios em que a homeostase do cálcio está alterada e promovendo aumento na densidade neuronal.

Suzuki *et al.* (2004) também isolaram o gene *Efhc1* de camundongo. Os experimentos de *Northern blotting* revelaram que o transcrito de 2,3 kb desse gene é mais expresso em cérebros de embriões que em cérebros de adultos. As análises imunohistoquímicas mostraram que a

proteína Efhc1 é abundante no cérebro destes animais, sendo encontrada no hipocampo, cerebelo, córtex cerebral, tálamo, hipotálamo e amígdala. Est proteína está presente em dendritos e corpos celulares de neurônios, mas está ausente em axônios.

O grupo de Kanya observou que Efhc1 de camundongo apresenta 40% de identidade com a proteína Rib72 (p72) de *Chlamydomonas reinhardtii* (Ikeda *et al.*, 2005). A proteína mRib72-1/Efhc1, também chamada mioclonina 1, foi identificada como sendo um componente de axonemas de flagelos (Patel-King *et al.*, 2002; Ikeda *et al.*, 2003). Acredita-se que Rib72 desempenhe papel importante na montagem e função dos axonemas em cílios e flagelos móveis (Ikeda *et al.*, 2005). Verificou-se também que maiores níveis de expressão de *mRib72-1/Efhc1* são observados nos cílios e flagelos de células de testículos e ovidutos de camundongos adultos. Assim, esses autores propõem que o funcionamento inadequado dos cílios, resultante de mutações no gene *mRib72-1/Efhc1*, seja a causa da EMJ, o que por si só é de fato intrigante (King, 2006). Entretanto, pode haver relação direta entre a deficiência na motilidade de células endodimais, altamente ciliadas, e o desenvolvimento neuronal (Ibanez-Tallon *et al.*, 2004). Evidentemente, a identificação da mRib72-1/Efhc1 em cílios e flagelos fornece nova luz na provável causa da mais comum epilepsia idiopática. Conhecer se Efhc1 está presente ou não em cílios de células endodimais e qual a relação entre defeitos nessas organelas e EMJ será de grande interesse.

Sabe-se, através de análises *in silico* e alinhamento em bancos de dados, que a proteína EFHC1 humana, de aproximadamente 74 kDa, apresenta três domínios DM10 em sequência (Ikeda *et al.*, 2005), de função desconhecida, e um motivo EF-hand de ligação a cálcio na porção C-terminal (Grabarek, 2006). A figura 2 mostra a organização dos domínios desta proteína.

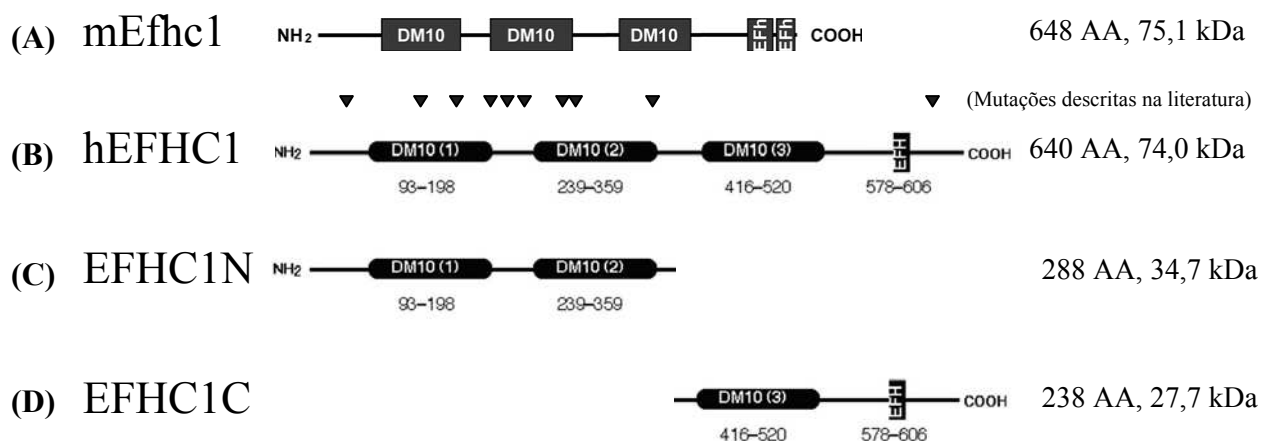


Figura 2: Organização estrutural da proteína mRib72/Efhc1 murina e EFHC1 humana. (A) Ilustração retirada de Ikeda *et al.* (2005), que assumem a existência de dois motivos EF-hand na proteína murina. (B) Esquema de Suzuki *et al.* (2004), que prevêem a existência de apenas um motivo, utilizando os bancos de dados InterPro, Pfam e SMART. Na parte superior do esquema apresentado em (B) estão indicadas as mutações que causam mudanças de aminoácidos. (C) Clone N-terminal, EFHC1N, e (D) clone C-terminal, EFHC1C, da proteína EFHC1.

O domínio DM10 apresenta aproximadamente 105 resíduos e foi identificado em uma única cópia na porção N-terminal da proteína humana nm23-H7, uma nucleosídeo difosfato quinase (NDK, *Nucleoside-Diphosphate Kinase*) (Munier *et al.*, 2003; Patel-King *et al.*, 2004) e na proteína associada a axonema flagelar p40 de *Chlamydomonas reinhardtii* (Patel-King *et al.*, 2002). Na proteína Efhc1 murina e na Rib72 de *Chlamydomonas* há a presença de três repetições do domínio DM10 (Patel-King *et al.*, 2002; Ikeda *et al.*, 2005). Proteínas que apresentam esse domínio interagem com microtúbulos (Ikeda *et al.*, 2003; Ikeda *et al.*, 2005), funcionando possivelmente na montagem do axonema ou na regulação de NDKs (King, 2006), e na regulação da motilidade flagelar dependente de cálcio (Patel-King *et al.*, 2002). O EF-hand é o motivo de ligação a cálcio mais amplamente utilizado por proteínas (Lewit-Bentley & Réty, 2000) e é

caracterizado por uma estrutura hélice-alça-hélice (Bhattacharya *et al.*, 2004; Grabarek, 2006). A superfamília EF-hand regula diversos aspectos do funcionamento celular, como tamponamento de cálcio no citosol, proliferação celular e transdução de sinal (Skelton *et al.*, 1994; Berridge *et al.*, 2000). No sistema nervoso central, a calmodulina e a proteína sensora de cálcio têm importante papel na sinalização neuronal (Levitan, 1999; Burgoyne *et al.*, 2004). A unidade funcional básica do EF-hand consiste em um par de motivos (Bhattacharya *et al.*, 2004), mas no caso da proteína humana, não está claro se está presente em par ou não (figura 2).

EFHC1 se liga ao aparato mitótico através de sua porção N-terminal e pode ter papel significativo na divisão celular (de Nijs *et al.*, 2006); todavia, sua função durante a mitose ainda precisa ser esclarecida (Grisar *et al.*, 2006). De qualquer maneira, sua associação com o fuso mitótico através de microtúbulos é um achado que se deve levar em consideração na fisiopatologia da EMJ. O grupo de Chanasa observou ainda, utilizando diversas construções truncadas, que os 92 resíduos iniciais da proteína são necessários para o endereçamento da EFHC1 ao fuso mitótico e ao corpo mediano (de Nijs *et al.*, 2006) e que não há envolvimento direto dos primeiros dois domínios DM10 (figura 2). Como o nível de expressão do RNAm é maior no período embrionário, levanta-se a hipótese de participação da EFHC1 no controle da migração neuronal nos estágios iniciais do desenvolvimento embrionário (Suzuki *et al.*, 2004; Grisar *et al.*, 2006).

Além da interação à porção citoplasmática do canal de cálcio Ca_v2.3, o grupo de Suzuki *et al.* (2004) também relatou a interação funcional e molecular da EFHC1 com o receptor transiente potencial do canal M2 (TRPM2), um canal de cátion permeável a cálcio (Mori *et al.*, 2005). Entretanto, a função fisiológica deste canal continua apenas especulativa (Mei *et al.*, 2006).

2. OBJETIVOS

Este trabalho teve por meta estudar os aspectos estruturais das proteínas humanas LGI1 e sua porção C-terminal e EFHC1 dividida em porção N- e C-terminal. De maneira específica, os seguintes objetivos foram estabelecidos:

1) Clonar e expressar em diferentes caudas de fusões, testando-se diferentes cepas de *Escherichia coli*, a fim de se obter proteína solúvel;

2) Clivar a cauda de fusão;

3) Acessar a estrutura secundária e terciária das proteínas por diferentes métodos, como dicroísmo circular, espalhamento de luz dinâmico, espalhamento de raios X a baixo ângulo, espectrometria de massas, difração de raios X, entre outros.

3. A PROTEÍNA LGI1

Os domínios LRRs e EPTs presentes na proteína LGI1 provavelmente são independentes estruturalmente, o que nos levou a buscar a clonagem da proteína tanto inteira, como sua porção N- e C-terminal, além de seus possíveis ligantes descritos na literatura. A seguir detalharemos a metodologia empregada e os resultados obtidos.

3.1 METODOLOGIA

Meios de cultura e cepas de Escherichia coli

As culturas de bactéria 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. Para os testes de purificação em resina de amilose, utilizou-se o meio LB acrescido de 0,2% de glicose.

As cepas de *Escherichia coli* DH5 α , BL21(DE3), BL21(DE3) pLysS, BL21(DE3) pTgroE, BL21(DE3) pRARE, Origami(DE3) (Novagen), Origami2(DE3) pRARE, C41(DE3), C41(DE3) pRARE, C41(DE3) pLysS e BL21Condon+RIL (Stratagene) foram transformadas por choque térmico de acordo com protocolo otimizado de Sambrook *et al.* (1989).

Extração de RNA total de hipocampo humano e reação de síntese de cDNA

Espécimes cirúrgicos e de necropsia de hipocampo humano foram congelados em nitrogênio líquido e macerados. A extração de RNA total foi realizada com Trizol (Invitrogen), seguindo especificações do fabricante.

Para a reação de síntese de cDNA total, utilizou-se o protocolo sugerido pelo fabricante (*Revertaid H Minus First Strand cDNA Synthesis Kit*, Fermentas). Resumidamente, utilizou-se 8 µg de RNA total, 1 µL de oligo(dT)₁₈ a 0,5 µg/µL e completou-se para 12 µL com H₂O livre de nuclease. Incubou-se por 5 min a 70°C e em seguida esfriou-se rapidamente em gelo. Adicionou-se o tampão de reação (concentração final Tris-HCl 50 mM, pH 8,3, KCl 50 mM, MgCl₂ 4 mM, DTT 10 mM), 1 µL de inibidor de ribonuclease (20 U/µL) e 2 µL de dNTPs (10 mM cada). Incubou-se a reação por 5 min a 37°C e em seguida adicionou-se 1 µL da enzima transcriptase reversa (200 U/µL), incubando-se por 1 h a 42°C. Após esse tempo, inativou-se a enzima aquecendo-se a 70°C durante 10 min e esfriando-se a reação em gelo.

O Comitê de Ética em Pesquisa (CEP) da Faculdade de Ciências Médicas (FCM) aprovou o adendo ao projeto nº 470/2003, que permite a análise e a extração de RNAs e proteínas dos espécimes cirúrgicos e de necrópsia.

Clonagem e construção dos vetores de expressão

As seqüências codificadoras das proteínas LGI1, LGI3, LGI4, Kvβ1.3 e ADAM22 foram amplificadas por PCR (Reação em Cadeia de Polimerase, da sigla em inglês) utilizando como molde cDNA total de hipocampo humano. Para as subclonagens subseqüentes das diferentes fusões de LGI1 utilizou-se como molde o clone pGEMT-easy LGI1 clone 1.7. A reação consistiu

em 1 µL de cDNA (até 10 ng), 0,5 µL da Taq DNA polimerase, 2 µL de dNTPs (1,25 mM cada), 1,5 µL do tampão de reação (10X) com MgCl₂, 1 µL de cada oligo (1 pmol/µL) e água deionizada para um volume final de 15 µL. A reação foi realizada em termociclador Eppendorf (*Mastercycler Gradient*) e constituiu-se basicamente de uma desnaturação inicial a 94°C por 2 min, seguida por 35-40 ciclos de desnaturação a 94°C por 30 s, anelamento por 30-45 s em gradiente e extensão a 72°C por 1 min, seguido de um ciclo de extensão final a 72°C por 10 minutos.

A sequência da proteína LGI2 foi amplificada por PCR tendo como molde o clone KIAA1916, gentilmente cedido por *Kazusa DNA Research Institute*.

Algumas características das proteínas selecionadas estão listadas na tabela 1. Os oligos utilizados em cada reação continham sítios de restrição específicos (tabela 2).

Os produtos de PCR obtidos foram purificados utilizando-se o sistema de purificação de DNA da Promega (*Wizard SV Gel and PCR Clean-Up System*) e analisados em gel de agarose. Em seguida, realizou-se a ligação do fragmento de DNA com o vetor de clonagem pGEM-T easy (Promega), de acordo com as instruções do fabricante. Confirmou-se a presença do inserto por análise com enzimas de restrição, ou PCR de colônia. Posteriormente transferiu-se o fragmento digerido em seus sítios de restrição específicos para os vetores de expressão pGEX-4T1 (GE Healthcare), pET32a (Novagen), pETNusA e pETMBP (gentilmente cedido pelo Dr. Ario de Marco, EMBL Heidelberg), previamente clivados e defosforilados. No caso da subclonagem com o vetor pETSUMO (Invitrogen), realizou-se a ligação direta ao vetor através da estratégia de clonagem do tipo TA.

Sigla	Nome	Inserto (pb)	AA	MW (kDa)
LGI1	<i>Leucine-rich Glioma Inactivated protein 1</i> (AA 1-557)	1674	557	63,8
LGI1N	Porção N-terminal da proteína LGI1 (AA 1-222)	669	222	24,8
LGI1C	Porção C-terminal da proteína LGI1 (AA 224-557)	1005	335	38,9
LGI1 isoforma	LGI1 Isoforma <i>splice</i> alternativo (AA 1-291) (12 aminoácidos finais diferentes)	892	291	33,1
LGI1 tr5	LGI1 construção truncada nº 5 (AA 316-557)	748	243	28,4
LGI1 tr6	LGI1 construção truncada nº 6 (AA 418-557)	442	141	16,3
LGI2	<i>Leucine-rich Glioma Inactivated protein 2</i> (AA 1-545)	1638	545	62,0
LGI3	<i>Leucine-rich Glioma Inactivated protein 3</i> (AA 1-548)	1647	548	61,7
LGI4	<i>Leucine-rich Glioma Inactivated protein 4</i> (AA 1-537)	1614	537	59,0
Kvβ1.3	<i>Potassium voltage-gated channel beta member 1</i> isoforma 3 (AA 1-401)	1206	401	44,7
ADAM22 Δ768-859	<i>ADAM metalloproteinase domain 22</i> Δ768-859 (domínio extracelular) (AA 1-767)	2304	767	85,1
ADAM22 (AA 233-718)	<i>ADAM metalloproteinase domain 22</i> (domínio extracelular) (AA 233-718)	1473	486	53,8

Tabela 1: Dados básicos das proteínas selecionadas. AA: número de aminoácidos. MW: massa molar predita pelo servidor Protparam.

ORF	Oligos	Direção	Seqüência	Sítio de restrição	Vetor ¹
LGI1	OMM10 OMM11	D R	5'-Cg AATT cATggAATCAgAAAgAAgC-3' 5'-TTgCTC gAg TCATgCgCTTAAGTCAAC-3'	EcoRI XhoI	pET28a pGEX-4T1
LGI1	OMM26 OMM29	D R	5'-ATgAAACCAgCgAAgCCAAAATg-3' 5'-TCATgCgCTTAAGTCAACTATg-3'	— —	pETSUMO
LGI1	OMM30 OMM32	D R	5'-CTg AggTCTCg CATgAAACCAgCgAAgCCAAAATg-3' 5'-gTT AAgCTTT CATgCgCTTAAGTCAACTAT-3'	Eco31I (NcoI) HindIII	pETNusA pETMBP
LGI1N	OMM17 OMM12	D R	5'-TC ggATCC ATggAATCAgAAAgAAgC-3' 5'-AAACTCgAgTCATgCAATCgAAATCCTTC-3'	BamHI XhoI	pGEX-4T1 pET32a
LGI1C	OMM14 OMM11	D R	5'-TGA ATTTCATG ACAGAAATTTGCAAAGTCTC-3' 5'-TTgCTC gAg TCATgCgCTTAAGTCAAC-3'	EcoRI XhoI	pET28a pGEX-4T1
LGI1C	OMM28 OMM29	D R	5'-ATgACAgAATTTgCAAAGTCTC-3' 5'-TCATgCgCTTAAGTCAACTATg-3'	— —	pETSUMO
LGI1C	OMM31 OMM32	D R	5'-TTC gggTCTCTC ATgACAgAATTTgCAAAGTCTC-3' 5'-gTT AAgCTTT CATgCgCTTAAGTCAACTAT-3'	Eco31I (NcoI) HindIII	pETNusA pETMBP
LGI1 isoforma	OMM10 OMM13	D R	5'-Cg AATT cATggAATCAgAAAgAAgC-3' 5'-TTACTC gAg TCATgACATgTTTgTAAATC-3'	EcoRI XhoI	pGEX-4T1
LGI1 tr5	OMM15 OMM11	D R	5'- CGAATTC ATGAATAAATTCATAAAATC-3' 5'-TTgCTC gAg TCATgCgCTTAAGTCAAC-3'	EcoRI XhoI	pGEX-4T1
LGI1 tr6	OMM16 OMM11	D R	5'- CGAATTCATG CAATTATTCATAACCAAAC-3' 5'-TTgCTC gAg TCATgCgCTTAAGTCAAC-3'	EcoRI XhoI	pGEX-4T1
LGI2	OMM4 OMM5	D R	5'-CATATgCgCTgCggAgAggC-3' 5'-CC ggATCC TCACAAACTTAAGTCAAC-3'	NdeI BamHI	pET28a
LGI3	OMM6 OMM7	D R	5'-CCATATGgCggggCTgCgggC-3' 5'-ggg AATTC CTAggCACTgAgATCCAC-3'	NdeI EcoRI	pET28a
LGI4	OMM8 OMM9	D R	5'-ggCATATgAggggCAgCATTC-3' 5'-TgAAgCTTTCAGgCACTgAggTCgATC-3'	NdeI HindIII	pET28a
Kvβ1.3	OMM36 OMM37	D R	5'-ATgCAAgtCTCCATAgCCTg-3' 5'-TTATgATCTATAgCTCTTCTTg-3'	— —	pETSUMO
ADAM22 Δ768-859	OMM38 OMM39	D R	5'-CATATgCAgCgCgAgTggCTg-3' 5'-g AATTC TCATCTCTgTTCTCgATAgTT-3'	NdeI EcoRI	pETSUMO pGADT7 pGBKT7
ADAM22 (AA 233-718)	OMM40 OMM41	D R	5'-CATATgAATgTAgAAgAAgAAACC-3' 5'-g AATTC CAATTgTgAgggAAgTAAgTg-3'	NdeI EcoRI	pETSUMO pGADT7 pGBKT7

Tabela 2: Oligos utilizados para as reações de PCR. Para cada seqüência codificadora (ORF, *Open Reading Frame*) específica está indicado o par de oligos utilizado, bem como os sítios de restrição, quando presentes, e o vetor de destino. Legenda: D: direto; R: reverso.

¹ Salientamos que outros vetores foram previstos, como pET29a, pET41a, pSV282 (gentilmente cedido pela Profa. Mizoue, Universidade de Vanderbilt), além de outros clones truncados (figura 15).

A integridade do inserto foi verificada por seqüenciamento automático (MegaBACE 1000 da GE Healthcare, ou ABI PRISM 377 *Genetic Analyser* da Applied Biosystems). Para o seqüenciador ABI 377, utilizou-se o sistema *Big Dye*, seguindo protocolo fornecido pelo fabricante. Para o seqüenciador automático MegaBACE, realizou-se a seguinte reação: 4 µL de *DYEnamic ET reagent premix* (*DYEnamic ET dye terminator cycle sequencing kit – MegaBACE*; GE Healthcare), 1 µL de DNA, 1 µL de cada oligo (5 pmol/µL) para um volume final de 15 µL. Utilizou-se o programa indicado pelo fabricante: 35 ciclos de 94°C por 20 s, 50°C por 15 s, 60°C por 1 min.

Análises in silico

As seqüências nucleotídicas das construções foram analisadas no servidor Webcutter 2.0 (ver endereços eletrônicos), para a construção do mapa de restrição do inserto de interesse, a fim de se verificar se não continham sítios de restrição internos para as enzimas selecionadas. Para a análise do “*codon usage*” da proteína heteróloga LGI1, Kvβ1.3 e ADAM22 foi utilizado o servidor *Rare Codon Caltor*.

A predição de massa molar, composição de aminoácidos, ponto isoeletrico, entre outros parâmetros, foi realizada no servidor Protparam (Gasteiger *et al.*, 2005).

A predição de estrutura secundária das proteínas foi feita por meio do servidor PSIPRED (Jones, 1999; McGuffin *et al.*, 2000), PHD (Rost *et al.*, 2004) e DICHROWEB (Lobley *et al.*, 2002; Whitmore & Wallace, 2004). Utilizou-se esta predição para a construção dos clones truncados N- (1-222) e C-terminal (224-557) da LGI1 (anexo 2). Pelo índice de instabilidade fornecido pelo Protparam, a porção N-terminal da LGI1 seria instável, enquanto a porção C-

terminal seria estável. A predição de estrutura secundária da porção N-terminal do gene *LGII* foi realizada pelo servidor Mfold (Zuker, 2003). Adicionalmente, a fim de se delimitar com maior precisão as bordas dos clones truncados, os servidores SMART, Pfam e PROSITE foram consultados.

Para a predição de solubilidade utilizou-se o servidor *Recombinant Protein Solubility Prediction* (Wilkinson & Harrison, 1991; Davis et al., 1999; Harrison, 2000).

Extração de DNA plasmidial

Os DNAs plasmidiais foram extraídos de acordo com o protocolo de lise alcalina descrito por Birnboim & Doly (1979). Uma vez confirmada a presença do inserto no vetor, o DNA foi isolado utilizando-se kits comerciais de purificação (*Wizard Plus SV Minipreps*, Promega), conforme especificações do fabricante.

Reações de clivagem e eletroforese de DNA

Para a digestão de vetores e insertos, foram utilizadas enzimas de restrição específicas. Em cada reação, utilizou-se 1 unidade de enzima por µg de DNA, incubando-se em média de 2 a 4 horas a 37°C. O DNA dos vetores foi precipitado com 10% do volume da reação com acetato de amônio 3M, pH 4,8 e 3 volumes de etanol 100% gelado, sendo posteriormente ressuspensionado em 10 µL de H₂O milli-Q estéril. No caso das subclonagens, os vetores, após clivagem, foram defosforilados e extraídos do gel com o kit *Wizard SV Gel and PCR Clean-Up System*

(Promega). As construções foram confirmadas por meio de clivagens enzimáticas e posteriormente por seqüenciamento automático.

Todas as análises de DNA foram feitas por eletroforese em gel de agarose a 0,8% (Sambrook *et al.*, 1989), utilizando tampão de corrida TBE 1X concentrado (Tris-borato 40 mM, pH 8,0, EDTA 1 mM). Os géis foram corados com brometo de etídio 0,1 µg/mL e a purificação dos fragmentos de DNA excisados do gel foi feita com o sistema de purificação da Promega, de acordo com o protocolo do fabricante.

Ligação de fragmentos de DNA

Para a clonagem em pGEM-T easy, utilizou-se no mínimo 20 ng de DNA do inserto, 50 ng de DNA do vetor, 5 µL de tampão de ligação 2X concentrado (Tris-HCl 60 mM, pH 7,8, MgCl₂ 20 mM, DTT 20 mM, ATP 2 mM, polietilenoglicol 8000 10%), 3 unidades de T4 DNA Ligase (Promega) e H₂O milli-Q estéril suficiente para 10 µL de reação. Utilizou-se o controle positivo fornecido pelo fabricante. A reação foi incubada a 4°C por 16h. Após esse período, adicionou-se a ligação em uma suspensão de células *E. coli* DH5α competentes e seguiu-se com o procedimento de transformação.

Para as subclonagens em pGEX-4T1, pET32a, pETSUMO, pETNusA e pETMBP, utilizou-se 100-200 ng de DNA do vetor, 200-300 ng de DNA do inserto, 2 µL de tampão de ligação 10X concentrado (Tris-HCl 300 mM, pH 7,8, MgCl₂ 100 mM, DTT 100 mM, ATP 10 mM), 1 unidade de T4 DNA Ligase (Promega ou USB) e água para um volume total de 20 µL. Como controle, realizou-se a reação como descrito acima, entretanto, sem a presença do inserto.

A reação foi incubada por 16h a 16°C. Posteriormente, células DH5 α competentes foram transformadas com a ligação.

Teste de expressão e de solubilidade em E. coli e preparação dos extratos

Os testes iniciais de expressão foram realizados em pequena escala. Para as construções em pGEX-4T1 utilizou-se a cepa BL21Codon+RIL e para as construções em pET32a utilizou-se a cepa Origami(DE3). Em ambos os casos foi co-transformado o vetor pRARE (Novagen), que transcreve códons raros em *E. coli*. Como controle positivo, utilizou-se o vetor pS219V, que codifica a protease TEV, em conjunto com a cepa BL21(DE3) pLysS (tabelas 3 e 4).

Para as construções em pETSUMO, pETNusA e pETMBP, foram testadas diversas cepas e temperaturas, além da co-expressão dos vetores pRARE (Novagen), pTgroE (Yasukawa *et al.*, 1995), que co-expressa o sistema de chaperoninas GroEL-GroES, ou ainda, conforme o caso, pLysS ou pLysE, que reprime a transcrição basal do sistema de expressão pET. Como controle positivo, utilizou-se o vetor pETNusA_GFP e pETMBP_GPF, que codifica para a proteína fluorescente verde fusionada a NusA ou MBP, respectivamente.

Os testes de indução foram realizados em 100 mL ou 250 mL de meio LB, ou LB glicose, contendo o antibiótico apropriado. Cresceu-se o pré-inóculo por 16h a 37°C sob agitação constante de 200 rpm. No dia seguinte utilizava-se de 1-3% de pré-inóculo no meio de cultura até atingir a absorbância a 600 nm (A_{600nm}) de 0,8. Em seguida, alterava-se para a temperatura de indução a ser testada (18-20°C, 25°C, 30°C, 37°C ou 42°C) e adicionava-se o indutor IPTG (isopropanil- β -D-tiogalactopiranosídeo) para concentração final 0,4 mM. Em 18-20°C aumentava-se o tempo de indução para 6h ou 16h, enquanto que nas temperaturas mais altas

(acima de 25°C), o tempo de indução foi fixado em 4h. Testou-se ainda uma concentração de indutor maior no caso das induções a 18-20°C (0,6 mM de IPTG).

Construção	Cepa de <i>E. coli</i>	Temperatura de indução testada
pGEX-4T1 LGI1	BL21Codon+RIL, BL21Codon+RIL pRARE	18-20°C, 25°C, 37°C e 42°C
pGEX-4T1 LGI1C	BL21Codon+RIL, BL21Codon+RIL pRARE	18-20°C, 25°C, 37°C e 42°C
pET32a LGI1	Origami(DE3) pRARE	18-20°C, 25°C, 30°C
pET32a LGI1C	Origami(DE3) pRARE	18-20°C, 25°C, 30°C
pS219V (TEV)	BL21(DE3) pLysS	15°C

Tabela 3: Construções testadas, cepas bacterianas e temperaturas de indução utilizadas.

Posteriormente, as células foram coletadas por centrifugação a 6000 g, por 5 min a 4°C e congeladas a -20°C até o momento da preparação dos extratos.

Para a análise dos testes de expressão, as células foram descongeladas, ressuspensas em tampão de extração (Tris-HCl 20 mM, pH 7,5, NaCl 150 mM, glicerol 5%, EDTA 1 mM, PMSF 0,5 mM), ou em tampão A (Na₂HPO₄ e NaH₂PO₄ 50 mM, pH 7,2, glicerol 5%, NaCl 100 mM, EDTA 0,1 mM) para a expressão com o vetor pET32a e em tampão PBS 1X (NaCl 140 mM, KCl 2,7 mM, Na₂HPO₄ 10 mM, KH₂PO₄ 1,8 mM, pH 7,3) para a expressão com o vetor pGEX. Acrescentou-se lisozima (5 µg/mL), incubou-se em gelo e sonicou-se durante 10 s por 3 a 6 vezes em um sonicador Sonics VCX-750. Em seguida realizou-se a clarificação do extrato por centrifugação a 20.000 g, 10 min 4°C. As frações insolúveis foram sonicadas novamente no mesmo tampão acrescido de 2% de Triton X-100, clarificadas por centrifugação e ressuspensas em tampão de extração. As alíquotas das frações solúveis e insolúveis foram analisadas por eletroforese em gel de poliacrilamida (SDS-PAGE), corado com azul de coomassie, conforme método descrito previamente (Laemmli, 1970).

Construção	Cepa de <i>E. coli</i> utilizada no teste de expressão	Temperatura de indução testada	Teste de solubilidade ¹	Testes de purificação ²
pETSUMO LGI1	BL21(DE3) pLysS, BL21(DE3) pRARE, C41(DE3), C41(DE3) pRARE, Origami2(DE3) pRARE	18° e 30°C	BL21(DE3) pRARE 25° e 30°C	Origami2(DE3) pRARE 18° e 30°C Ni-NTA
pETSUMO LGI1C	BL21(DE3) pRARE, BL21(DE3) pLysS, C41(DE3), C41(DE3) pRARE	18° e 30°C	BL21(DE3) pRARE 25° e 30°C	Origami2(DE3) pRARE 18° e 30°C Ni-NTA
pETNusA LGI1C	BL21(DE3) pRARE, BL21(DE3) pLysS, C41(DE3) pRARE	18° e 30°C	BL21(DE3) pRARE 25° e 30°C	BL21(DE3) pRARE 30°C Ni-NTA nativo (3 tampões) Ni-NTA desnaturante
pETMBP LGI1C	BL21(DE3) pRARE, BL21(DE3) pTgroE	18°, 25° e 30°C	BL21(DE3) pRARE 30°C BL21(DE3) pTgroE 30°C	BL21(DE3) pRARE ou pTgroE 25° e 30°C Ni-NTA nativo (4 tampões) Ni-NTA desnaturante <i>Refolding</i> em coluna Amilose (4 tampões)
pETSUMO Kvβ1.3	BL21(DE3) pRARE	18° e 30°C	–	BL21(DE3) pRARE 30°C Ni-NTA e TALON
pETSUMO ADAM22	BL21(DE3) pRARE	30°C	–	Origami2(DE3) pRARE 30°C
pETNusA GFP	BL21(DE3)	37°C	–	BL21(DE3) 37°C Ni-NTA
pETMBP GFP	BL21(DE3)	37°C	–	BL21(DE3) 37°C Ni-NTA

Tabela 4: Construções testadas, cepas bacterianas e temperaturas de indução utilizadas.

¹ Realizado em diferentes tampões, alterando pH, força iônica e aditivos (como aminoácidos e detergentes).

² Apenas para cromatografia de afinidade. Indicamos a cepa bacteriana utilizada no teste de purificação, além da temperatura utilizada.

Para os testes de solubilidade, 12 alíquotas de 2 mL do meio de cultura cada foram ressuspendidas em tampões específicos, variando pH, força iônica e aditivos (como aminoácidos e detergentes, tabela 5), de forma a se aproximar da composição do tampão em que as proteínas de interesse poderiam ser extraídas em máxima quantidade na fração solúvel. As frações solúveis e insolúveis foram tratadas como acima descrito.

Nº do tampão	Bis-Tris pH 6,8	Tris-HCl pH 8,0	NaCl	Arginina	Glicina	NP-40	Triton X-100
1	20 mM	–	150 mM	–	–	–	–
2	20 mM	–	150 mM	100 mM	–	–	–
3	20 mM	–	150 mM	–	100 mM	–	–
4	20 mM	–	150 mM	–	–	0,1%	–
5	20 mM	–	500 mM	–	–	–	0,1%
6	20 mM	–	500 mM	–	–	–	–
7	–	20 mM	150 mM	–	–	–	–
8	–	20 mM	150 mM	100 mM	–	–	–
9	–	20 mM	150 mM	–	100 mM	–	–
10	–	20 mM	150 mM	–	–	–	0,1%
11	–	20 mM	500 mM	–	–	0,1%	–
12	–	20 mM	500 mM	–	–	–	–

Tabela 5: Tampões para teste de solubilidade.

Testes de purificação protéica

O vetor de expressão pGEX-4T1 permite a expressão de proteínas recombinantes fusionadas à GST, glutathione-*S*-transferase, de *Schistosoma japonicum* (Nygren *et al.*, 1994). Sua purificação é possível através de cromatografia de afinidade em resina Glutathione Sepharose 4B (GE Healthcare), em que a proteína é eluída em condições pouco estridentes, com glutathione reduzida (Smith & Johnson, 1988). Os extratos protéicos foram incubados com 100 µL da resina por 1 h a 4°C com agitação constante. Após esse período, a matriz foi transferida para coluna BioSpin (BioRad) e lavou-se a resina com 1 mL de PBS 1X por 3 vezes, correspondendo a 30 volumes de coluna (cv). A eluição se dava com 100 µL de glutathione reduzida 10 mM em Tris-HCl 50 mM, pH 8,0, por três vezes. As frações foram analisadas em gel SDS-PAGE.

Em relação ao vetor pET32a, sua expressão permite a fusão com Trx, tioredoxina, de *E. coli* (LaVallie *et al.*, 1993), seguido de uma cauda de poli-histidina (His₆) na porção N-terminal

da proteína de interesse, a mais utilizada cauda para sistema de expressão heterólogo (Gaberc-Porekar & Menart, 2001). Esta cauda His₆ também está presente nas fusões expressas pelos vetores pETSUMO, pETNusA e pETMBP, que codificam as caudas SUMO, pequeno modificador semelhante à ubiquitina de *Homo sapiens* (Marblestone *et al.*, 2006); NusA, substância A de utilização N de *E. coli* (Davis *et al.*, 1999) e MBP, proteína ligadora de maltose de *E. coli* (di Guan *et al.*, 1988; Fox & Waugh, 2003; Nallamsetty & Waugh, 2006), respectivamente.

Por meio da cauda His₆, as proteínas recombinantes podem ser isoladas dos extratos celulares por cromatografia de afinidade em metal imobilizado (IMAC, do inglês *Immobilized Metal Affinity Chromatography*). O extrato protéico foi incubado com 100 µL da resina Ni-NTA Superflow (Qiagen) por 1h, em temperatura ambiente (25°-30°C), em um agitador rotatório. Após esse período, a matriz foi transferida para uma coluna BioSpin (BioRad) e as proteínas eluídas em um gradiente crescente de imidazol (de 5 a 200 mM). No caso da captura em condição desnaturante, tanto para NusA-LGI1C quanto para MBP-LGI1C, utilizou-se uréia 8M no tampão de extração. Para o teste de renaturação em coluna, apenas para MBP-LGI1C, seguiu-se o protocolo de Oganessian *et al.* (2004; 2005).

Para os testes de purificação da construção NusA-LGI1C por troca iônica, utilizou-se as colunas HiTrap DEAE FF e HiTrap Q FF, ambas de 1 mL (GE Healthcare). A proteína foi extraída no tampão A (Tris-HCl 20 mM, pH 8,0, NaCl 20 mM, DTT 0,5 mM, PMSF 0,5 mM) e a eluição se deu em um gradiente de sal (NaCl até 1M). Para os testes de captura por interação hidrofóbica desta mesma construção, utilizou-se as colunas HiTrap Phenyl FF (high sub) e HiTrap Phenyl FF (low sub) e HiTrap Butyl FF (GE Healthcare), todas de 1 mL. Utilizou-se o tampão A acrescido de 1 M de sulfato de amônio e a eluição se deu em gradiente decrescente de

sal. Todas as colunas da GE Healthcare foram utilizadas no sistema FPLC (Amersham Biosciences).

Para os testes de captura da construção MBP-LGI1C através da resina de amilose (New England Biolabs), foram testados quatro diferentes tampões com a co-expressão das chaperoninas GroEL-GroES: tampão 6 (Bis-Tris 20 mM, pH 6,8, NaCl 500 mM); tampão 7 (Tris-HCl 20 mM, pH 8,0, NaCl 150 mM), tampão 8 (Tris-HCl 20 mM, pH 8,0, NaCl 150 mM, arginina 100 mM) e tampão 10 (Tris-HCl 20 mM, pH 8,0, NaCl 150 mM, Triton X-100 0,1%), todos acrescidos de DTT 1 mM (tabela 5), pois a proteína LGI1C apresenta três cisteínas (anexo 2). Em relação à expressão desta construção apenas com o vetor pRARE, testou-se a captura com os tampões 6 e 7. Passou-se o extrato celular pela resina de amilose (1,0-1,5 mL para cada 100 mL de extrato), em coluna de gravidade BioSpin (BioRad), sendo que as proteínas foram eluídas com maltose 10 mM em 15 volumes de coluna (*cv*).

No caso da cromatografia por filtração em gel, utilizou-se a coluna Superdex 200 10/300 GL de 24 mL (GE Healthcare), no sistema FPLC, para a amostra MBP-LGI1C/GroEL. A coluna foi equilibrada em tampão TEV modificado (Tris-HCl 50 mM, pH 8,0, NaCl 150 mM, DTT 1 mM, EDTA 0,5 mM). O fluxo utilizado foi 0,4 mL/min, sendo o comprimento da corrida 1,5 *cv*, tendo sido aplicado aproximadamente 2,2 mg de proteína.

Para os testes de purificação da construção MBP-LGI1C/GroEL por troca aniônica, utilizou-se as colunas HiTrap DEAE FF e HiTrap Q FF de 1 mL, e as colunas HiTrap CM FF e HiTrap SP FF, ambas de 1 mL também (GE Healthcare), para troca catiônica. Foram utilizados ainda os tampões 6 e 7 com baixa concentração de sal (tampões 6A e 7A, respectivamente, com NaCl 20 mM) e a eluição em um gradiente crescente de sal (NaCl 1M). Em relação aos testes de purificação da construção MBP-LGI1C, utilizou-se as colunas HiTrap DEAE FF e HiTrap Q FF,

ambas de 1 mL, utilizando-se apenas o tampão 7A. Todas essas corridas foram realizadas em FPLC.

SDS-PAGE não reduzido

Para se avaliar o estado de agregação da construção MBP-LGI1C com ou sem a co-expressão do sistema GroEL-GroES, realizou-se eletroforese em gel de poliacrilamida sem a adição de agente redutor ao tampão de amostra (Ausubel *et al.*, 1998). Além disso, as amostras analisadas não foram desnaturadas a 90°C, como comumente se realiza no sistema SDS-PAGE reduzido.

Análise por western blot

Utilizou-se um antisoro anti-LGI1 gentilmente fornecido pelo Dr. Jordi Pérez-Tur (Instituto de Biomedicina de Valência; veja Morante-Redolat *et al.*, 2002). O antisoro foi purificado utilizando uma coluna HiTrap rProteína A FF de 1 mL (GE Healthcare), em cromatógrafo FPLC (Amersham Biosciences).

O anticorpo secundário (*anti-rabbit* IgG) conjugado à fosfatase alcalina foi adquirido da Sigma-Aldrich e a membrana Hybond-C da GE Healthcare.

Seguiu-se protocolo de Sambrook *et al.* (1989). De maneira sucinta: entre 40 e 60 µg de proteína, quantificados através de absorbância a 280 nm, foram fracionados em gel de SDS-PAGE 10% e transferidos para as membranas utilizando um aparato horizontal *electroblot* semi-seco (Amersham Biosciences). Após a transferência, as membranas foram bloqueadas em temperatura ambiente por 2 h com TBS leite (TBS 1X: Tris-HCl 20 mM, pH 7,6, NaCl 140 mM,

acrescido de leite desnatado desidratado 5% (m/v) para reduzir ligações inespecíficas). Para o antisoro anti-LGI1 purificado utilizou-se a diluição 1:250 em TBS leite, incubando-se com as membranas por 16 h em temperatura ambiente, em agitação constante. As membranas foram lavadas duas vezes com TBS 1X antes de incubá-las com o anticorpo secundário (1:20.000) por 1 h em temperatura ambiente. As bandas imunoreativas foram detectadas usando como substrato BCIP/NBT (Fermentas) para a fosfatase alcalina.

Digestão em gel e análises por espectrometria de massas

Realizou-se a digestão das bandas ou *spots* de interesse essencialmente como descrito por Smolka (2002). Resumidamente, os *spots* foram picotados, lavados por 10 min com acetonitrila 50% e bicarbonato de amônio 25 mM, em seguida lavados com água Milli-Q, e depois lavados duas vezes 10 min cada com acetonitrila 100% (para remoção do corante azul de coomassie). Desidratou-se os mesmos em *speedvac* por 40 min. Os *spots* foram reidratados em solução de tripsina modificada grau seqüenciamento (Promega) 1 µg em 80 µL de solução básica de bicarbonato de amônio. A reação foi incubada em gelo por outros 40 min e em seguida aquecida a 37°C por 15 h.

Após esse período de digestão enzimática, os peptídeos trípticos foram extraídos do gel por eluição passiva. Recuperou-se a solução em tubo eppendorf novo, contendo acetonitrila 50% e ácido fórmico 5%. Ao gel restante, adicionou-se bicarbonato de amônio 50 mM por 20 min. Esta solução foi recuperada. Ao gel adicionou-se ainda acetonitrila 50% e ácido fórmico 5% por duas vezes e recuperou-se a solução. Os peptídeos foram concentrados em *speedvac* até o volume de aproximadamente 20 µL e congelados a -20°C.

As análises de *fingerprinting* dos peptídeos trípticos (PMF, *peptide mass fingerprinting*) (Shevchenko *et al.*, 1996) foram realizadas no espectrômetro de massas Voyager-DE PRO MALDI-TOF (Applied Biosystems) ou Q-TOF Ultima API (Waters/Micromass). No caso da análise por MALDI-TOF, 0,5 µL da amostra foi plaqueada com 0,5 µL da solução de matriz, composta de ácido α -ciano-4-hidroxi-cinâmico (Sigma) a 10 mg/mL em água/acetonitrila (1:1, v/v) e 0,1% ácido fórmico. No caso da análise por electrospray, os peptídeos trípticos foram aplicados e desalinizados numa coluna Waters Opti-Pak C18. A amostra foi eluída da coluna a 250 nL/min utilizando como fase móvel água/acetonitrila (1:1, v/v) e ácido fórmico 0,1%. As condições utilizadas do espectrômetro foram 3 kV para o *spray*, 100 V para o cone e o gás do cone num fluxo de 30 L/h, sendo a temperatura da fonte 100°C. O espectro final foi processado usando o programa MaxEntI (Waters/Micromass).

3.2 RESULTADOS E DISCUSSÃO

Clonagens iniciais dos genes LGIs e subclonagem de LGI1 e porções truncadas

As seqüências codificadoras das proteínas LGI1 (AA 1-557), suas formas truncadas N- (AA 1-222) e C-terminal (AA 224-557), LGI2, LGI3 e LGI4 foram clonadas em pGEMT-easy, sendo que a presença dos insertos foi confirmada por digestão enzimática ou PCR de colônia e sua integridade por seqüenciamento automático. Encontrou-se especial dificuldade na clonagem das seqüências codificadoras de LGI3 e LGI4. De acordo com Senechal e colaboradores (2005), o cDNA completo da LGI4 mostrou-se bastante instável nas cepas bacterianas utilizadas por eles.

Subclonou-se LGI1, LGIN e LGIC em pET32a e LGI1 e LGIC em pGEX-4T1. As tentativas de subclonagem do fragmento N-terminal em pGEX-4T1 falharam, assim como em pSV282. Muito provavelmente devido à formação de estrutura secundária do inserto de DNA (Zuker, 2003), o que impossibilitou a subclonagem desta porção da proteína.

Procurou-se ainda a subclonagem do gene inteiro e da porção correspondente ao domínio epitempina no vetor pSV282, cedido pela Prof^a. Laura Mizoue, do Centro de Biologia Estrutural da Universidade de Vanderbilt. Este vetor permite a expressão de proteínas fusionadas ao C-terminal de MBP, e entre a proteína carreadora e a proteína passageira, há a presença de um sítio de clivagem para a protease TEV. Entretanto, o sítio de clonagem 5' do vetor é *BamHI*, e o gene *LGI1* apresenta este mesmo sítio no final de sua seqüência codificadora, região correspondente à porção C-terminal da proteína (anexo 1, sítio indicado). Por isso a impossibilidade de se utilizar desse sítio para clonagem. Diversas tentativas de digestão parcial foram infrutíferas.

Testes de expressão e de purificação

Os clones dos vetores pET32a não expressaram a proteína recombinante. A fim de sanar qualquer dúvida quanto à sua expressão, foi realizado um teste de purificação por afinidade, através da resina Ni-NTA, tanto para a proteína inteira quanto para as construções LGI1N e LGI1C. Entretanto, confirmou-se a não expressão destas mesmas (figuras 3 e 4).

Como a proteína inteira fusionada à GST é solúvel em sistema de expressão eucariótico (site da Abnova; veja endereços eletrônicos), investiu-se na expressão da forma inteira e do C-terminal fusionados a GST, mas, utilizando sistema de expressão procariótico. Como dito anteriormente, o clone N-terminal não foi obtido. A proteína inteira fusionada à cauda de GST tem massa molar predita de 90 kDa e a construção LGI1C fusionado à cauda têm massa molar predita de 65 kDa. Todavia, em ambos os casos, a proteína foi expressa na forma insolúvel (figura 3). Testaram-se diversas temperaturas, privilegiando as mais baixas, de forma a favorecer sua solubilização. Realizou-se um teste de purificação com a resina Glutathione Sepharose 4B, mas não foi possível recuperar a proteína na fração solúvel (dados não mostrados).

A expressão de proteínas humanas em sistema procariótico é de fato desafiadora (Hannig & Makrides, 1998; Baneyx, 1999; Terpe, 2003; Sorensen & Mortensen, 2005a e 2005b). A proteína humana LGI1 ainda apresenta algumas particularidades que dificultam consideravelmente sua expressão na forma solúvel em *Escherichia coli*: é uma proteína relativamente grande (aproximadamente 63 kDa); sabe-se que é secretada (Senechal *et al.*, 2005) e glicosilada em três sítios (Sirerol-Piquer *et al.*, 2006) e possui ainda treze resíduos de cisteína (anexo 2). Além das características intrínsecas da proteína, a bactéria *E. coli* é incapaz de realizar modificações pós-traducionais, como a N-glicosilação, sendo sua principal desvantagem a possibilidade de haver um incorreto enovelamento das proteínas heterólogas, o que exige muitas

vezes a aplicação de um protocolo de solubilização e re-enovelamento (Clark, 1998; Lilie *et al.*, 1998; Carrio & Villaverde, 2002).

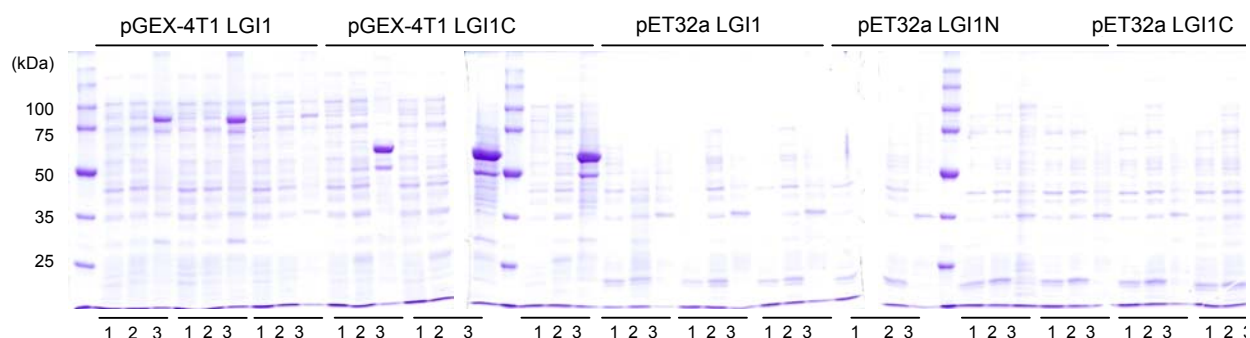


Figura 3: Análise em gel SDS-PAGE 10% para verificação das frações solúveis e insolúveis dos testes de expressão com os clones fusionados a GST (pGEX-4T1) e Trx (pET32a). Para cada construção testou-se a expressão de três clones distintos. Legenda: 1, fração não induzida; 2, fração induzida; 3, fração insolúvel. As massas moleculares previstas para cada construção são as seguintes: GST-LGI1: 90,4 kDa; GST-LGI1C: 65,7 kDa; Trx-LGI1: 81,7 kDa; Trx-LGI1N: 42,6 kDa e Trx-LGI1C: 57,0 kDa. Observa-se que as fusões com GST são insolúveis (canaletas 3 de cada clone testado) e que não houve expressão protéica com as fusões com Trx. Analisou-se a temperatura de 25°C neste gel.

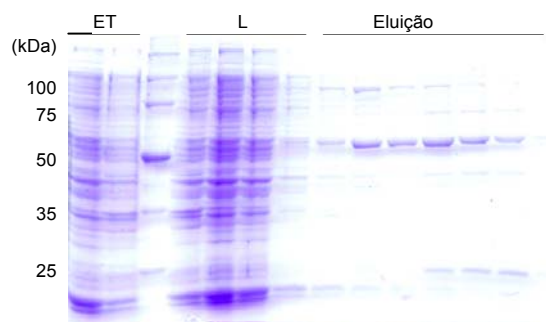


Figura 4: Análise em gel SDS-PAGE 10% do teste de purificação da proteína Trx-LGI1. A massa molar esperada é de 81,7 kDa. Legenda: ET, extrato total; L, lavagens com tampão de extração. Eluição: com gradiente crescente de imidazol (de 5 a 200 mM). Observa-se que não há a proteína no tamanho esperado, apenas os contaminantes habituais de *E. coli* estão presentes no gel. Os testes de purificação com as construções Trx-LGI1N e Trx-LGI1C apresentaram o mesmo perfil protéico, indicando que também não foram expressas.

Deve-se ainda levar em consideração o caráter redutor do citoplasma desta bactéria (Clark, 1998; Lilie *et al.*, 1998), já que a formação de pontes dissulfeto normalmente não ocorre neste compartimento celular. Outra característica relevante de *E. coli* é seu “*codon usage*”: muitos códons utilizados em organismos eucariotos são raros em procariotos (Nakamura *et al.*, 2000). Os aminoácidos menos freqüentemente utilizados em *E. coli* são arginina, isoleucina, prolina e serina, bastante comuns em organismos eucariotos (Zhang *et al.*, 1991; Kane, 1995; Novy *et al.*, 2001). A sequência codificadora da LGI1 apresenta muitos códons raros em *E. coli* (anexo 1).

Como tentativa de favorecer a correta formação das ligações dissulfeto, utilizamos a cepa Origami(DE3), deficiente nos genes *gor* (glutathione redutase) e *trxB* (tioredoxina redutase), que facilitam a formação deste tipo de ligação no citoplasma de *E. coli* (Bessette *et al.*, 1999). Além disso, utilizamos freqüentemente o plasmídio pRARE (Novy *et al.*, 2001) co-transformado nas cepas utilizadas, de forma a aliviar a maquinaria de transcrição com códons raramente usados em *E. coli*.

Novas subclonagens

Subclonou-se com sucesso o gene *LGII* inteiro no vetor de expressão pETSUMO, assim como a porção correspondente ao domínio C-terminal (*LGII*C) nos vetores pETSUMO, pETNusA e pETMBP. Além disso, subclonou-se os genes humanos *Kvβ1.3* e *ADAM22* no vetor pETSUMO para ensaios de interação proteína-proteína.

A proteína Kvβ1.3 (AA 1-401) apresenta solubilidade razoável em bactéria (dados não mostrados). Todavia, não foi possível sua captura através de cromatografia de afinidade em níquel imobilizado. Em relação à proteína ADAM22 (AA 1-859), adotamos como estratégia

inicial a amplificação do fragmento correspondente a sua porção N-terminal extracelular (AA 1-767). Entretanto, a porção 5' da sequência codificadora apresentava porcentagem de GC muito alta, o que impossibilitou a amplificação por PCR. Novo par de oligos foi sintetizado, compreendendo a porção correspondente aos aminoácidos 233-718. A amplificação por PCR foi alcançada e o produto foi corretamente ligado ao vetor pETSUMO. Mas, nos testes de expressão e solubilidade iniciais, a proteína mostrou-se insolúvel (dados não mostrados).

Schulte *et al.* (2006) observaram a associação da LGI1 com o canal de potássio Kv1.1 intracelularmente, através da subunidade Kv β 1.3. Apesar de seu grupo mostrar evidências dessa interação, Fukata e colaboradores (2006) questionaram esses dados e mostraram de maneira bastante elegante e robusta que a LGI1 é secretada pela célula e interage com a proteína ADAM22.

A exportação da LGI1 para fora da célula foi confirmada pelo grupo de Jordi Pérez-Tur (Sirerol-Piquer *et al.*, 2006) e por Senechal e colaboradores um ano antes (2005). Curiosamente, Schulte também identificou a proteína ADAM22 em suas frações de membrana plasmática purificadas através de anticorpo anti-LGI1 imobilizado.

Testes de expressão e de solubilidade dos novos clones

Nos últimos anos, novas caudas de fusão têm sido testadas, não apenas como forma de facilitar a detecção ou a purificação das proteínas de interesse através de cromatografia de afinidade, como se tem tornado claro um efeito positivo no rendimento, solubilidade e mesmo no enovelamento de suas proteínas parceiras (Waugh, 2005). Como evidentemente não existe uma cauda única que satisfaça todos os parâmetros citados, diferentes caudas são combinadas tomando-se partido das vantagens de cada uma. Em uma estratégia genérica (Tropea *et al.*,

2007), utiliza-se uma pequena cauda de poli-histidina (His₆) para facilitar a purificação da proteína recombinante através de IMAC, seguida de MBP para aumentar o rendimento e solubilidade da proteína parceira (Pryor & Leiting, 1997). A proteína, após sua purificação, pode ser clivada *in vitro* com a protease TEV (*tobacco etch virus*), também com cauda His₆ para separar a proteína carreadora e a protease. Em uma abordagem alternativa, pode ser realizada a clivagem *in vivo*, através do que se denominou processamento interno controlado (CIP, *Controlled Intracellular Processing*; Kapust & Waugh, 2000). É possível, através desta metodologia *in vivo*, avaliar se a clivagem da cauda causará problemas futuros de precipitação quando realizados *in vitro*.

Utilizamos esta estratégia genérica, não só com MBP como cauda “solubilizadora”, mas também NusA e SUMO. Estas proteínas, em especial MBP e NusA, são bem conhecidas em sua capacidade de aumentar a solubilidade da proteína parceira, como favorecem também o correto enovelamento das mesmas (Marblestone *et al.*, 2006; Nallamsetty & Waugh, 2006). Assim, diversas cepas, construções e condições de expressão foram testadas a fim de se superar os obstáculos apresentados na expressão da proteína humana LGI1 em sistema procariótico (tabelas 3 e 4).

Testou-se ainda doze diferentes tampões de extração protéica (tabela 5), utilizando a cepa BL21(DE3) pRARE, uma vez que a LGI1 apresenta diversos códons raros em *E. coli* (anexo 1). As construções fusionadas a SUMO foram as mais insolúveis (figura 5), tanto a proteína inteira como sua porção C-terminal. Das três fusões, NusA foi de fato a que apresentou maior capacidade de solubilização para a proteína LGI1C. Dos doze tampões testados, observou-se proteína solúvel em sete condições (figura 6). Entretanto, em uma condição específica (tampão 2), a presença do aditivo arginina tornou sua solubilidade enorme. No tampão 8 também há arginina, mas a quantidade de proteína na fração solúvel é comparável à existente em outras

condições, excetuando-se a condição 2. O tampão 8 é alcalino (pH 8,0), enquanto o tampão 2 é levemente ácido (pH 6,8). Esta solubilidade enorme no tampão 2 por si só atrai atenção, mas há ainda o fato da arginina ser reconhecida como supressora de agregação e reduzir interações proteína-proteína (Tsumoto *et al.*, 2004; Tsumoto *et al.*, 2005). Esta é uma evidência indireta do estado de agregação desta construção: sem a presença do agente supressor de agregação, a maior parte da proteína vai para a fração insolúvel do extrato protéico. Fukata *et al.* (2006) propõe em seu trabalho que a LGI1 é um oligômero, o que vai de acordo com esta evidência.

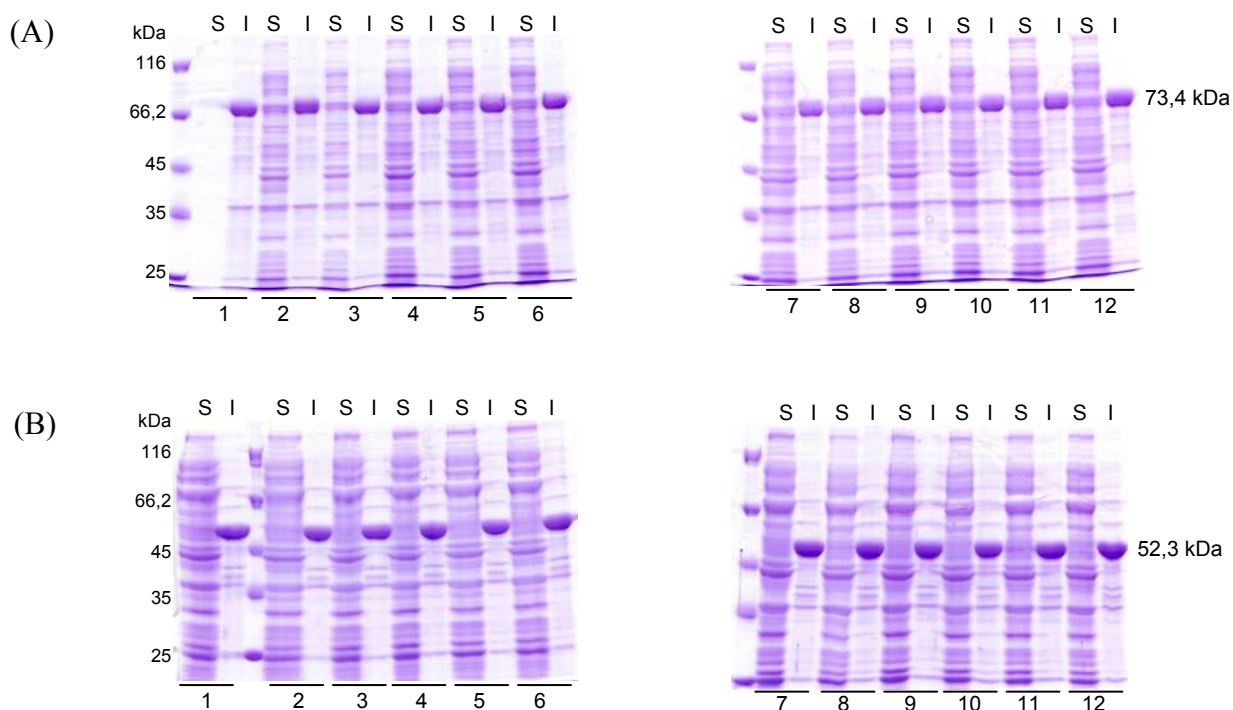


Figura 5: Testes de solubilidade com as proteínas (A) SUMO-LGI1, de aproximadamente 73,4 kDa e (B) SUMO-LGI1C, de aproximadamente 52,3 kDa. Gel SDS-PAGE 10%. Estão indicados abaixo de cada gel os tampões utilizados, numerados de 1 a 12, de acordo com a tabela 5 (p. 24). Legenda: S, fração solúvel; I, fração insolúvel. Marcador de massa molar indicado. A temperatura de indução testada foi de 30°C. Foram ainda feitos testes de purificação por IMAC, mas não se capturou proteína solúvel.

Em relação ao teste de solubilidade da fusão com MBP, não se verificou o mesmo efeito da arginina na solubilidade. A diferença em relação aos outros tampões é apenas modesta, quando comparada com NusA-LGI1C. Todavia, dos doze tampões testados, seis apresentaram proteína solúvel (figura 7). No caso do teste de solubilidade de MBP-LGI1C co-expressando o vetor pTgroE, observou-se que GroEL-GroES é capaz de tornar a construção um pouco mais solúvel que a co-expressão de códons raros em *E. coli*, através do vetor pRARE (figura 8).

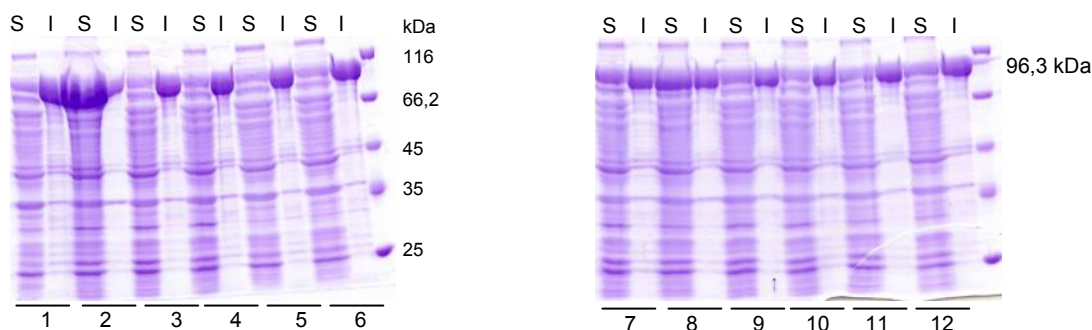


Figura 6: Teste de solubilidade com a proteína NusA-LGI1C, de aproximadamente 96,3 kDa. Gel SDS-PAGE 10%. Estão indicados os tampões utilizados, numerados de 1 a 12 (p. 24), na parte de baixo do gel. Legenda: S, fração solúvel; I, fração insolúvel. Marcador de massa molar indicado. A temperatura de indução testada foi de 30°C. Apesar de apresentar proteína na fração solúvel em mais de um tampão, a condição 2, com arginina, foi a que chamou mais atenção.

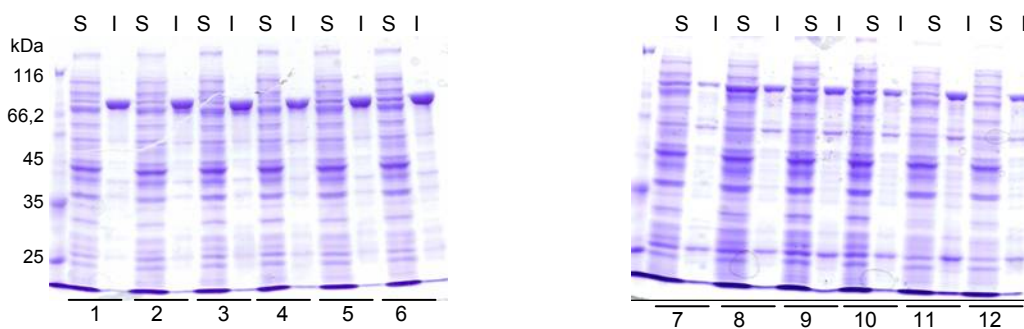


Figura 7: Teste de solubilidade com a proteína MBP-LGI1C, de aproximadamente 82,0 kDa. Gel SDS-PAGE 10%. Estão indicados os tampões utilizados, numerados de 1 a 12 (p. 24), na parte de baixo do gel. Legenda: S, fração solúvel; I, fração insolúvel. Marcador de massa molar indicado. A temperatura de indução testada foi de 30°C. Observou-se apenas uma diminuição na fração insolúvel na condição 7 sem um correspondente aumento da fração solúvel. A fração mais solúvel foi a apresentada pelo tampão 8.

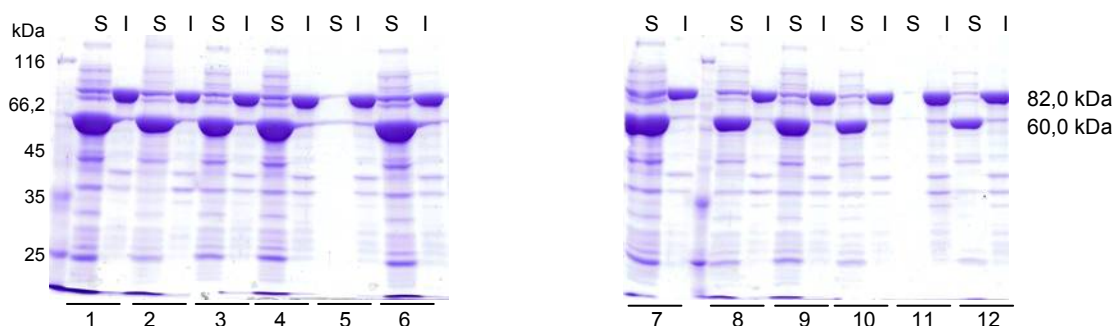


Figura 8: Teste de solubilidade com a proteína MBP-LGI1C, de 82,0 kDa, co-expressando pT-GroE. GroEL têm aproximadamente 60 kDa. GroES têm aproximadamente 14 kDa e não aparece no gel. Gel SDS-PAGE 10%. Estão indicados os tampões utilizados, numerados de 1 a 12, na parte de baixo do gel. Legenda: S, fração solúvel; I, fração insolúvel. Marcador de massa molar indicado. A temperatura de indução testada foi de 30°C. Observa-se um discreto incremento geral na quantidade de proteína solúvel no extrato em relação à utilização do vetor pRARE (figura 7).

Testes de purificação das novas construções

Não foi possível a captura, através de IMAC, das fusões SUMO-LGI1 e SUMO-LGI1C, confirmando o que os testes de solubilidade mostraram: essas construções não são solúveis. Também não foi possível capturar a construção SUMO-Kv β 1.3, apesar de apresentar bom nível de solubilidade no extrato protéico, como julgado por SDS-PAGE. Por se tratar de uma proteína relacionada a canal de membrana, sua expressão e purificação, aliada a um passo cristalográfico, se torna especialmente desafiadora. Em virtude dessas dificuldades associadas a proteínas integrais ou associadas à membrana, o número de estruturas tridimensionais resolvidas em nível atômico permanece insignificante, quando comparado às outras classes de proteínas (Niegowski *et al.*, 2006).

Estas mesmas dificuldades se aplicam à LGI1. Por se tratar de uma proteína humana que sofre modificação pós-traducional (Sirerol-Piquer *et al.*, 2006), sua expressão em procarioto encontra grandes obstáculos (Hannig & Makrides, 1998; Baneyx, 1999; Terpe, 2003; Sorensen

& Mortensen, 2005a e 2005b). A proteína completa apresenta ainda treze resíduos de cisteína, e como o citoplasma bacteriano possui caráter redutor (Clark, 1998; Lilie *et al.*, 1998), a correta formação das ligações dissulfeto é prejudicada, senão impedida.

Já mencionamos que a proteína LGI1 tem sua produção dificultada pelo “*codon usage*” de *E. coli* (Nakamura *et al.*, 2000). Como o número de cisteínas presente pode trazer complicações adicionais, decidiu-se investir sobretudo na região C-terminal da proteína (AA 224-557), que engloba o domínio EPTP. Nesta porção, há apenas três cisteínas e o número de códons raros é aliviado em comparação com a forma completa (quarenta e sete códons raros neste último contra vinte e sete do primeiro; anexo 1). Adicionalmente, as mutações descritas na literatura, os dados apresentados por Fukata *et al.* (2006) e Sirerol-Piquer *et al.* (2006) e as poucas informações à respeito do domínio EPTP foram argumentos decisivos que demonstraram a importância de se conhecer melhor este domínio.

Os testes de captura da proteína NusA-LGI1C foram todos infrutíferos, apesar de seu nível de expressão ter sido o melhor dentre todas as fusões construídas e testadas (figura 6). A presença de arginina no tampão de extração inibiu a captura da cauda His₆ da proteína por IMAC. E, dado inesperado, mesmo em condição desnaturante, não foi possível sua captura (dados não mostrados), provavelmente devido à formação de complexos agregados com ácidos nucleicos. Como a cauda de histidina não estava acessível, buscou-se sua captura através de cromatografia convencional: troca iônica e interação hidrofóbica. Entretanto, igualmente não se obteve sucesso. A provável oligomerização impede sua captura, seja explorando interações iônicas ou hidrofóbicas.

As mesmas dificuldades foram enfrentadas com a afinidade a níquel pela fusão MBP-LGI1C. Seja em condição nativa ou desnaturante, seja partindo-se de corpúsculos de inclusão para renaturação em coluna, nada foi obtido. Tomou-se partido da natureza dupla desta

construção: a proteína ligadora de maltose (MBP) é uma cauda natural de afinidade (Nallamsetty & Waugh, 2006). Testou-se a ligação desta proteína a seu substrato natural maltodextrina (Routzahn & Waugh, 2002) e, finalmente, após essa longa série de testes, obteve-se a captura da proteína recombinante (figura 9). Observou-se níveis diferentes de captura entre MBP-LGI1C pRARE e MBP-LGI1C/GroEL (figura 9). Como num primeiro momento importa mais a quantidade do que a pureza da proteína para estudos estruturais, buscamos separar a chaperona GroEL de MBP-LGI1C. Realizou-se inicialmente uma corrida de filtração em gel na coluna Superdex200 (GE Healthcare), mas não alcançamos a separação entre as duas por este método no tampão utilizado (figura 10); deve-se levar em consideração ainda que são proteínas de tamanhos muito próximos (~82,0 kDa para MBP-LGI1C e ~60,0 kDa para GroEL), diferença pequena para a capacidade de resolução da coluna utilizada.

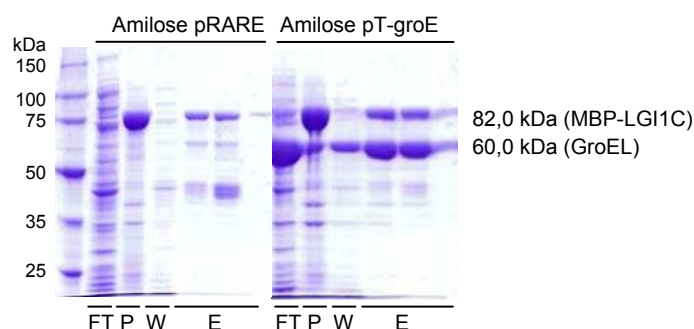


Figura 9: Perfil cromatográfico da purificação de MBP-LGI1C por amilose (à esquerda co-expressando códons raros com pRARE e à direita co-expressando GroEL-GroES com pT-groE). Gel SDS-PAGE 10%. Legenda: FT, proteínas não ligadas após passarem pela coluna de afinidade; P, fração insolúvel; W, lavagem; E, eluição em três passos com maltose 10 mM. Marcador de massa molar indicado. A temperatura de indução utilizada foi de 25°C GroEL auxilia um pouco o aumento da solubilidade da proteína de fusão, quando comparada com pRARE. É possível observar a quantidade habitual de contaminantes que eluem juntamente com a proteína alvo.

Utilizamos em seguida duas colunas de troca aniônica e duas de troca catiônica, em dois tampões diferentes, um alcalino e outro ácido. Oito diferentes corridas cromatográficas foram

realizadas e os melhores resultados ocorreram com a troca aniônica: a amostra interage com a coluna, entretanto, não obtivemos a separação desejada entre GroEL e MBP-LGI1C. (figura 11). Realizou-se o mesmo teste de purificação por troca aniônica com a amostra MBP-LGI1C pRARE, mas não se obteve a pureza, a separação e a quantidade almejadas.

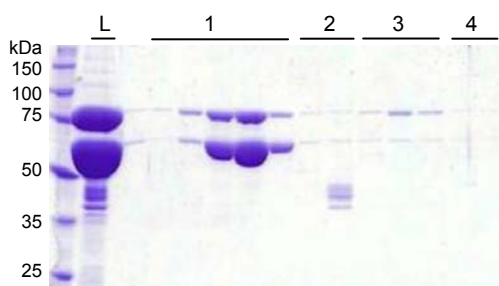


Figura 10: Perfil cromatográfico da purificação de MBP-LGI1C/GroEL por filtração em gel na coluna Superdex 200. Gel SDS-PAGE 10%. Legenda: L, amostra aplicada na coluna; 1, 2, 3 e 4, picos da corrida cromatográfica. Marcador de massa molar indicado. Não foi possível a separação entre GroEL e MBP-LGIC. Apesar de não ter havido uma corrida de calibração da coluna com marcadores de massas conhecidas, é possível verificar que a presença de três picos distintos com perfis semelhantes indicam conformações distintas, ou, em outras palavras, uma população heterogênea de macromoléculas.

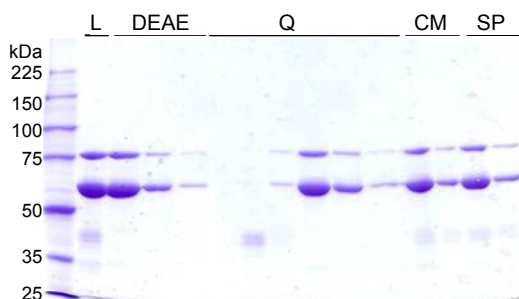


Figura 11: Amostras fracionadas em cromatografia de troca iônica. Gel SDS-PAGE 10%. Legenda: L, amostra aplicada na coluna; DEAE e Q, colunas de troca aniônica; CM e SP, colunas de troca catiônica. Marcador de massa molar indicado. Não se obtém a separação entre GroEL e MBP-LGI1C. Obtêm-se apenas a separação das bandas contaminantes abaixo de 50 kDa, provenientes da purificação por amilose, em coluna HiTrap Q FF (indicado por Q). Adicionalmente, observou-se que, no meio ácido utilizado (pH 5,2), ocorria precipitação mais forte da amostra.

Western blot e gel desnaturante não reduzido

Com o anticorpo anti-LGI1 cedido pelo Dr. Jordi Pérez-Tur e purificado em nosso laboratório, observou-se que as construções NusA-LGI1C e MBP-LGI1C apresentavam diversas bandas de proteínas degradadas (figuras 12). Além disso, nota-se que o padrão de degradação é muito semelhante. No caso da MBP-LGI1C, mesmo com a utilização de inibidor de protease PMSF não se verificou diferença no padrão de degradação. O resultado obtido com *western blot* indica que: a) PMSF não foi eficiente para inibir a degradação observada; b) a proteína é instável. Inesperadamente, o anticorpo anti-LGI1 reconheceu também a chaperonina GroEL. Análises por espectrometria de massas confirmaram a identidade da chaperonina. Provavelmente há a formação de um complexo inespecífico MBP-LGI1C/GroEL/MBP-LGI1C ou GroEL/MBP-LGI1C/GroEL.

Para tentarmos compreender em que estado de oligomerização estavam as amostras de MBP-LGI1C produzidas, realizamos uma corrida de SDS-PAGE não reduzido (figura 13). Após alguns dias a 4°C, algumas amostras passaram a apresentar bandas de alta massa molar, um forte indicativo de agregação (figura 13).

Tivemos assim um quadro bastante amplo do estado da construção MBP-LGI1C: não só apresenta diversas bandas de degradação, como ainda agrega em solução.

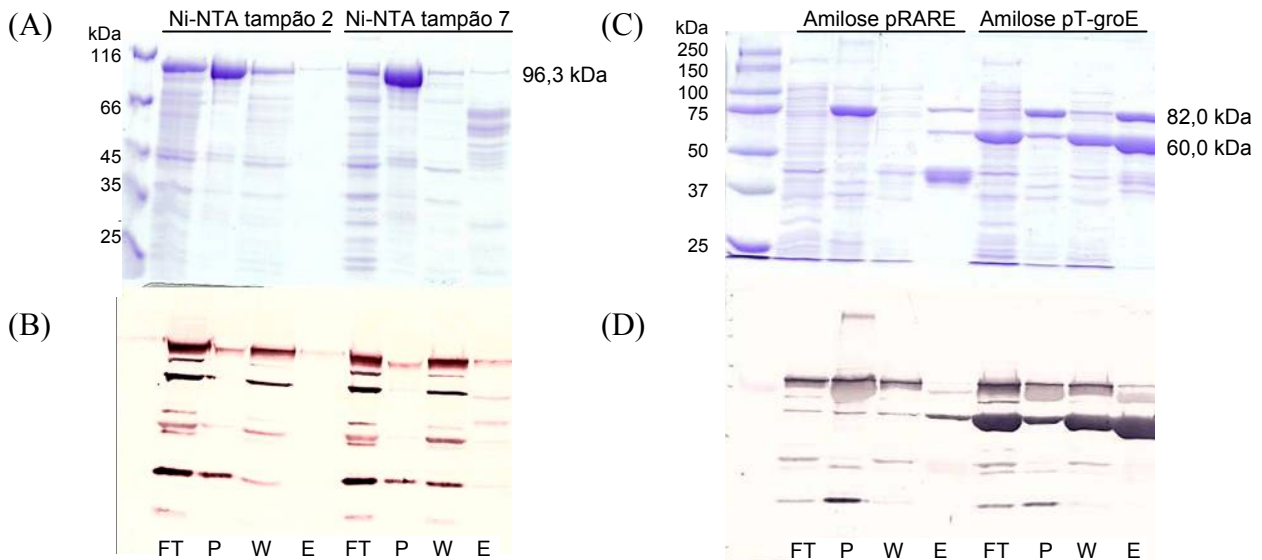


Figura 12: *Western blot* anti-LGI1 contra NusA-LGI1C e contra MBP-LGI1C. (A) Gel SDS-PAGE 10% do teste de purificação de NusA-LGI1C em resina Ni-NTA, utilizando-se os tampões 2 (Bis-Tris 20 mM, pH 6,8, NaCl 150 mM e arginina 100 mM) e 7 (Tris-HCl 20 mM, pH 8,0, NaCl 150 mM), conforme tabela 5. (B) *Western blot* anti-LGI1 do mesmo gel em (A). (C) Gel SDS-PAGE 10% da purificação de MBP-LGI1C por amilose com pRARE e pT-groE. (D) *Western blot* anti-LGI1 do mesmo gel em (C). Legenda: FT, proteínas não ligadas após passarem pela coluna de afinidade; P, fração insolúvel; W, lavagem; E, eluição com eluente específico. Marcador de massa molar indicado. São observadas diversas bandas de degradação, tanto com NusA-LGI1C quanto com MBP-LGI1C. Não foi possível a captura da construção NusA-LGI1C por IMAC. A proteína GroEL (~60,0 kDa) também foi reconhecida pelo anticorpo.

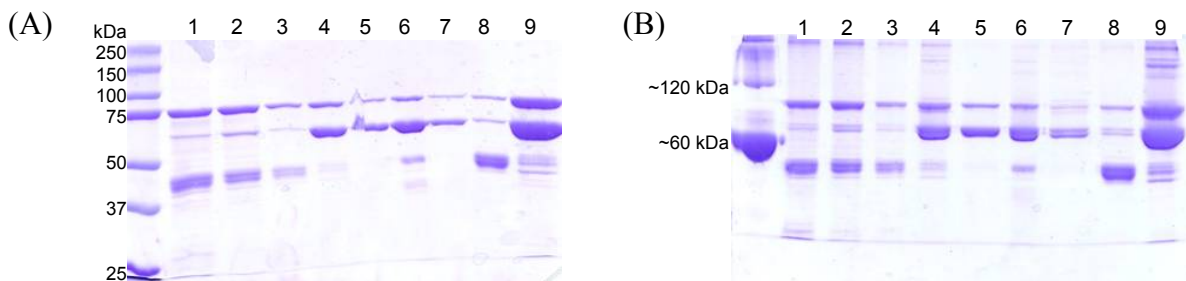


Figura 13: Gel SDS-PAGE 10% de MBP-LGI1C. (A) Gel desnatante reduzido. (B) Gel desnatante não reduzido das mesmas amostras em (A). Legenda: 1, 2 e 3, MBP-LGI1C sem GroEL por 22, 20 e 15 dias, respectivamente, a 4°C; 4, MBP-LGI1C com GroEL por 15 dias a 4°C; 5 e 6, MBP-LGI1C com GroEL, em tampão 6 e 7 (p. 24); 7, MBP-LGI1C com GroEL aplicada em DEAE FF; 8, MBP-LGI1C sem GroEL aplicada em Q FF; 9, MBP-LGI1C com GroEL por 6 dias a 4°C. Marcador de massa molar indicado em (A) e albumina bovina, BSA, como padrão em (B), que apresenta forma monomérica de aproximadamente 60 kDa e outra dimérica de aproximadamente 120 kDa. Observa-se em (B) a presença de diversas bandas de alta massa molar.

3.3 ANÁLISES ADICIONAIS FINAIS

Na busca por dados estruturais da proteína LGI1, submetemos a amostra MBP-LGIC sem GroEL a análises por espalhamento de luz dinâmico (DLS). A proteína fresca foi dialisada por 16 h em diferentes tampões, a fim de se determinar um tampão ótimo, ou ao menos se aproximar de um (tabela 6). Adicionalmente, obteríamos informações diretas de seu estado de agregação. As amostras foram centrifugadas previamente a 20.000 g por 10 min a 4°C e analisadas a 25°C no equipamento DynaPRO-MS/X (Protein Solutions) com temperatura controlada. Os dados foram obtidos através do programa DYNAMICS versão 6.2.

Nº da amostra	Condição	Tampão	Volume utilizado
1	fresca	Tris-HCl 20 mM, pH 8,0	150 µL
2	fresca	Tris-HCl 20 mM, pH 8,0, NaCl 20 mM	150 µL
3	fresca	Tris-HCl 20 mM, pH 8,0, NaCl 20 mM, DTT 1 mM	150 µL
4	fresca	Tris-HCl 20 mM, pH 8,0, NaCl 150 mM	150 µL
5	fresca	Tris-HCl 20 mM, pH 8,0, NaCl 500 mM	150 µL
6	fresca	Tris-HCl 20 mM, pH 8,0, glicerol 5%	150 µL
7	fresca	ácido acético 20 mM, pH 5,2	150 µL

Tabela 6: Amostras analisadas por DLS.

O espalhamento de luz é uma das maneiras mais comumente utilizadas para se caracterizar macromoléculas. Utilizando equipamentos modernos de DLS, é possível obter uma massa molar média, o coeficiente de difusão e o raio hidrodinâmico da proteína (Bloomfield, 2000). Adicionalmente, requer volumes pequenos (no mínimo 60 µL, conforme a cubeta utilizada) e quantidades modestas de proteína (em geral se utiliza a 1 mg/mL), além da rapidez na aquisição e simplicidade de manipulação da técnica.

Por ser uma técnica não destrutiva e não invasiva, este método possibilita ainda o estudo da proteína em solução, em condições próximas da fisiológica (Murphy, 1997); o estudo de agregação protéica (Philo, 2006), uma ampla gama de processos dinâmicos (Suijsa *et al.* 2007), como método diagnóstico para o crescimento de cristais de proteína (Wilson, 2003), além de permitir o estudo de estabilidade estrutural, desnaturação térmica, conformação, entre outros. Adicionalmente, o DLS é bastante sensível a pequenas quantidades de agregados (Papish *et al.*, 2002).

O comportamento hidrodinâmico e a intensidade de luz espalhada são funções não lineares do tamanho, devido a interações hidrodinâmicas no caso de partículas compostas por múltiplas subunidades, e interferência intrapartícula caso a partícula apresente tamanho comparável ao comprimento de onda da luz espalhada (Bloomfield, 2000). O coeficiente de difusão depende da concentração protéica. Através deste coeficiente é possível determinar o raio hidrodinâmico da proteína. Assim, a intensidade do espalhamento mensurado é diretamente proporcional à massa molar, que por sua vez é proporcional ao raio.

Em todas as amostras de proteína analisadas obtivemos um raio de giro entre 10 e 100 nm (figura 14), indicando agregação.

Foram realizadas ainda análises de espalhamento de raios X a baixo ângulo (SAXS), que permite o conhecimento da estrutura tridimensional protéica em baixa resolução. É possível obter parâmetros moleculares, como massa molar, raio de giro (R_g) e dimensão máxima (D_{max}).

O desenvolvimento de novos algoritmos na década passada permitiu a construção *ab initio* dos mapas de densidade eletrônica de baixa resolução tridimensional a partir do perfil de espalhamento unidimensional (Svergun, 1999; Walther *et al.*, 2000; Svergun *et al.*, 2001), obtendo-se informações estruturais além dos parâmetros usuais. Foram criadas, recentemente, ferramentas para modelar proteínas e complexos moleculares a partir das informações obtidas

por SAXS, basendo-se nas estruturas conhecidas ou parcialmente conhecidas através de experimentos de alta resolução (Lipfert & Doniach, 2007).

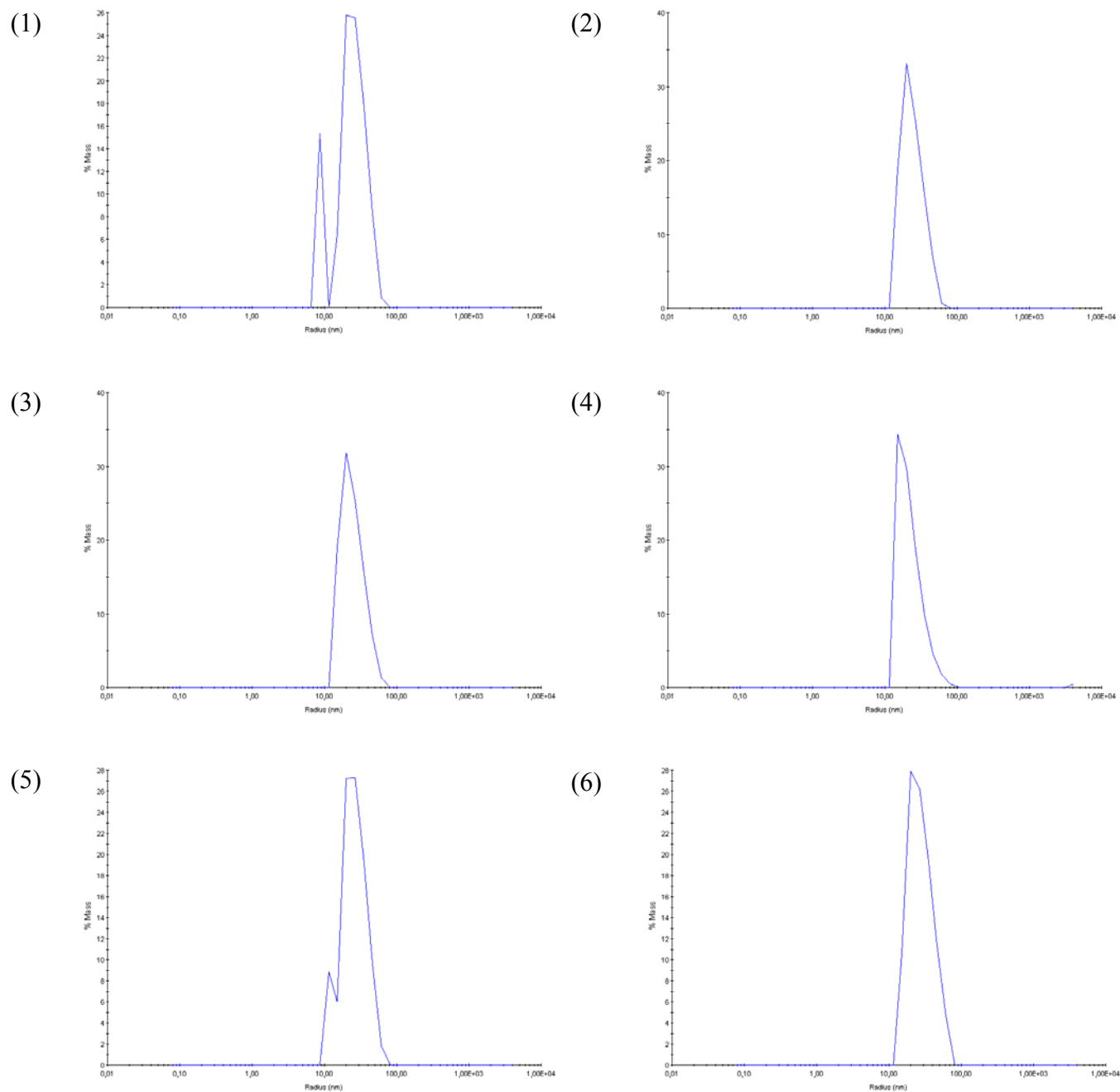


Figura 14: Espectros raio (nm) *versus* porcentagem de massa. Números correspondentes ao número das amostras da tabela 6. Não foi possível adquirir os dados apenas para a amostra 7 (em ácido acético 20 mM, pH 5,2).

O perfil de intensidade de raios X espalhado também pode ser descrito em termos da função de distribuição de distâncias $p(r)$ (Glatter & Kratky, 1982). Esta função 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 das distâncias r que podem ser encontradas a partir da combinação de qualquer par de elementos de volume. A função $p(r)$ pode ser obtida através da transformada indireta de Fourier a partir do perfil de intensidade de raios X espalhado usando o software GNOM (Svergun, 1992). Atualmente, a função $p(r)$ pode ser analisada por algoritmos que fornecem o mapa da densidade eletrônica da molécula a partir do perfil de SAXS unidimensional. A visualização e manipulação de estruturas atômicas e modelos de baixa resolução, a partir de modelos de alta resolução, podem ser obtidas através dos programas CRY SOL (Svergun *et al.*, 1995) e GNOM (Svergun, 1992). Porém, quando não há 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) e outros.

Nossas análises na nova linha de SAXS do Laboratório Nacional de Luz Síncrotron, utilizando a amostra MBP-LGI1C sem a presença de GroEL, confirmaram mais uma vez que a proteína estava fortemente agregada. Como indicado não apenas por este experimento, há necessidade de se otimizar o tampão final em que a proteína recombinante estará dissolvida, de forma a mantê-la no estado monodisperso, ainda que forme oligômero.

Os últimos experimentos realizados buscaram analisar a interação *in vivo* entre as proteínas GST-LGI1C X SUMO-Kv β 1.3 e GST-LGI1C X SUMO-ADAM22 (porção extracelular). GST-LGI1C apresentou-se insolúvel (figura 3, pGEX-4T1 LGI1C), enquanto SUMO-Kv β 1.3, apesar de solúvel, não foi possível ser capturada por IMAC (dados não

mostrados). Caso a presença do possível ligante auxiliasse a solubilizar a proteína de interesse, poderíamos esperar que sua captura passasse a se tornar possível através de cromatografia de afinidade, ou ainda, que seu estado de agregação se alterasse. Um resultado semelhante poderia ser esperado para a interação GST-LGI1C X SUMO-ADAM22, apesar de construção SUMO-ADAM22 ter se apresentado completamente insolúvel (dados não mostrados). Algumas questões interessantes poderiam ser respondidas, trazendo evidências das possíveis interações, e outras questões mais poderiam ser formuladas. Entretanto, os testes iniciais não foram conclusivos e mais controles serão necessários (dados não mostrados).

3.4 CONCLUSÕES

Baseando-nos na estratégia adotada pelo Laboratório de Biologia Molecular Europeu (EMBL, *European Molecular Biology Laboratory*) (Dümmler *et al.*, 2005; Peti & Page, 2007) buscamos implementar em nosso laboratório a clonagem em diversos vetores de expressão bacterianos, que codificam diferentes caudas de fusão, a fim de se obter não somente mais proteína solúvel, mas também corretamente enovelada. Busca-se não apenas utilizar a mesma enzima de clonagem no sítio 5', mas testar diferentes cepas de *E. coli* e otimizar diferentes condições de expressão, como concentração do indutor gratuito, temperatura de expressão, co-expressão de chaperoninas, expressão de fragmentos da proteína, repressão da expressão basal, entre outras. Houve uma mudança de foco na produção de proteínas heterólogas em bactéria, passando-se da alta produção protéica, que geralmente acaba sobrecarregando o sistema bacteriano de transcrição e tradução, para um correto enovelamento protéico (Nominé *et al.*, 2001; de Marco, 2004). Evita-se assim, quando possível, a produção de intermediários desenovelados agregados ou não enovelados corretamente, que podem resultar na formação de corpúsculos de inclusão (Schröder & de Marco, 2005).

Inicialmente escolhemos três diferentes fusões para a expressão da LGI1, seus clones truncados (figura 15), e dos outros membros da família LGI (LGI2, LGI3 e LGI4): His₆, no N-terminal da proteína de interesse e, conforme o caso, também no C-terminal da mesma, GST e Trx. As dificuldades encontradas em se obter proteína solúvel logo de início com GST e a não expressão das proteínas fusionadas à Trx nos levaram a abandonar a clonagem em pET28a e pET29a e a buscar caudas alternativas, que não apenas ajudassem na purificação da proteína, como auxiliassem o correto enovelamento e o aumento da solubilidade do alvo.

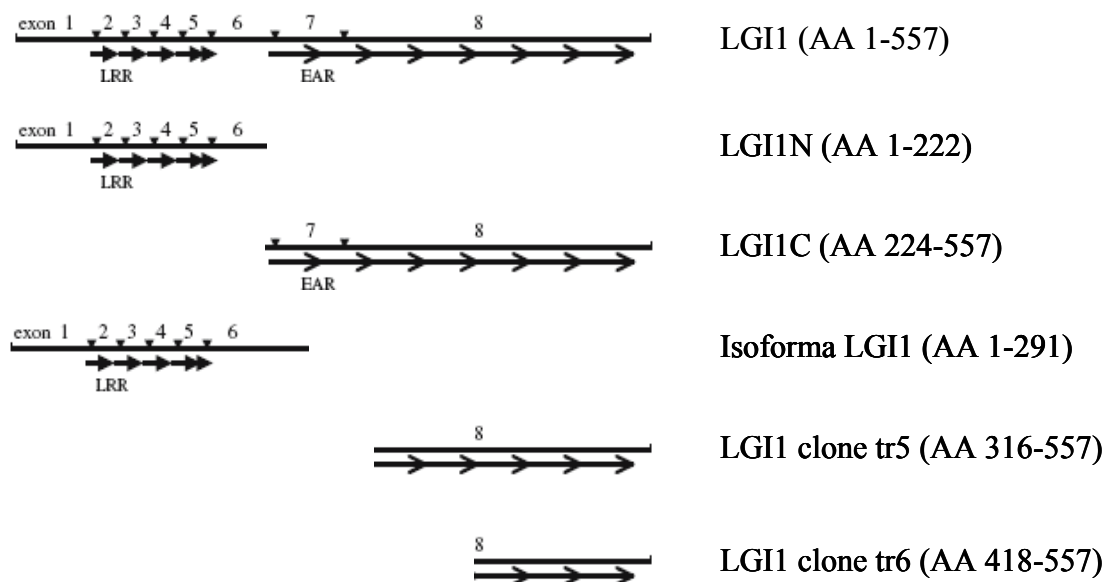


Figura 15: Esquema inicialmente adotado dos clones truncados da proteína humana LGI1.

Concentramos nossos esforços em três clones: LGI1 em sua forma inteira, LGI1N e LGI1C. Abandonou-se os outros membros da família LGI, os clones truncados e a isoforma de aproximadamente 800 bp que foi amplificado a partir de RNAm total de hipocampo humano. Partimos para a clonagem no vetor pSV282. Entretanto, o sítio de clonagem 5' do vetor era justamente uma enzima que clivava internamente nosso inserto de interesse (*BamHI*, anexo 1). Outros sítios de restrição do vetor ou aumentavam demasiadamente a quantidade de aminoácidos estranhos à proteína após a clivagem com a protease específica, ou tiravam da pauta de leitura nossa sequência codificadora. Diversas tentativas foram feitas visando clivagem parcial do inserto, mas sem sucesso, tanto para a proteína inteira quanto para a porção C-terminal. Não obtivemos nenhum clone da porção N-terminal neste vetor, pela dificuldade da ligação entre inserto e vetor, aparentemente devido à formação de estruturas secundárias do inserto.

Novos vetores de expressão foram obtidos (pETSUMO, pETMBP e pETNusA) e à medida em que clones positivos foram confirmados, realizou-se testes de expressão protéica. Somando-se a essa estratégia de clonagem em diferentes vetores paralelamente, obtivemos resultados interessantes com testes de solubilidade, empregando-se diferentes tampões de extração (tabela 5).

Obtivemos expressão protéica com as seguintes construções: SUMO-LGI1, SUMO-LGI1C, NusA-LGI1C e MBP-LGI1C. A porção N-terminal da proteína, apesar de interessante, apresenta dez cisteínas (anexo 2, resíduos de cisteína indicados), o que dificulta sobremaneira a expressão de proteína solúvel no citoplasma bacteriano, uma vez que esse local não é adequado para a formação de pontes dissulfeto (Bulaj, 2005; Singh & Panda, 2005; Clark, 1998; Lilie *et al.*, 1998). Por isso, os esforços foram dirigidos para a porção C-terminal da LGI1.

Apesar da obtenção de expressão protéica, apenas com as fusões NusA-LGI1C e MBP-LGI1C foi possível observar proteína solúvel, mostrando que a proteína carreadora ajuda de maneira caso-específico a proteína passageira (Nallamsetty & Waugh, 2006). Além disso, um dado indireto mostrou que a fusão NusA-LGI1C estava bastante agregada: no tampão com arginina 100 mM, em pH 6,8, observou-se uma quantidade enorme de proteína solúvel. O mesmo não foi verificado em pH 8,0. Adicionalmente, a presença da cauda NusA favoreceu a solubilização da fusão, uma vez que a construção MBP-LGI1C nessa mesma condição não resultou em aumento de solubilidade. A arginina é conhecido supressor de agregação (figura 6, tampão 2; Tsumoto *et al.*, 2004; Tsumoto *et al.*, 2005); todavia, apesar de favorecer a não agregação da proteína de interesse, impede sua captura por afinidade a níquel. Testes de captura através de cromatografia de troca iônica e interação hidrofóbica foram igualmente infrutíferos. Mesmo a cromatografia de afinidade em condição desnaturante, seguindo o protocolo da Qiagen, não conseguiu capturar a proteína NusA-LGI1C.

Investiu-se na fusão MBP-LGI1C, co-expressando as chaperoninas GroEL-GroES ou procurando aliviar a maquinaria de tradução bacteriana com a transcrição de códons raros de eucariotos, a partir do vetor pRARE. Novamente, apesar da proteína estar na fração solúvel, sua captura por afinidade a níquel foi impossível. Mesmo testando o re-enovelamento em coluna não se obteve um resultado satisfatório (dados não mostrados). Testou-se ainda a captura por cromatografia de troca aniônica, catiônica e interação hidrofóbica. Igualmente sem sucesso. Buscou-se, assim, a captura por afinidade através da resina de amilose, tirando proveito da afinidade natural da proteína carreadora MBP à maltodextrina (Boos & Shuman, 1998). Obtém-se uma quantidade apenas razoável quando há co-expressão das chaperoninas moleculares GroEL-GroES; porém, superior quando comparada com a co-expressão do vetor pRARE (figura 9). O mecanismo pelo qual o complexo molecular de chaperonas cataliza o correto enovelamento do polipeptídeo nascente (Yasukawa *et al.*, 1995) ainda não é completamente conhecido (Lin & Rye, 2006). Apesar do nível de impurezas eluídas juntamente com a proteína passageira neste tipo de purificação ser inaceitável para estudos visando cristalização (figura 9) e a eficiência da ligação da proteína MBP à resina de amilose não ser tão grande (Routzahn & Waugh, 2002), a captura da mesma foi um bom resultado alcançado após as mais diversas tentativas frustradas de captura da construção.

Outra dificuldade encontrada foi separar a proteína MBP-LGI1C e a chaperonina GroEL. GroEL se liga a porções hidrofóbicas da proteína nascente, auxiliando seu correto enovelamento (Zhang *et al.*, 2002; Tang *et al.*, 2006). Provavelmente tínhamos em solução um complexo GroEL/MBP-LGI1C/GroEL ou MBP-LGI1C/GroEL/MBP-LGI1C. De qualquer maneira, as tentativas de separação das duas proteínas por meio de troca iônica ou interação hidrofóbica falharam, assim como filtração em gel. Este complexo foi observado através de *western blot*

(figura 12), onde o anticorpo contra a LGI1 reconheceu a proteína GroEL. A identidade da GroEL foi confirmada através de espectrometria de massas.

Testou-se ainda a clivagem da proteína carreadora MBP com a protease TEV em duas concentrações distintas. Apesar da reação de clivagem ter sido bastante eficiente, contrariamente ao esperado (dados não mostrados), e a maior parte da proteína LGI1C clivada ter permanecido em solução, a quantidade final foi pequena, e não se consegue pureza suficiente para testes posteriores.

Foram realizados experimentos de SAXS com a proteína MBP-LGI1C produzida sem a presença da GroEL, assim como DLS. Obtivemos, assim, dados diretos do estado de agregação da mesma, confirmando as evidências indiretas observadas anteriormente.

Nos últimos experimentos realizados, co-expressamos na cepa Origami2(DE3) a fusão GST-LGI1C (Amp^R) juntamente com SUMO-ADAM22 (Kan^R), assim como GST-LGI1C com SUMO-Kvβ1.3, ambos na presença do vetor pRARE, uma vez que não apenas LGI, mas também ADAM22 e Kvβ1.3 possuíam códons raros em bactéria. Todavia, infelizmente os resultados não foram conclusivos (dados não mostrados), necessitando de mais controles, tanto positivos quanto negativos.

Empreendeu-se grande esforço na tentativa de se obter a proteína humana LGI1, ou sua porção C-terminal, em estado solúvel em sistema procariótico, a fim de se obter dados estruturais da mesma visando maior compreensão da epileptogênese associada à EPADSA. Utilizando as mais diversas construções (anexo 3) com as proteínas de fusão mais comumente utilizadas em biologia molecular estrutural, além de testar diferentes cepas de *E. coli* e condições de expressão e purificação, pudemos obter a proteína LGI1C solúvel com duas diferentes fusões, NusA e MBP, abrindo possibilidades de caracterização estrutural desta proteína relacionada a um distúrbio neurológico importante como a epilepsia.

3.5 PERSPECTIVAS

O próximo passo importante a ser dado em relação à proteína LGI1 num contexto de estudo de biologia molecular estrutural é superar os problemas decorrentes do estado de agregação da mesma. Sua forte agregação ocorre dentro da célula hospedeira (uma vez que não éramos capazes de separar GroEL e MBP-LGI1C, indicando forte associação entre as duas proteínas) e não se encontrou dentre as condições utilizadas nesta tese de doutoramento uma que fosse capaz de superar esse obstáculo. Ou deverá ser focalizada essa questão ainda dentro da célula, diminuindo, por exemplo, a velocidade do metabolismo bacteriano (fazendo a expressão protéica a 4°C, por exemplo), a concentração de indutor gratuito, ou novas células hospedeiras poderão ser testadas, como leveduras ou células de inseto, uma vez que estas últimas são capazes de realizar modificações pós-traducionais, inexistentes em *E. coli*.

Neste sentido, Nominé *et al.* (2001) sugerem a utilização da monodispersidade como critério para checar a qualidade das proteínas de fusão e para a otimização da expressão e purificação, uma vez que solubilidade por si só nem sempre corresponde a correto enovelamento. Caso os testes futuros de clivagem da cauda MBP com a protease TEV continuem mostrando baixo rendimento e se a questão de agregação (ou monodispersidade) da amostra for superada, a construção MBP-LGI1C poderá ser submetida diretamente a testes de cristalização (Smyth *et al.*, 2003), o que poderá facilitar a resolução do problema das fases numa eventual cristalização da proteína quimera (Taylor, 2003), uma vez que a estrutura tridimensional de MBP é conhecida. Adicionalmente, deve ser realizada ainda uma caracterização espectroscópica da proteína clivada, pois poderá complementar os dados de cristalografia e auxiliar a compreensão da estrutura tridimensional.

Como não poderia deixar de ser, uma vez que os estudos estruturais têm como fim a compreensão funcional das proteínas, experimentos de duplo híbrido de levedura devem ser realizados (Fields & Song, 1989), a fim de se conhecer com quais proteínas a LGI1 interage e mesmo para confirmar as interações descritas na literatura. Apesar dos problemas deste sistema, por exemplo, a alta taxa de falso-positivos e a questão da limitada análise proteína-proteína, ainda é um método bastante poderoso para se estabelecer, ainda que um a um, os componentes de uma rede de comunicação protéica (Brent & Finley Jr., 1997; Crieckinge & Beyaert, 1999). Além disso, podem ser utilizados clones truncados, por exemplo, pode-se utilizar apenas o domínio LRR ou o domínio EPTP, ou ainda a isoforma com *splice* alternativo capturada através de PCR.

Ainda em relação aos possíveis parceiros de LGI1, caso seja confirmada sua interação com a subunidade do canal de potássio Kv β 1.3, esforços deverão ser direcionados para a captura desta proteína. Wennerstrand *et al.* (2005) resolveram a estrutura tridimensional da proteína humana Kv β 2 (PDB 1ZSX), que possui mais de 80% de similaridade com a proteína Kv β 1.3, o que facilitaria um eventual estudo da estrutura cristalográfica da Kv β 1.3.

4. A PROTEÍNA EFHC1

Para se estudar a proteína EFHC1, decidiu-se construir dois clones truncados, buscando aumentar as chances de se obter proteína solúvel. Descreveremos, em seções separadas, as porções N- e C-terminal da proteína EFHC1 humana.

4.1 A PORÇÃO N-TERMINAL DA PROTEÍNA HUMANA EFHC1

4.1.1 Metodologia

Clonagem e subclonagem do gene de interesse

Para a amplificação da região que codifica a porção N-terminal da proteína EFHC1, denominada aqui EFHC1N (AA 78-364, figura 2), utilizou-se cDNA sintetizado a partir de RNA total de hipocampo humano. Os oligos utilizados para a reação em cadeia da polimerase (PCR) foram os seguintes:

OMM22: 5'-ACTTGGTCTCCCATGAAACAAGCCCCACCTGCGG-3' (*Eco3II/NcoI*)

OMM23: 5'-ATCCTCGAGTCATAAATCAGTGATTCCAAAC-3' (*XhoI*)

O tamanho do produto de PCR esperado era de aproximadamente 876 bp, tendo sido extraído do gel de agarose utilizando-se o kit *Wizard SV Gel and PCR Clean-Up System* (Promega). Realizou-se a ligação com o vetor de clonagem pGEMT-easy (Promega) de acordo

com o protocolo sugerido pelo fabricante. Os clones positivos foram confirmados por digestão enzimática ou PCR direto de colônia, e a integridade da sequência foi verificada por sequenciamento automático. Posteriormente, utilizou-se o clone positivo como molde para a amplificação de novo produto de PCR para subclonagem no vetor de expressão pETSUMO (Invitrogen), cuja reação de ligação foi realizada conforme instruções do fabricante. Estavam previstos ainda, caso houvesse necessidade, novas clonagens nos vetores pETNusA e pETMBP. Por isso a presença dos sítios de restrição nos oligos. Metade do volume da reação de ligação foi transformada em *Escherichia coli* DH5 α competente através de choque térmico, de acordo com protocolo otimizado de Sambrook *et al.* (1989). Os clones positivos foram verificados através de digestão enzimática.

Importa salientar que o Comitê de Ética em Pesquisa (CEP) da Faculdade de Ciências Médicas (FCM) aprovou o adendo ao projeto nº 470/2003, que permite a análise e a extração de RNAs e proteínas de espécimes cirúrgicos e de necrópsia para pesquisa.

Teste de expressão

Para o teste de expressão, foram testados dois clones positivos, nomeados 1 e 8. Transformou-se a cepa BL21(DE3) juntamente com o plasmídio pRARE, já que o inserto apresentava códons raramente utilizados em bactéria (anexo 4). A partir de uma colônia isolada em meio de cultura sólido contendo os antibióticos apropriados (canamicina e cloranfenicol), fez-se um pré-inóculo em meio de cultura LB a 37°C, a 200 rpm de rotação por no mínimo 16 h seguidas. Inoculou-se de 1% a 3% do pré-inóculo em 100 mL de meio LB suplementado com canamicina (50 μ g/mL) e cloranfenicol (34 μ g/mL), a 37°C, 200 rpm. Ao atingir 0,6 de A_{600nm} , induziu-se a expressão protéica com IPTG 0,5 mM. Retirou-se uma alíquota no tempo zero de

indução (T0) para ser analisado em gel de poliacrilamida. Baixou-se a temperatura para 30°C e manteve-se a cultura por mais 4 h a 200 rpm. Retirou-se mais uma alíquota neste tempo (T4). Após o que, centrifugou-se as células 6.000 g por 10 min a 4°C, descartando-se o sobrenadante e congelando-se a -20°C.

O tamanho esperado da proteína recombinante era de aproximadamente 47,6 kDa. As amostras congeladas foram ressuspensas em tampão de extração (Tris-HCl 20 mM, pH 7,5, NaCl 150 mM, glicerol 5%, DTT 1 mM, EDTA 1 mM, PSMF 0,5 mM), adicionou-se lisozima (5 µg/mL de extrato) e incubou-se em gelo por no mínimo 20 min. Após esse período, as amostras foram sonicadas 3 vezes durante 10 s a 25% de amplitude em sonicador Sonics VCX-750. As frações solúveis e insolúveis foram separadas por centrifugação a 20.000 g por 10 min a 4°C. A amostra precipitada foi ressuspensa em tampão de extração acrescido de Triton X-100 0,1% (v:v), a fim de solubilizar as proteínas de membrana, e novamente sonicadas por 10 s três vezes, a 25% de amplitude. Essa amostra foi novamente centrifugada, sendo o sobrenadante descartado e a fração insolúvel limpa era novamente ressuspensa em tampão de extração. As amostras solúveis e insolúveis foram analisadas em gel SDS-PAGE 13,5%.

Purificação por IMAC e testes de polimento

As células induzidas e congeladas foram ressuspensas em tampão afinidade (Na₂HPO₄ e NaH₂PO₄ 50 mM, pH 7,2, glicerol 5%, NaCl 300 mM). Incubou-se com lisozima (estoque a 10 mg/mL) por 30 min em gelo. Sonicou-se 10 s com intervalo de 20 s dez vezes. Centrifugou-se por 10 min a 4°C 20.000 g. O sobrenadante foi incubado com 400 µL da resina Ni-NTA (Qiagen) por 1 h em temperatura ambiente. A matriz foi transferida para uma coluna Biospin (BioRad). Lavou-se com 20 volumes de coluna (cv), e a eluição foi por passos: imizadol 5 mM, 25 mM, 50

mM e 250 mM adicionados ao tampão afinidade. Todas as frações foram analisadas em gel de poliacrilamida (SDS-PAGE).

Após a captura inicial da proteína por IMAC, testou-se ainda um segundo passo de purificação, a fim de realizar o polimento da amostra. Dialisou-se a mesma em tampão A IEX (Tris-HCl 20 mM, pH 8,0, NaCl 20 mM, DTT 1 mM) por 16 h a 4°C. Após esse período, a amostra foi centrifugada por 10 min a 20.000 g 4°C. Testou-se duas colunas de troca iônica, HiTrap Q FF e HiTrap DEAE FF, ambas de 1 mL (GE Healthcare). Aplicou-se a amostra nas colunas previamente equilibradas com o mesmo tampão de diálise e a eluição se deu em um gradiente crescente de sal, utilizando o cromatógrafo FPLC. As amostras aplicadas e o eluído foram analisados em gel SDS-PAGE.

Testou-se ainda uma cromatografia de filtração em gel, utilizando a coluna Superdex75 10/300 GL (GE Healthcare), tendo como fase móvel o tampão SAXS (Tris-HCl 20 mM, pH 7,5, NaCl 200 mM, DTT 1 mM). Realizou-se a corrida em 1,5 cv a 0,5 mL/min, no sistema FLPC.

Teste de clivagem da cauda SUMO

A amostra purificada por afinidade a níquel foi quantificada por coeficiente de extinção a 280 nm (ϵ_{280}), pois a porção N-terminal da EFHC1 apresenta quatro resíduos de triptofano. Adicionou-se à solução de proteína purificada a protease Ulp1, produzida em nosso laboratório, numa razão enzima:substrato 1:100 (m/m) ao mesmo tempo em que se realizava a diálise em tampão A IEX por 16 h a 4°C.

Após diálise e clivagem, a amostra foi centrifugada por 10 min a 20.000 g a 4°C, e aplicada numa coluna de Ni-NTA previamente equilibrada com o tampão A IEX. A amostra foi coletada e as proteínas que possivelmente se ligaram por afinidade à coluna (His₆-SUMO e His₆-

Ulp1 e possíveis contaminantes) foram eluídas com imidazol 400 mM. Todas as frações foram analisadas em gel SDS-PAGE. Caso a reação de clivagem fosse bem sucedida e na eventualidade da amostra apresentar ainda contaminantes, poderia ser realizado um segundo passo de purificação por troca iônica.

Análise por Western blot

A fim de se conhecer a identidade da proteína recombinante, realizou-se análise de *western blot*, com anticorpo policlonal anti-Efhc1 de camundongo, gentilmente cedido pelo Prof. Ritsu Kamiya, da Universidade de Tóquio.

Diferentes amostras (extrato total; fração insolúvel; lavagem da IMAC; e eluição de IMAC) foram fracionadas em gel SDS-PAGE 13,5% e transferidas para membrana Hybond-C (GE Healthcare) por 2 h a 37 mA, utilizando aparato semi-seco horizontal da Amersham Biosciences.

Após a transferência, a membrana foi bloqueada em temperatura ambiente por 2 h com TBS (Tris-HCl 20 mM, pH 7,6, NaCl 140 mM), acrescido de leite desnatado desidratado 5% (m/v). Para o anticorpo primário, utilizou-se a diluição 1:350 em TBS leite, incubando-se com a membrana por 16 h em temperatura ambiente, com agitação constante. Lavou-se a membrana duas vezes com TBS 1X e, em seguida, incubou-se o anticorpo secundário conjugado à fosfatase alcalina (Sigma), diluído 1:20.000, por 1 h em agitação constante à temperatura ambiente. As bandas foram detectadas utilizando-se BCIP/NBT (Fermentas) como substrato para a fosfatase.

4.1.2 Resultados e discussão

Teste de expressão

As frações não induzidas (T0, tempo zero de indução) e induzidas (T4, 4 h da indução) foram analisadas em gel SDS-PAGE 13,5% (figura 16). Observou-se a presença de proteína nas frações solúvel e insolúvel, no tempo 4 h de indução, apenas para o clone 1. Em relação ao clone 8, observou-se que há uma banda correspondente ao tamanho esperado da proteína na fração solúvel apenas (tanto em T0 quanto em T4) e não na fração insolúvel com 4 h de indução.

Seguiram-se os testes seguintes com o clone 1.

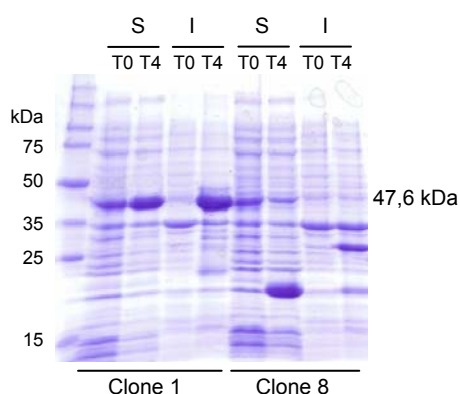


Figura 16: Teste de expressão a 30°C da construção SUMO-EFHC1N. Gel SDS-PAGE 13,5%. Legenda: S, fração solúvel; I, fração insolúvel; T0: tempo de indução zero; T4: tempo de indução 4 h. Marcador de massa molar indicado. O tamanho esperado da proteína era de aproximadamente 47,6 kDa. Observa-se que há proteína solúvel e insolúvel em 4 h de indução para o clone 1 e que não há proteína no tamanho esperado para o clone 8.

Purificação por IMAC e corridas cromatográficas subseqüentes

A proteína de fusão SUMO-EFHC1N foi capturada através de cromatografia de afinidade em níquel imobilizado. Observa-se uma grande quantidade de proteína que não interagiu com a resina (figura 17, canaleta 1), e que a fração eluída da coluna com imidazol 250 mM apresentava alguns contaminantes de baixa massa, mesmo com a extensiva lavagem da resina com 20 cv de tampão afinidade.

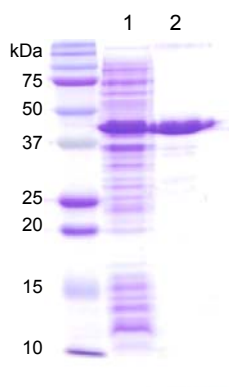


Figura 17: Gel SDS-PAGE 13,5% do perfil de purificação da proteína SUMO-EFHC1N em IMAC. Legenda: 1, amostra que passou diretamente pela resina de Ni-NTA; 2, proteína eluída com imidazol 250 mM. Marcador de massa molar indicado. Observa-se que ainda há muita proteína que não interagiu com a coluna (canaleta 1).

Os testes que se seguiram buscando o polimento da proteína, seja por troca iônica, seja por filtração em gel, falharam em remover as proteínas contaminantes. Seguiu-se com o teste de clivagem da cauda SUMO, mesmo com a proteína apresentando tais contaminantes, pois a proteína clivada poderia necessitar de um passo final de purificação.

Teste de clivagem da cauda SUMO

A protease Ulp1 suporta uma ampla gama de condições de tampão na reação de clivagem da cauda SUMO, como alteração na força iônica (NaCl entre 100 e 300 mM), temperatura (sua temperatura ótima é 30°C, mas suporta reações a 4°C, 16°C ou 21°C), pH (de 5,5 a 9,5) e aditivos, como agentes redutores e desnaturantes, detergentes e outros (Malakhov *et al.*, 2004). A recomendação dos fabricantes desta protease é manter a concentração do imidazol abaixo de 150 mM e a concentração de sal entre 150 mM e 300 mM.

Utilizamos o tampão A IEX para realizar a reação (Tris-HCl 20 mM, pH 8,0, NaCl 20 mM, DTT 1 mM), conjuntamente com a diálise a 4°C no mesmo tampão. Conforme figura 18, infelizmente não obtivemos a clivagem da proteína. Tanto a temperatura como a concentração de sal utilizados estão abaixo do ótimo recomendado pelos fabricantes, pois buscou-se manter as condições ótimas para um passo de cromatografia de troca iônica posterior à clivagem. Nessa mesma condição foi possível a clivagem da construção SUMO-EFHC1C (figura 18, canaleta 5), mostrando que a composição do tampão não impediu o funcionamento da protease Ulp1.

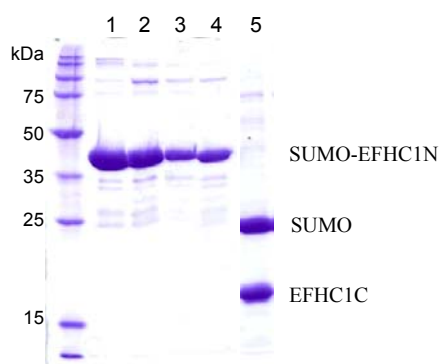


Figura 18: Gel SDS-PAGE 13,5% do teste de clivagem da construção SUMO-EFHC1N. Legenda: 1, controle negativo (SUMO-EFHC1N sem a presença da Ulp1); 2, reação de clivagem; 3, amostra que passou direto pela resina Ni-NTA; 4, amostra eluída com imidazol 400 mM; 5, controle positivo (SUMO-EFHC1C clivada com Ulp1). Marcador de massa molar indicado.

Mais de um teste foi realizado, mas não se obteve a clivagem da cauda SUMO. Observamos que em altas concentrações (acima de 20 mg/mL) a proteína SUMO-EFHC1N tendia a agregar fortemente. Esta condição de agregação pode impedir a protease de se aproximar do sítio de clivagem.

Análise por western blot

As construções SUMO-EFHC1N e SUMO-EFHC1C apresentam tamanhos aparentes muito próximos no gel SDS-PAGE (dados não mostrados). Realizamos uma análise por *western blot* contra Efhc1 a fim de confirmar a identidade das proteínas em estudo. Como esse anticorpo (Ikeda *et al.*, 2005) reconhece a porção C-terminal da proteína Efhc1 de camundongo, no motivo EF-hand, apenas o clone C-terminal seria reconhecido. Além disso, é preciso dizer que as proteínas humana e de camundongo apresentam mais de 80% de aminoácidos idênticos. Isso explica porque o anticorpo reconheceu a porção C-terminal da proteína recombinante humana e, como era de se esperar, não reconheceu a porção N-terminal da mesma (dados não mostrados), confirmando a identidade de ambas.

Adicionalmente, foram iniciadas análises de espectrometria de massas, para determinar se o sítio de clivagem da proteína está intacto, assim como para determinar o local preciso de clivagem da porção C-terminal. Todavia, devido a problemas de preparação das amostras, não se concluíram as análises.

4.2 MANUSCRITO

“The C-terminal half of juvenile myoclonic epilepsy human EFHC1: preliminary structural analysis and dimerization evidence”

Marcelo Jun Murai, André Henrique Zamboni, Fábio Frangiotti Conte, Ricardo Aparício and
Ischia Lopes-Cendes

Submetido para a Revista

Archives of Biochemistry and Biophysics

**The C-terminal half of juvenile myoclonic epilepsy human EFHC1:
preliminary structural analysis and dimerization evidence**

Marcelo Jun Murai¹, André Henrique Zamboni¹, Fábio Frangiotti Conte¹, Ricardo Aparício² and
Ischia Lopes-Cendes^{1,¶}

¹Department of Medical Genetics, State University of Campinas (Unicamp), Campinas, SP, Brazil; ²Chemistry Institute, State University of Campinas (Unicamp), Campinas, SP, Brazil.

¶ Corresponding author:

Ischia Lopes-Cendes, M.D., Ph.D.

Department of Medical Genetics

State University of Campinas (Unicamp)

Rua Tessália Vieira de Camargo, 126

Campinas, SP, Brazil

CEP 13083-970

Phone: +55 19 35218907

Fax: +55 19 35218909

e-mail: icendes@unicamp.br

Abstract

Human EFHC1 is a member of the EF-hand superfamily of Ca^{2+} -binding proteins that possess three DM10 domains of unclear function. Point mutations in the *EFHC1* gene are related to juvenile myoclonic epilepsy, a fairly common idiopathic generalized epilepsy. Here we report the first structural characterization of the EFHC1 C-terminus (residues 403 to 640; named EFHC1C), comprising the last DM10 domain and the EF-hand motif. Protein was expressed as a fusion tag in bacteria, isolated by affinity chromatography, cleaved and purified by ion exchange chromatography. Circular dichroism revealed that the secondary structure of EFHC1C is composed mainly by α -helices and some β -strand contributions. Thermal unfolding showed a melting temperature of 54.5 °C. Interestingly, EFHC1C may form a dimer by disulfide bonding. Our results demonstrated that properly folded EFHC1C can be produced in high amount and quality by using bacterial expression system for further structural and functional studies.

Keywords: EFHC1; juvenile myoclonic epilepsy; DM10 domain; EF-hand motif.

Introduction

Different *EFHC1* gene mutations occur in juvenile myoclonic epilepsy (JME) [1], the most common form of idiopathic generalized epilepsy (IGE) [2]. The gene encodes a non-ion channel protein of 640 amino acids with three DM10 domains of unknown function and a putative EF-hand Ca^{2+} -binding motif at the C-terminal portion [3].

Although the existence of very compelling evidence for the involvement of EFHC1 in JME [2], its functions are much less clear. Suzuki *et al.* [1] showed that the EFHC1 protein localizes in different brain regions in the dendrites and soma of neuronal cells and it is associated to a regulatory apoptotic system with an R-type voltage-dependent Ca^{2+} channel $\text{Ca}_v2.3$. The overexpression of the mutated forms decreased cell death in primary neuronal cultures, probably mediated by a reduction of Ca^{2+} influx through the channel. According to the authors, the EFHC1 apoptotic activity may be related to elimination of certain neurons during development of the central nervous system. Ikeda *et al.* [4] proposed another functional role of EFHC1 based on sequence similarity (40% identity) and domain composition suggesting that EFHC1 is closely related to the flagella axoneme component Rib72. Expression and immunofluorescence data showed that it is located in the cilia of tracheal epithelia, sperm flagella and within motile cilia of others tissues, including brain, but in a less extended way [4]. In addition, there is experimental evidence showing that EFHC1 associates to the mitotic apparatus and centrosome, suggesting a possible role during cell division [5].

In this paper we report the first structural data on human EFHC1. The purified C-terminus of the recombinant protein (named EFHC1C), containing the last DM10 domain and the EF-hand motif, was produced in bacteria. Circular dichroism spectroscopy indicates that α -helices is the main secondary structure of the EFHC1C. The protein has a high thermal stability

and unfolds in a sharp transition centered at 54.5 °C. In addition, a dimer conformation was also observed in some of the experimental conditions.

Materials and methods

Plasmid construction

Total RNA was extracted from human hippocampus using TRIzol reagent (Invitrogen), according to manufacturer instructions. cDNA was generated by reverse transcriptase reaction using ImProm-II Reverse Transcription System (Promega). The C-terminus of EFHC1, EFHC1C, comprising amino acids 403-640 was PCR amplified using forward primer OMM24 (5'-TTTGGGTCTCCCATGAAAGCTCCAAAAAAGACG-3') and reverse primer OMM25 (5'-ATCCTCGAGTCAGTTTGAGAAAGCACGA-3'). The resulting amplicon was cloned into pGEM-T easy vector system (Promega), creating plasmid pGEM-EFHC1C that was used as template for subcloning into pET SUMO (Champion pET SUMO Protein Expression System, Invitrogen), generating pETSUMO-EFHC1C. Both constructs were verified by restriction enzyme digestion and automatic DNA sequencing. Normal hippocampal samples were obtained from autopsy material. This study was approved by Research Ethics Committee of our Institution.

Protein expression and purification

Protein was expressed using *Escherichia coli* strain BL21(DE3) *slyD*⁻, kindly provided by Dr. Ryland Young (Texas A&M University). Competent cells carrying the plasmid pRARE (Novagen) were transformed with recombinant plasmid. Single bacterial colonies were cultured overnight at 37 °C in LB broth added to 50 µg/mL kanamycin and 34 µg/mL chloramphenicol. This pre-innoculum was transferred to 2 L of LB broth supplemented with kanamycin and chloramphenicol at the same concentration. When cultures reached an A_{600} of 0.6, expression was induced by adding 0.5 mM IPTG (isopropyl β-D-thiogalactopyranoside), 30 °C for 4 h and

shaking at 200 rpm. Cells were then harvested by centrifugation at 5,000 *g* for 10 min at 4 °C and stored at −20 °C. Induced bacterial cells were resuspended in affinity chromatography buffer (50 mM sodium phosphate buffer, pH 7.2, 300 mM NaCl, 5% glycerol, 2,5 mM imidazole) containing PMSF (phenylmethylsulphonyl fluoride) and lysozyme to concentrations of 0.5 mM and 1.0 mg/mL, respectively. The suspension was sonicated, clarified by centrifugation at 20,000 *g* and the supernatant was used for protein purification. The extract from 1 L of LB cultures was incubated with 3 mL of Ni-NTA Superflow resin (Qiagen) for 1 h at 25 °C; transferred to a spin column, washed with 20 column volumes (*cv*) of affinity buffer and eluted with increasing imidazole concentrations (up to 250 mM) in the same buffer.

Cleavage and separation of fusion tag from human EFHC1C

The affinity purified protein was incubated with SUMO protease-1 (Life Sensors) in a 1:100 (w/w) enzyme-to-substrate ratio; and dialyzed overnight at 4 °C against buffer A (20 mM Tris-HCl, pH 7.5, 20 mM NaCl, 1 mM DTT). The cleavage reaction was passed by gravity through 3 mL of Ni-NTA column pre-equilibrated with buffer A. Flow-through fraction was collected and applied onto a 1 mL HiTrap Q FF column (Amersham Biosciences) pre-equilibrated with buffer A. The column was washed with 2 *cv* using ÄKTA-FPLC system (Amersham Biosciences) and proteins were eluted with 15 *cv* linear gradient (0 to 30%) of buffer B (20 mM Tris-HCl, pH 7.5, 1 M NaCl, 1 mM DTT). Fractions enriched with the protein were pooled and concentrated using Amicon Ultra-4 centrifugal filter device (Millipore). Purification efficiency was analyzed by SDS–PAGE [6].

Protein quantitation

Protein concentration was estimated measuring the absorbance at 205 nm (A_{205}), according to method described by Scopes [7], and at 280 nm (A_{280}) considering extinction coefficient calculated from the composition of the protein using Protparam tool (<http://www.expasy.org/tools/protparam.html>).

Far-UV circular dichroism (CD) spectroscopy and thermal stability studies

CD spectra were acquired on a Jasco J-810 spectropolarimeter (Jasco) coupled to a Peltier Jasco PFD-425 system for temperature control. Protein samples were prepared at 2 μ M in 10 mM sodium phosphate buffer pH 7.4. The spectra were collected at 20 °C using 0.2 cm path length cell over the range of 190–260 nm, and a 1-nm bandwidth with step size of 0.5 nm and speed of 50 nm/min. For each measurement, the mean values of fifteen spectra were taken to improve signal-to-noise ratio. The ellipticity was expressed as the mean residue ellipticity [θ] ($\text{deg.cm}^2.\text{dmol}^{-1}$). For thermal denaturation studies, CD measurements were made from 20 °C to 90 °C (by using scan rates of 60 °C/h) following the loss of secondary structure at 222 nm using 0.5 °C intervals and spectra were collected at 5 °C intervals over the range λ 190–260 nm. Midpoint transition temperature was calculated as the center of the Gaussian fit of the first derivative of the denaturation curve. Prediction of secondary structure elements was performed using PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred/psiform.html>), PHD (<http://www.predictprotein.org/>) and DICHROWEB (<http://public-1.cryst.bbk.ac.uk/cdweb>). Thermal stability study was carried out in thermocycler Mastercycler (Eppendorf). 60 μ g of sample for each temperature were incubated 5 min over the range of 40 °C to 100 °C, with 10 °C intervals. Precipitated fractions were separated from soluble fractions by centrifugation and both were analyzed by SDS-PAGE.

Western blot analysis

Protein samples were fractionated by SDS–PAGE and transferred onto a nitrocellulose filter using a horizontal semi-dry electroblot apparatus (Amersham Biosciences). The filter was incubated for 2 h at 25 °C in blocking solution (3% nonfat dry milk powder in 25 mM Tris-buffered saline, pH 7.4). The membrane was processed through sequential incubations with primary antibody for 16 h, and then with an alkaline phosphatase-conjugated, affinity-purified goat anti-rabbit IgG antibody (Sigma-Aldrich). Immunoreacted bands were detected using BCIP/NBT (5-bromo-4-chloro-3-indolylphosphate/nitroblue tetrazolium) substrate (Fermentas). Polyclonal anti-Efhc1 antibody was a gift from Dr. Ritsu Kamiya (University of Tokyo).

Results and Discussion

Expression and purification of human EFHC1C fusion protein

The EFHC1 protein belongs to a poorly characterized class of DM10 domain-containing proteins with an unknown function. The DM10 domain contains ~105 residues and has been identified in only one copy in the N-terminal portion of nucleoside-diphosphate kinase (NDK) of human nm23-H7 [8, 9] and *Chlamydomonas reinhardtii* flagellar axoneme associated protein p40 [10]. In mouse Efhc1 and *Chlamydomonas* Rib72 there are three repeats within this domain [4, 10]. Proteins that have DM10 domains interact with doublet microtubules [4, 11], raising a possible function in axonemal assembly or regulating NDKs [12]. EF-hand motif is the largest utilized Ca^{2+} -binding motif found in proteins [13] and is characterized by a helix-loop-helix structure [3, 14]. The EF-hand superfamily regulates many aspects of cell function, like Ca^{2+} buffering in the cytosol, cell proliferation and signal transduction [15, 16]. In the nervous central system, calmodulin and the related EF-hand containing neuronal Ca^{2+} sensor proteins have many important roles in neuronal signalling [17, 18]. The basic functional unit consists in a pair of EF-hand motifs [14], but in the case of human EFHC1 it is not clear if it is present in pairs or not.

In order to facilitate purification from *E. coli*, the C-terminal half of EFHC1 (EFHC1C) was expressed as a fusion protein with an N-terminal SUMO tag. SUMO has been described as a solubility enhancer, favoring proper folding of the passenger protein [19]. Recombinant vector was cotransformed with pRARE into BL21(DE3) *slyD*⁻, because EFHC1C sequence presents codons that are rare in bacteria. The expression of SUMO-EFHC1C was induced by 0.5 mM IPTG at 30 °C by four hours. The recombinant protein yielded ~70 mg of protein per liter of culture after a single affinity step (Fig. 1A, lane 2).

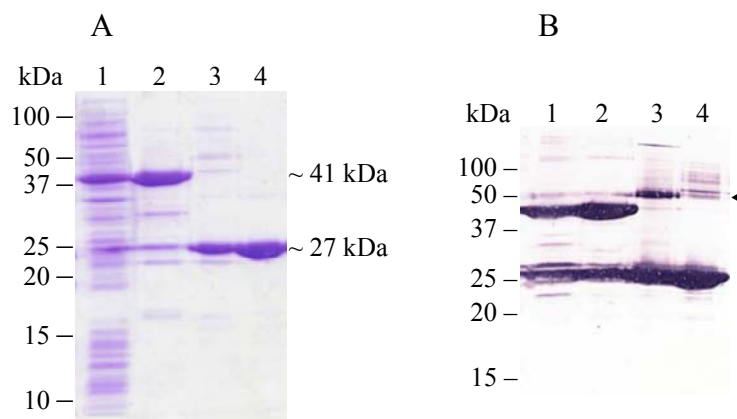


Fig. 1

Figure 1 (A) SDS-PAGE analysis of purification steps. See legend below. (B) Western blot analysis of the separated proteins in (A) using polyclonal mouse anti-Efhc1 antibody [4]. Lane 1, soluble fraction extracted from bacterial lysis; lane 2, fraction collected from affinity chromatography during the elution step at imidazole concentration of ~100 mM; lane 3, cleaved EFHC1C that passed through affinity chromatography column to remove tag; lane 4, final purified protein eluted from ion-exchange chromatography. Molecular weight markers are indicated. Filled arrow head indicates a probable EFHC1C dimer.

The crude cell lysate was incubated in a batch preparation procedure with Ni-NTA resin for 1 h at room temperature, followed by washing the packed resin with affinity chromatography buffer. The human EFHC1C protein was eluted from the Ni-NTA column by a gradient of 5 mM to 250 mM imidazole and exhibited a satisfactory purity level for cleavage and remove of tag (Fig. 1A, lane 2).

Cleavage of SUMO-EFHC1C fusion protein and EFHC1C purification

Eluted fractions from first affinity purification step were successfully cleaved with SUMO protease-1. Both SUMO fusion tag and SUMO protease-1 contain an N-terminal polyhistidine tag that facilitates capture with Ni-NTA resin. EFHC1C protein passed through the column and was collected in the flow-through fraction (Fig. 1A, lane 3). Final purification steps were performed using ion exchange Q FF chromatography column, separating trace contaminant proteins and resulting in a highly purified target protein, as judged by SDS-PAGE (Fig. 1A, lane 4). The final amount of purified EFHC1C protein was ~2.7 mg/L of initial bacterial broth. Dynamic light scattering analysis showed that under room temperature cleaved EFHC1C was monodisperse (data not shown).

CD spectroscopy and secondary structure

Far-UV CD spectrum at 20 °C of cleaved EFHC1C shown in Fig. 2A exhibits two minima centered around 208 nm and 225 nm, suggesting the presence of helical structure along with other secondary structure elements. However, the CD signal at 225 nm is not so pronounced as the 208 nm signal, probably due to the presence of Ca^{2+} bound to the protein, as reported.

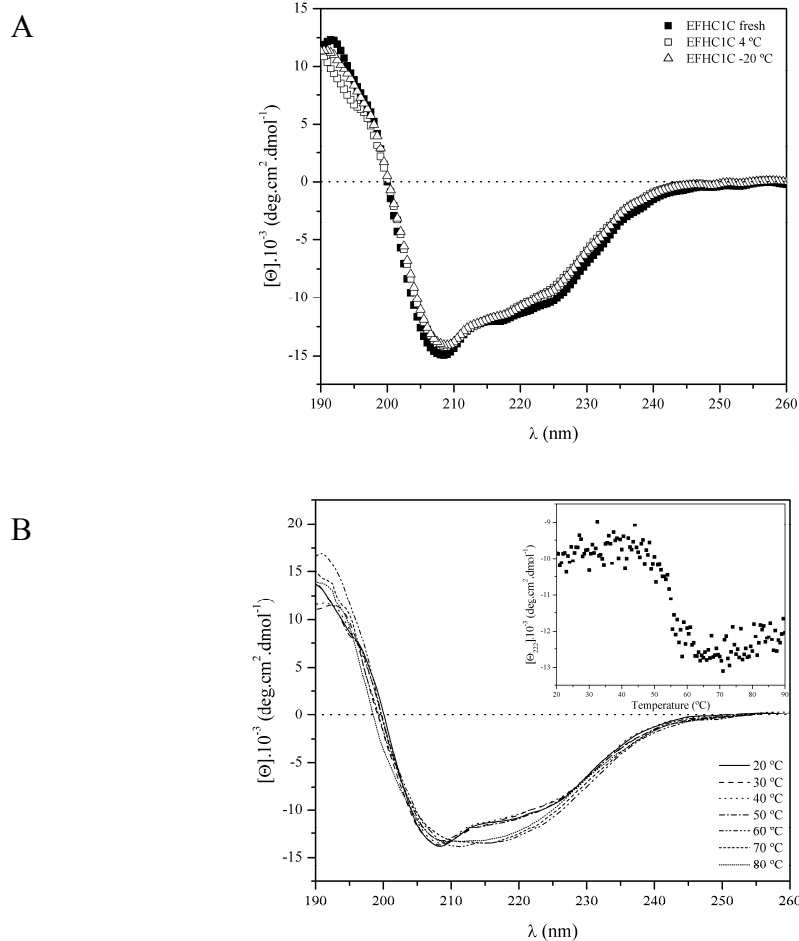


Fig. 2

Figure 2 (A) Far-UV CD spectra of cleaved purified EFHC1C. Three different conditions were measured: freshly prepared protein, 19-days protein stored at 4 °C and freeze and thaw protein. (B) Thermal unfolding analysis in absence of reducing agent. For clarity, just some of the recorded spectra are shown. Unfolding was also followed by measuring the ellipticity at 222 nm using 0.5 °C intervals in the same temperature range (inset).

elsewhere [20, 21]. It is important to mention that throughout protein characterization no treatment with metal chelators was used, such as EDTA. EFHC1C does not contain any tryptophan, only the aromatic residues phenylalanine (13) and tyrosine (13). We also observed that both frozen protein at $-20\text{ }^{\circ}\text{C}$ and stored protein at $4\text{ }^{\circ}\text{C}$ for 19 days showed just modest decrease in secondary structure elements (Fig. 2A), evidencing a stable conformation.

In order to estimate the content of different types of secondary structures in EFHC1C, the far-UV CD spectrum of the protein was analyzed using the routines available at DICHROWEB. The values obtained for the secondary structures are: α -helix (42.1%) and β -sheet (15.2%), is in agreement with secondary structure predictions by PSIPRED and PHD (Fig. 3). The high α -helical content is characteristic of the EF-hand family of proteins [22]. It is predicted that DM10 domain contributes totally with β -sheet elements (see Fig. 3) [12] but unfortunately there is no CD information published on this domain, so it is not possible to correlate our data with other DM10 domain proteins.

Thermal stability studies

Thermal unfolding of EFHC1C was investigated by far-UV CD. Spectra collected at $5\text{ }^{\circ}\text{C}$ intervals over the range of $20\text{ }^{\circ}\text{C}$ to $90\text{ }^{\circ}\text{C}$ at a scan rate of $60\text{ }^{\circ}\text{C/h}$ are shown in Fig. 2B. Only moderate changes in the spectral features are seen in the spectra recorded at $20\text{ }^{\circ}\text{C}$ until $50\text{ }^{\circ}\text{C}$ but significantly larger changes were observed in the spectra recorded at $60\text{ }^{\circ}\text{C}$ until $90\text{ }^{\circ}\text{C}$, showing the occurrence of a thermally-induced unfolding phase transition between $50\text{ }^{\circ}\text{C}$ and $60\text{ }^{\circ}\text{C}$.

Loss of secondary structure were followed at 222 nm in the same temperature range and showed that EFHC1C possess midpoint transition temperature of $54.5\text{ }^{\circ}\text{C}$ (Fig. 2B, inset). Refolding was found to be irreversible (data not shown), probably due to an aggregation process,

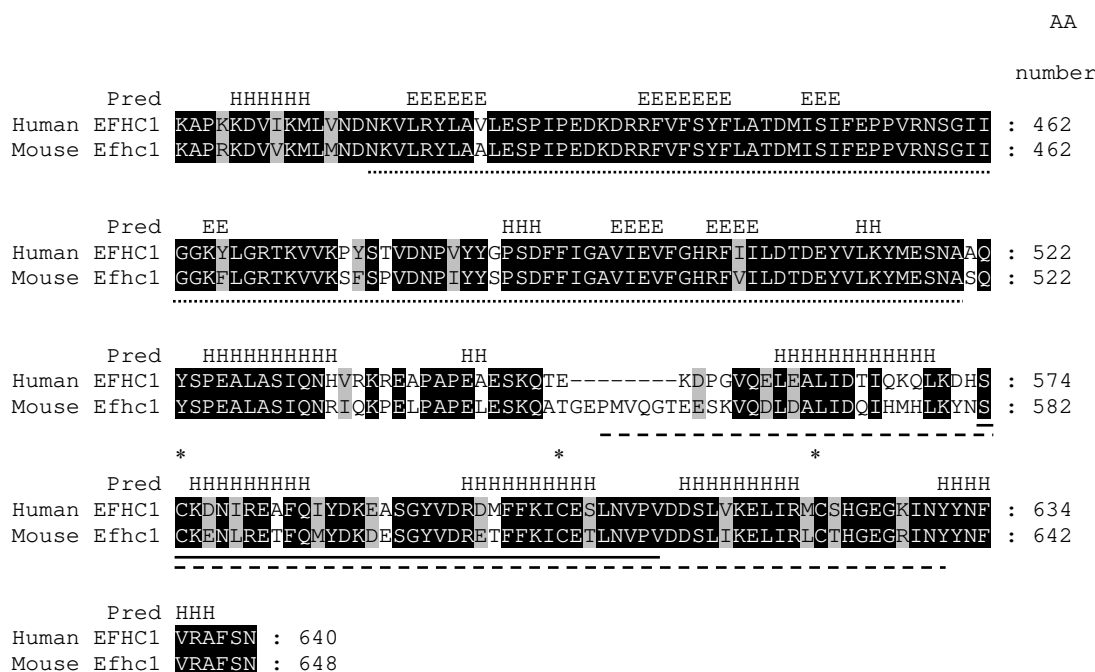


Figure 3 Amino acid sequence comparison of human EFHC1 (amino acids 403–640; GenBank accession no. **NM018100**) and mouse Efhc1 (amino acids 403–648; GenBank accession no. **XM990358**). The sequence alignment was carried out using CLUSTALW. Letters shaded in black indicate amino acids that are identical. Letters shaded in gray indicate similar amino acids. Secondary structure elements (E = β -sheet, H = α -helix) are indicated above the predict structure deduced from the human EFHC1 sequence (by PSIPRED and PHD servers). Dotted line, DM10 domain; solid line, EF-hand motif; dashed line, mouse Efhc1 sequence recognized by polyclonal antibody [4]. DM10 domain and EF-hand motif were predicted by PROSITE (<http://expasy.org/prosite/>) and SMART (<http://smart.embl-heidelberg.de/>). Cysteine residues are marked with an asterisk.

as a progressive increase in dynode voltage was observed [23]. In addition, a clear 50 kDa protein band appeared in SDS-PAGE analysis of CD thermally denatured protein (Fig. 4C, lane 2). The thermal transition curve in the presence or absence of reducing agent (1 mM DTT) was virtually the same and no turbidity was visualized in the unfolding process (data not shown); although the presence of reducing agent inhibits oligomer formation (see below and Fig. 4B).

Additional denaturation study was performed, checking soluble and insoluble fractions of protein aliquots incubated in a thermocycler (Fig. 4A). Thermal stability was tested over the temperature range of 40 °C to 100 °C. Most of EFHC1C denatured protein remains soluble until 80 °C, and then, a stronger precipitation occurs above 90 °C. Strikingly, it is observed the presence of an oligomer of ~50 kDa, another evidence of the aggregation process seen by CD denaturation (Fig. 4A). Taken together, these thermal stability studies show that EFHC1C is a rather stable protein.

Western blot

Western blot analysis confirmed that the human EFHC1C was identified by an anti-Efhc1 antibody [4] (Fig. 1B). This polyclonal antibody was raised against the EF-hand motif sequence of mouse Efhc1 (Fig. 3). Mouse Efhc1 and human EFHC1 share 76% amino acid identity and 86% similarity, which explains why the polyclonal mouse antibody recognized the human protein, although the mouse protein presents some extra amino acids. Peptide mass fingerprinting was also performed by mass spectrometry and correctly identified our clone construction (data not shown). Unexpectedly, intracellular cleavage of SUMO fusion tag was observed (Fig. 1B, lane 1 and 2), probably in the same cleavage site of SUMO protease-1, as judged by relative migration position on SDS-PAGE, indicating that protease inhibitor PMSF was not sufficient to

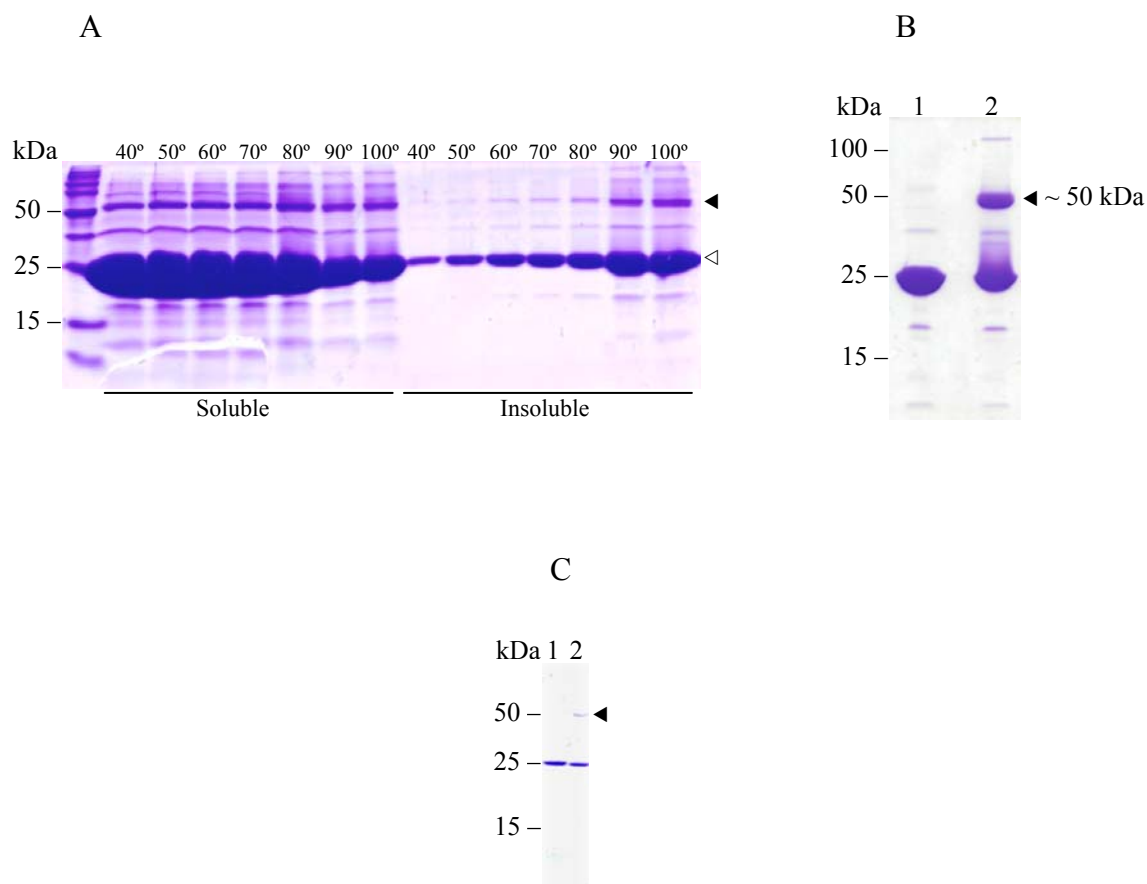


Fig. 4

Figure 4 (A) Thermal stability study analyzed by SDS-PAGE. 60 μ g of protein were incubated for each temperature ($^{\circ}$ C) indicated for 5 min. (B) Dimerization evidence of EFHC1C. Lane 1, EFHC1C in the presence of reducing agent; lane 2, EFHC1C in the absence of reducing agent. (C) EFHC1C sample measured by CD. Lane 1, 20 $^{\circ}$ C measured protein; lane 2, thermally denatured protein. Empty arrow head, EFHC1C monomer; Filled arrow head, indicating an EFHC1C dimer of $\sim$ 50 kDa. Molecular weight markers are indicated.

block these unspecific cleavage. Again it is possible to visualize the probable EFHC1C dimer (Fig. 1B, lane 3), even though the polyacrilamide gel utilized was reduced and denaturing one.

Evidence of EFHC1C dimerization

It is not known yet if the putative C-terminal EF-hand motif of EFHC1 occurs in adjacent pairs or not, but secondary structure prediction (Fig. 3) shows two sets of two α -helices in tandem, and amino acid variation in the roughly 30 positions of the consensus sequence [24] indicate that it is probably a pair of motifs. In addition, *Chlamydomonas* Rib72 presents as a pair of EF-hands [4, 10].

Many EF-hand proteins naturally occur as dimers [25]. Throughout this work we were able to detect that EFHC1C adopts a probable dimeric conformation of ~50 kDa. Thermally unfolded protein by Peltier device of spectropolarimeter was analyzed by SDS-PAGE (Fig. 4C, lane 2). The same sample was also subjected to decreasing temperature until starting condition, but the dimer observed was not broken. Another thermal denaturing experiment until boiling point was done by thermocycler and again it was observed the presence of the dimer (Fig. 4A). Even anti-Efhc1 polyclonal antibody raised against mouse protein was capable of detecting the human EFHC1C dimer (Fig. 1B, lane 3). Interestingly, the presence or absence of reducing agent in the protein solvent alters the oligomerization state of EFHC1C, passing from monomeric to dimeric conformation (Fig. 4B), probably by forming a disulfide bridge, as EFHC1C presents three cysteine residues (Fig. 3). A similar effect was also observed by the group of Hidaka, some years ago, with another member of the EF-hand superfamily [26]. Nonetheless, Ca^{2+} binding induces protein dimerization and promote folding stability of the EF-hand [21, 27], mediated by hydrophobic contacts through individual helices [25].

Further experiments must be conducted in order to determine if the dimeric state is a consequence of EF-hand pairing [13, 25] or a DM10 domain aggregation characteristic, as human EFHC1 N-terminal half, consisting of two DM10 domains in tandem, has a strong tendency of aggregation (unpublished data). Evidently our clone encodes just half of the protein; however, our results have shown that this could not be simply an artifact. Whether dimerization is a physiological conformation remains to be elucidated its homodimeric or heterodimeric nature.

Conclusions

Most of the genes responsible for IGE encode ion channel proteins [28]. *EFHC1* gene is one in a few that is a non ion channel protein related to epilepsy. We reported here original structural data about human EFHC1 C-terminus (referred as EFHC1C), containing the last DM10 domain and the EF-hand motif. The target protein was expressed mainly in the soluble form and purification protocol including tag removal was established. Final purified protein presented high stability; CD spectroscopic showed a high α -helical content, despite some β -strand also present and melting temperature was determined. We clearly showed evidence for dimerization that must be further confirmed. Finally, crystallization trials are in progress in our lab in order to determine the three-dimensional structure of EFHC1C, shedding light onto DM10 domain function in JME.

Acknowledgments

The authors thank Dr. Fábio Gozzo (IQ/Unicamp and LNLS) and Luiz Santos (LNLS) for mass spectrometry analysis; Tatiany Marques for technical support. The use of the CD facility at Brazilian Synchrotron Light Laboratory (LNLS) is acknowledged. This work was supported by grants from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP: 04/03600-0 I.L.-C. and 05/03234-6 R.A.). M.J.M. is recipient of fellowship from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). I.L.-C. is a recipient of a research fellowship from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

References

- [1] T. Suzuki, A.V. Delgado-Escueta, K. Aguan, M.E. Alonso, J. Shi, Y. Hara, M. Nishida, T. Numata, M.T. Medina, T. Takeuchi, R. Morita, D. Bai, S. Ganesh, Y. Sugimoto, J. Inazawa, J.N. Bailey, A. Ochoa, A. Jara-Prado, A. Rasmussen, J. Ramos-Peek, S. Cordova, F. Rubio-Donnadieu, Y. Inoue, M. Osawa, S. Kaneko, H. Oguni, Y. Mori, K. Yamakawa, Mutations in EFHC1 cause juvenile myoclonic epilepsy, *Nat. Genet.* 36 (2004) 842–849.
- [2] M. Gardiner, Genetics of idiopathic generalized epilepsies, *Epilepsia* 46 Suppl 9 (2005) 15–20.
- [3] Z. Grabarek, Structural basis for diversity of the EF-hand calcium-binding proteins, *J. Mol. Biol.* 359 (2006) 509–525.
- [4] T. Ikeda, K. Ikeda, M. Enomoto, M.K. Park, M. Hirono, R. Kamiya, The mouse ortholog of EFHC1 implicated in juvenile myoclonic epilepsy is an axonemal protein widely conserved among organisms with motile cilia and flagella, *FEBS Lett.* 579 (2005) 819–822.

- [5] L.de Nijs, B. Lakayea, B. Coumansa, C. Léona, T. Ikeda, A.V. Delgado-Escueta, T. Grisar, G. Chanasa, EFHC1, a protein mutated in juvenile myoclonic epilepsy, associates with the mitotic spindle through its N-terminus, *Exp. Cell. Res.* 312 (2006) 2872–2879.
- [6] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature* 227 (1970) 680–685.
- [7] R.K. Scopes, Measurement of protein by spectrophotometry at 205 nm, *Anal. Biochem.* 59 (1974) 277–282.
- [8] R.S. Patel-King, O. Gorbatyuk, S. Takebe, S.M. King, Flagellar radial spokes contain a Ca^{2+} -stimulated nucleoside diphosphate kinase, *Mol. Biol. Cell* 15 (2004) 3891–3902.
- [9] A. Munier, C. Serres, M. Kann, M. Boissan, C. Lesaffre, J. Capeau, J. Fouquet, M. Lacombe, Nm23/NDP kinases in human male germ cells: role in spermiogenesis and sperm motility?, *Exp. Cell Res.* 289 (2003) 295–306.
- [10] R.S. Patel-King, S.E. Benashski, S.M. King, A bipartite Ca^{2+} -regulated nucleoside-diphosphate kinase system within the *Chlamydomonas* flagellum, The regulatory subunit p72, *J. Biol. Chem.* 277 (2002) 34271–34279.
- [11] K. Ikeda, J.A. Brown, T. Yagi, J.M. Norrander, M. Hirono, E. Eccleston, R. Kamiya, R.W. Linck, Rib72, a conserved protein associated with the ribbon compartment of flagellar A-microtubules and potentially involved in the linkages between outer doublet microtubules, *J. Biol. Chem.* 278 (2003) 7725–7734.
- [12] S.M. King, Axonemal protofilament ribbons, DM10 domains, and the link to juvenile myoclonic epilepsy, *Cell Motil. Cytoskeleton* 63 (2006) 245–253.
- [13] A. Lewit-Bentley, S. Réty, EF-hand calcium-binding proteins, *Curr. Opin. Struct. Biol.* 10 (2000) 637–643.

- [14] S. Bhattacharya, C.G. Bunick, W.J. Chazin, Target selectivity in EF-hand calcium binding proteins, *Biochim. Biophys. Acta* 1742 (2004) 69–79.
- [15] N.J. Skelton, J. Kördel, M. Akke, S. Forsén, W.J. Chazin, Signal transduction versus buffering activity in Ca^{2+} -binding proteins, *Nat. Struct. Biol.* 1 (1994) 239–245.
- [16] M.J. Berridge, P. Lipp, M.D. Bootman, The versatility and universality of calcium signalling, *Nature Rev. Mol. Cell. Biol.* 1 (2000) 11–21.
- [17] I.B. Levitan, It is calmodulin after all! Mediator of the calcium modulation of multiple ion channels, *Neuron* 22 (1999) 645–648.
- [18] R.D. Burgoyne, D.W. O’Callaghan, B. Hasdemir, L.P. Haynes, A.V. Tepikin, Neuronal calcium sensor proteins: multitalented regulators of neuronal function, *Trends Neurosci.* 27 (2004) 203–209.
- [19] T.R. Butt, S.C. Edavettal, J.P. Hall, M.R. Mattern, SUMO fusion technology for difficult-to-express proteins, *Protein Expr. Purif.* 43 (2005) 1–9.
- [20] K. Julenius, J. Robblee, E. Thulin, B.E. Finn, R. Fairman, S. Linse, Coupling of ligand binding and dimerization of helix-loop-helix peptides: spectroscopic and sedimentation analyses of calbindin D9k EF-hands, *Proteins* 47 (2002) 323–333.
- [21] J. Wójcik, J. Góral, K. Pawlowski, A. Bierzynski, Isolated calcium-binding loops of EF-hand proteins can dimerize to form a native-like structure, *Biochemistry* 36 (1997) 680–687.
- [22] M.J. Hunter, W.J. Chazin, High level expression and dimer characterization of the S100 EF-hand proteins, migration inhibitory factor-related proteins 8 and 14, *J. Biol. Chem.* 273 (1998) 12427–12435.
- [23] S. Benjwal, S. Verma, K.-H. Röhm, O. Gursky, Monitoring protein aggregation during thermal unfolding in circular dichroism experiments, *Protein Sci.* 15 (2006) 635–639.

- [24] T. Cedervall, I. André, C. Selah, J.P. Robblee, P.C. Krecioch, R. Fairman, S. Linse, K.S. Akerfeldt, Calbindin D28k EF-hand ligand binding and oligomerization: four high-affinity sites—three modes of action, *Biochemistry* 44 (2005) 13522–13532.
- [25] R. Donato, Functional roles of S100 proteins, calcium-binding proteins of the EF-hand type, *Biochim. Biophys. Acta*, 1450 (1999) 191–231.
- [26] H. Todoroki, R. Kobayashi, M. Watanabe, H. Minami, H. Hidaka, Purification, characterization, and partial sequence analysis of a newly identified EF-hand type 13-kDa Ca^{2+} -binding protein from smooth muscle and non-muscle tissue, *J. Biol. Chem.* 266 (1991) 18668–18673.
- [27] J.N. Wingard, J. Chan, I. Bosanac, F. Haeseleer, K. Palczewski, M. Ikura, J.B. Ames, Structural analysis of Mg^{2+} and Ca^{2+} binding to CaBP1, a neuron-specific regulator of calcium channels, *J. Biol. Chem.* 280 (2005) 37461–37470.
- [28] J. Turnbull, H. Lohi, J.A. Kearney, G.A. Rouleau, A.V. Delgado-Escueta, M.H. Meisler, P. Cossette, B.A. Minassian, Sacred disease secrets revealed: the genetics of human epilepsy, *Hum. Mol. Genet.* 14 (2005) 2491–2500.

4.3 RESULTADOS COMPLEMENTARES

A fim de se verificar não apenas o estado de oligomerização da porção C-terminal da proteína EFHC1, mas se há de fato ligação a cálcio pelo motivo EF-hand, foram realizados experimentos de dicroísmo circular (CD). Após clivagem da cauda SUMO e purificação por troca iônica, a proteína clivada EFHC1C foi dialisada em três diferentes tampões:

- 1) Na_2HPO_4 e NaH_2PO_4 10 mM, pH 7,2;
- 2) Na_2HPO_4 e NaH_2PO_4 10 mM, pH 7,2, DTT 1 mM;
- 3) Na_2HPO_4 e NaH_2PO_4 10 mM, pH 7,2, EDTA 1 mM.

Após a diálise, a concentração das amostras foi estimada pelo método descrito por Scopes (1974) e considerando-se também o coeficiente de extinção (ϵ_{280}) calculado pelo Protparam.

Realizamos, ainda, uma curva de concentração com CaCl_2 com a finalidade de se determinar até que concentração de cálcio a proteína poderia se ligar, mantendo a concentração da EFHC1C constante.

Todavia, tivemos problemas com o branco da amostra contendo EDTA. À medida que acrescentávamos cloreto de cálcio, a solução se tornava particulada (CaCl_2 acima de 1,5 mM). Evidentemente, essa suspensão particulada impossibilitou a comparação da proteína com e sem o agente redutor.

Apesar deste problema, pudemos observar que a porção C-terminal da proteína de fato se liga a cálcio. Somente a partir de 2,5 mM de CaCl_2 a solução protéica passou a ficar igualmente particulada, enquanto a solução “branco” apresentava-se particulada a partir de 1,5 mM de cálcio. Apesar desta constatação, os dados coletados de CD não são de todo confiáveis,

necessitando de mais controles; além disso, temos a questão da solução “branco” se apresentar particulada com a presença do CaCl_2 , problema este que precisa ser contornado.

Realizamos ainda experimentos com SAXS, das amostras SUMO-EFHC1N e EFHC1C. Obtivemos as curvas $I(q) \times 2\theta$ (intensidade de espalhamento em função do ângulo de espalhamento) e a função de distribuição de distâncias $p(r)$ (figura 19).

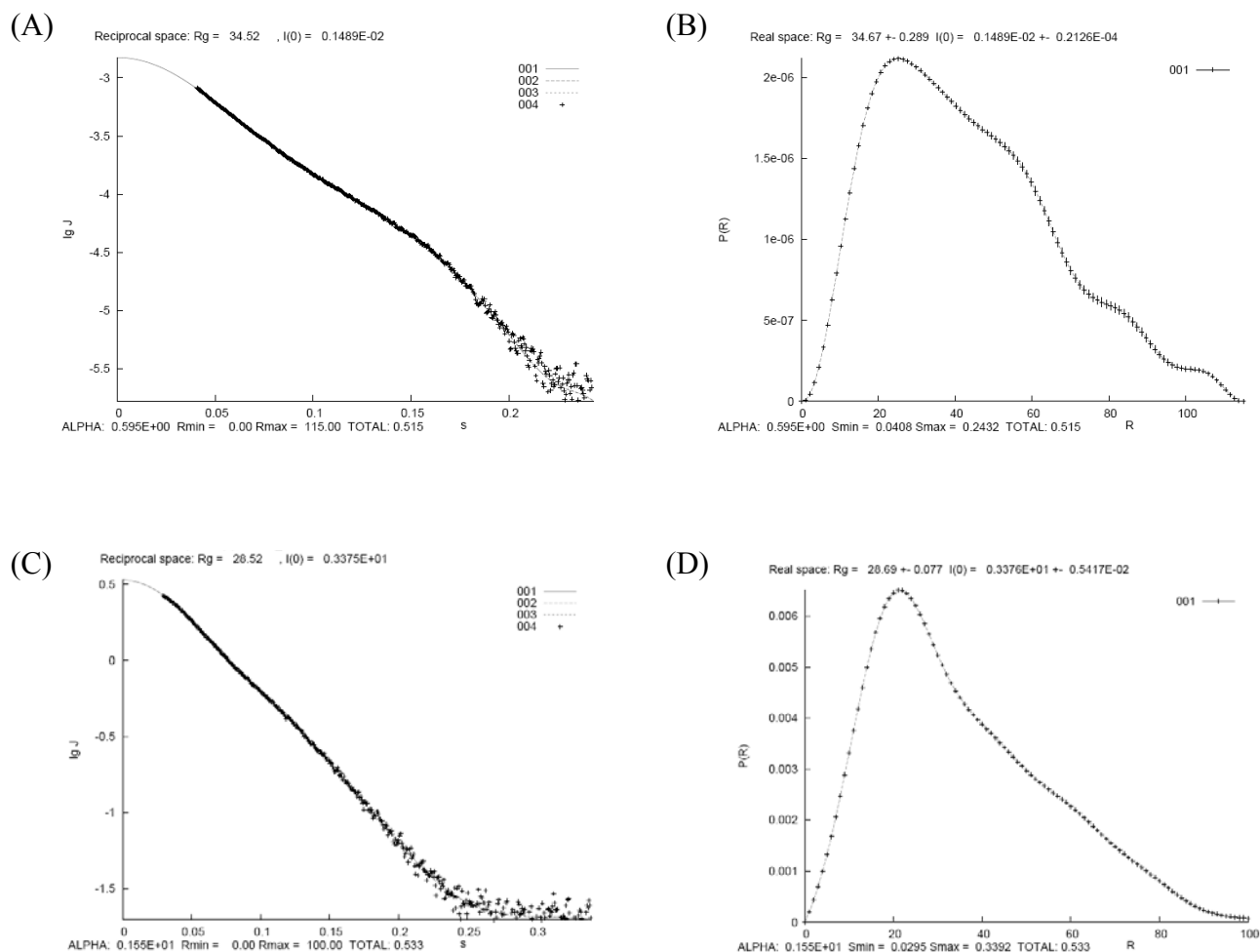


Figura 19: Dados experimentais preliminares de difração de raios X a baixo ângulo (SAXS) para as proteínas EFHC1N e EFHC1C. (A) e (C) gráfico de espalhamento de raios X gerado pela proteína em solução em função do ângulo de espalhamento, de EFHC1N e EFHC1C, respectivamente. (B) e (D) gráfico indicativo da conformação da proteína [função de distribuição das distâncias $p(r)$], de EFHC1N e EFHC1C, respectivamente.

Foi possível reconstituir o envelope da fusão SUMO-EFHC1N (figura 20) e através de estruturas cristalográficas de proteínas contendo o motivo EF-hand, foi possível criar um modelo *ab initio* da porção C-terminal (figura 21).

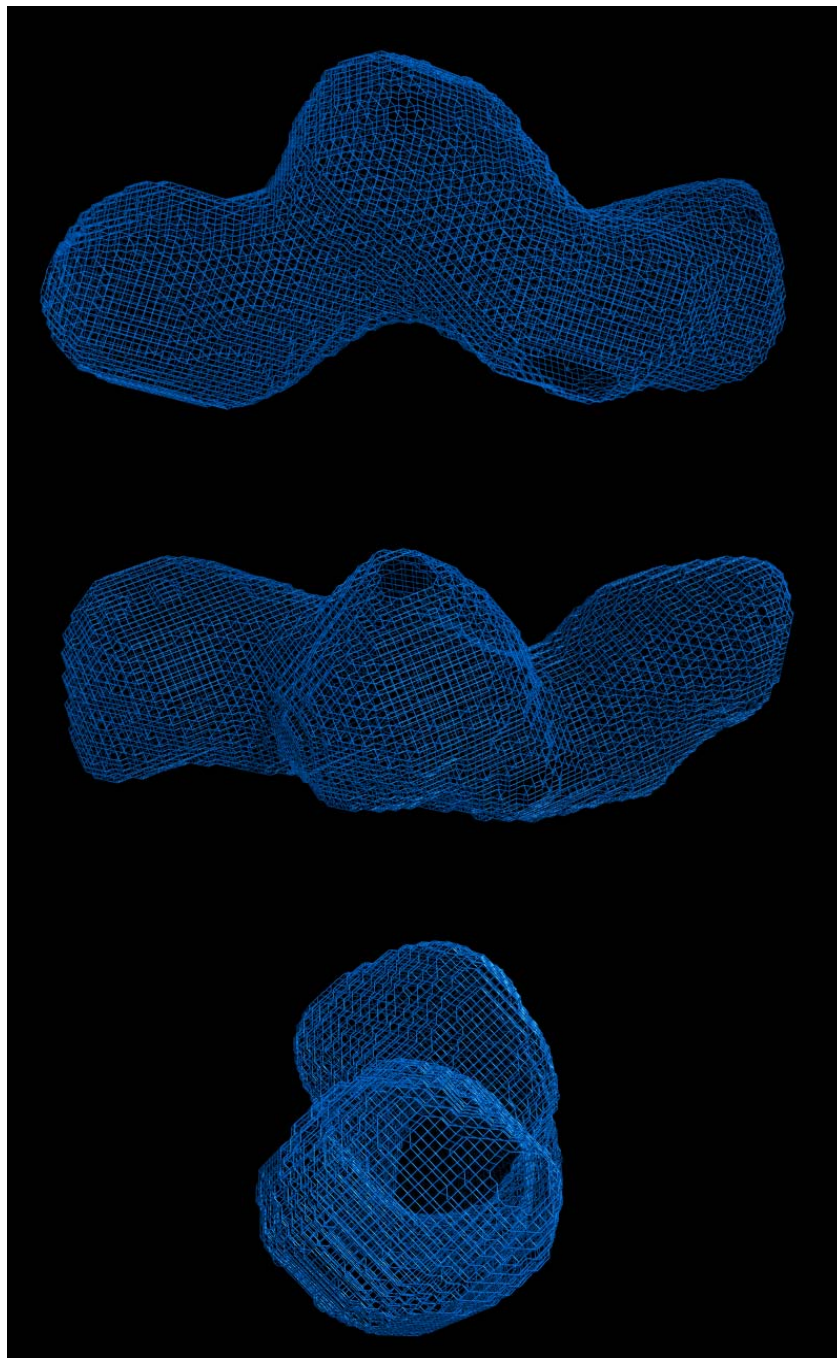


Figura 20: Envelope gerado através dos dados experimentais de SUMO-EFHC1N.

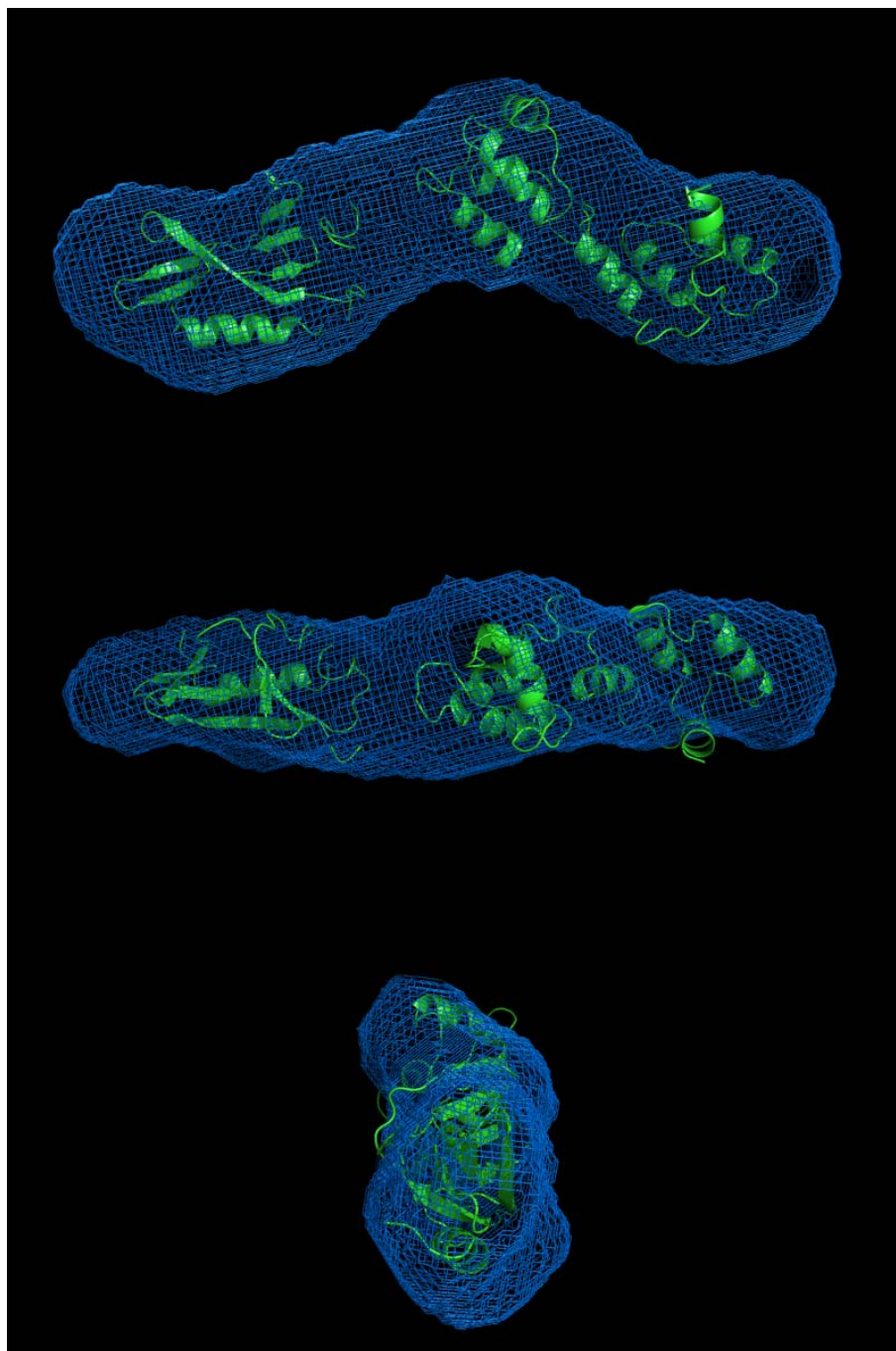


Figura 21: Modelo *ab initio* gerado da proteína EFHC1C.

Todavia, não há informação estrutural alguma do domínio DM10, o que diminui a confiança no modelo gerado. Temos dois domínios nesta proteína truncada, o EF-hand presente em diversas proteínas de uma superfamília com sua estrutura cristalográfica resolvida, e o DM10 apenas com predição de estrutura secundária disponível (King *et al.*, 2006).

No caso da proteína SUMO-EFHC1N, as medidas preliminares de SAXS mostraram que há duas populações da proteína presente na amostra. Este dado foi confirmado por filtração em gel, onde se obtêm dois picos cromatográficos que se sobrepõem, sem, contudo, se resolverem na cromatografia (dados não mostrados). Há ainda a necessidade de clivar a cauda e realizar novas medidas de SAXS, que trarão subsídios no caso de uma possível cristalização desta porção da EFHC1. Ressaltamos ainda que se não houver nenhuma proteína com sua estrutura tridimensional resolvida com uma certa porcentagem de similaridade, não será possível criar um modelo *ab initio* de baixa resolução de EFHC1N.

4.4 CONCLUSÃO

Obtivemos proteína solúvel tanto para a porção N- quanto para a C-terminal da proteína humana EFHC1, testando uma única cauda de fusão em sistema procariótico. Observamos diferentes graus de afinidade à resina de níquel Ni-NTA, devido possivelmente ao fato da cauda de histidina se apresentar mais ou menos exposta em uma ou outra construção. Poderíamos especular ainda se este diferente nível de captura não estaria relacionado às diferentes quantidades de proteína expressas pela bactéria, uma vez que a porção do gene correspondente ao N-terminal apresenta 31 códons raros, enquanto a porção do gene relativo ao C-terminal apresenta apenas 16 (anexos 4 e 6).

As diferenças observadas entre EFHC1N e EFHC1C não foram apenas em sua captura, mas principalmente no momento da clivagem da cauda SUMO. Não foi possível a clivagem da cauda nas condições testadas para a EFHC1N. Apesar disto, algumas informações foram obtidas através de filtração em gel e SAXS: a amostra apresenta-se com duas populações de proteínas distintas. Seriam diferentes estados de oligomerização?

Em relação à EFHC1C, a purificação apresenta alto rendimento e a proteína, além de bastante solúvel, é bastante estável. A clivagem da cauda foi bem sucedida e a etapa final de purificação alcançou um grau elevado de pureza. Uma caracterização espectroscópica parcial foi realizada através de dicroísmo circular, que mostrou que a proteína é composta de aproximadamente 42% de α -hélices e 15% de folhas- β . Como determinado nos experimentos de desnaturação térmica realizados, a proteína apresenta uma temperatura de transição calculada em 54,5°C. Não se realizaram experimentos de fluorescência, pois esta porção da proteína não apresenta nenhum resíduo de triptofano.

Obtivemos ainda um modelo de baixa resolução por SAXS, mostrando que a proteína em solução apresenta forma prolata. E observamos ainda evidências de que a proteína forma dímeros: além da presença de três resíduos de cisteína presentes, essa dimerização possivelmente tem envolvido outros fatores, em especial sua conformação.

Como a proteína está relacionada à epilepsia mioclônica juvenil? Esta pergunta ainda não foi respondida, mas a resolução tridimensional da estrutura da EFHC1 poderá trazer novas luzes para a compreensão desta importante forma de epilepsia, se não a mais comum, uma das que mais afetam adultos e adolescentes. Caso fique provada que sua interação com canal de cálcio, desregulando-o, leva ao desenvolvimento de crises epiléticas, fármacos poderão ser desenhadas para especificamente controlar essa interação protéica atenuando ou mesmo eliminando as crises.

Testes para cristalização de ambas as porções da proteína estão atualmente em andamento através de colaboração estabelecida com o Prof. Dr. Ricardo Aparício, do Instituto de Química da Unicamp.

4.5 PERSPECTIVAS

Novos testes de clivagem da SUMO-EFHC1N estão já em andamento em nosso laboratório. Caso seja bem sucedida, uma caracterização espectroscópica por CD e fluorescência deverá ser realizada, além de análise por SAXS.

Em relação à EFHC1C, as evidências de dimerização deverão ser confirmadas, por experimentos como DLS ou mesmo SAXS, com e sem a presença do agente redutor DTT. Além disso, experimentos mais controlados de CD deverão ser conduzidos em relação à interação da proteína com cálcio.

Como mencionado acima, ensaios de cristalografia estão em andamento, num esforço para se obter a resolução tridimensional da porção N- e C-terminal de EFHC1.

Adicionalmente, mas não menos importante, poderão ser confirmadas as interações da EFHC1 com as proteínas $\text{Ca}_v2.3$ e TRPM2, seja por métodos *in vitro*, seja por métodos *in vivo*, como o duplo híbrido de leveduras. Este último método poderá esclarecer o papel da EFHC1 na interação com proteínas do fuso mitótico (de Nijs *et al.*, 2006; Grisar *et al.*, 2006).

5. REFERÊNCIAS BIBLIOGRÁFICAS

- Ausubel, F. A., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., Struhl, K. (1998). *Current protocols in molecular biology*. In: John Wiley & Sons, New York.
- Ayerdi-Izquierdo, A., Stavrides, G., Sélles-Martínez, J. J., Larrea, L., Bovo, G., López de Munain, A., Bisulli, F., Martí-Massó, J. F., Michelucci, R., Poza, J. J., Tinuper, P., Stephani, U., Striano, P., Striano, S., Staub, E., Sarafidou, T., Hinzmann, B., Moschonas, N., Siebert, R., Deloukas, P., Nobile, C., Pérez-Tur, J. (2006). Genetic analysis of the LGI/Epitempin gene family in sporadic and familial lateral temporal lobe epilepsy. *Epilepsy Research* **70**:118-126.
- Baneyx, F. (1999). Recombinant protein expression in *Escherichia coli*. *Current Opinion in Biotechnology* **10**:411-421.
- Berridge, M. J., Lipp, P., Bootman, M. D. (2000). The versatility and universality of calcium signalling. *Nature Reviews Molecular Cell Biology* **1**:11-21.
- Bessette, P. H., Aslund, F., Beckwith, J., Georgiou, G. (1999). Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. *Proceedings of the National Academy of Science USA* **96**:13703-13708.
- Bhattacharya, S., Bunick, C. G., Chazin, W. J. (2004). Target selectivity in EF-hand calcium binding proteins. *Biochimica et Biophysica Acta* **1742**:69-79.
- Birnboim, H. C. & Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research* **7**:1513-1523.
- Bloomfield, V. A. (2000). Static and dynamic light scattering from aggregating particles. *Biopolymers* **54**:168-172.
- Boos, W. & Shuman, H. (1998). Maltose/maltodextrin system of *Escherichia coli*: transport, metabolism, and regulation. *Microbiology and Molecular Biology Reviews* **62**:204-229.
- Brent, R. & Finley Jr., R. L. (1997). Understanding gene and allele function with two-hybrid methods. *Annual Review of Genetics* **31**:663-704.
- Bulaj, G. (2005). Formation of disulfide bonds in proteins and peptides. *Biotechnology Advances* **23**:87-92.
- Burgoyne, R. D., O'Callaghan, D. W., Hasdemir, B., Haynes, L. P., Tepikin, A. V. (2004). Neuronal calcium sensor proteins: multitasking regulators of neuronal function. *Trends in Neurosciences* **27**:203-209.

- Carrio, M. M. & Villaverde, A. (2002). Construction and deconstruction of bacterial inclusion bodies. *Journal of Biotechnology* **96**:3-12.
- Chen, L., Chetkovich, D. M., Petralia, R. S., Sweeney, N. T., Kawasaki, Y., Wenthold, R. J., Brecht, D. S., Nicoll, R. A. (2000). Stargazin regulates synaptic targeting of AMPA receptors by two distinct mechanisms. *Nature* **408**:936-943.
- Chernova, O. B., Somerville, R. P., Cowell, J. K. (1998). A novel gene, LGI1, from 10q24 is rearranged and downregulated in malignant brain tumors. *Oncogene* **17**:2873-2881.
- Clark, E. D. B. (1998). Refolding of recombinant proteins. *Current Opinion in Biotechnology* **9**:157-163.
- Cossette, P., Liu, L., Brisebois, K., Dong, H., Lortie, A., Vanasse, M., Saint-Hilaire, J. M., Carmant, L., Verner, A., Lu, W. Y., Wang, Y. T., RPuleau, G.A. (2002). Mutation of GABRA1 in an autosomal dominant form of juvenile myoclonic epilepsy. *Nature Genetics* **31**:184-189.
- Crawford, H. C. & Matrisian, L. M. (1994). Tumor and stromal expression of matrix metalloproteinases and their role in tumor progression. *Invasion Metastasis* **14**:234-245.
- Criekinge, W. V. & Beyaert, R. (1999). Yeast two-hybrid: state of the art. *Biological Procedures Online* **2**:1-38. www.biologicalprocedures.com
- Davis, G. D., Elisee, C., Newham, D. M., Harrison, R. G. (1999). New fusion protein systems designed to give soluble expression in *Escherichia coli*. *Biotechnology and Bioengineering* **65**:382-388.
- de Marco, A. (2004). A step ahead: combining protein purification and correct folding selection. *Microbial Cell Factories* **3**:12.
- de Nijs, L., Lakayea, B., Coumansa, B., Léona, C., Ikeda, T., Delgado-Escueta, A. V., Grisar, T., Chanasa, G. (2006). EFHC1, a protein mutated in juvenile myoclonic epilepsy, associates with the mitotic spindle through its N-terminus. *Experimental Cell Research* **312**:2872-2879.
- Delgado-Escueta, A. V. & Enrile-Bacsal, F. (1984). Juvenile myoclonic epilepsy of Janz. *Neurology* **34**:285-294.
- di Guan, C., Li, P., Riggs, P. D., Inouye, H. (1988). Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein. *Gene* **67**:21-30.
- Dümmler, A., Lawrence, A. M., de Marco, A. (2005). Simplified screening for the detection of soluble fusion constructs expressed in *E. coli* using a modular set of vectors. *Microbial Cell Factories* **13**:34.

- Fields, S. & Song, O. (1989). A novel genetic system to detect protein-protein interaction. *Nature* **340**:245-246.
- Fletcher, C. F., Lutz, C. M., O'Sullivan, T. N., Shaughnessy, J. D. Jr., Hawkes, R., Frankel, W. N., Copeland, N. G., Jenkins, N. A. (1996). Absence epilepsy in tottering mutant mice is associated with calcium channel defects. *Cell* **87**:607-617.
- Fox, J. D. & Waugh, D. S. (2003). Maltose-binding protein as a solubility enhancer. *Methods in Molecular Biology* **205**:99-117.
- Fukata, Y., Adesnik, H., Iwanaga, T., Brecht, D. S., Nicoll, R. A., Fukata, M. (2006). Epilepsy-related ligand/receptor complex LGI1 and ADAM22 regulate synaptic transmission. *Science* **313**:1792-1795.
- Furlan, S., Roncaroli, F., Forner, F., Vitiello, L., Calabria, E., Piquer-Sirerol, S., Valle, G., Perez-Tur, J., Michelucci, R., Nobile, C. (2006). The LGI1/epitempin gene encodes two protein isoforms differentially expressed in human brain. *Journal of Neurochemistry* **98**:985-991.
- Gaberc-Porekar, V. & Menart, V. (2001). Perspectives of immobilized-metal affinity chromatography. *Journal of Biochemical and Biophysical Methods* **49**:335-360.
- Gardiner, M. (2005). Genetics of idiopathic generalized epilepsies. *Epilepsia* **46** Suppl 9:15-20.
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server; In: John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press. pp. 571-607.
- Genton, P. & Gelisse, P. (2001). Juvenile myoclonic epilepsy. *Archives of Neurology* **58**:1487-1490.
- Glatter, O. & Kratky, O. (1982) Small angle X-ray scattering. Londres: Academic Press.
- Grabarek, Z. (2006). Structural basis for diversity of the EF-hand calcium-binding proteins. *Journal of Molecular Biology* **359**:509-525.
- Greenberg, D. A., Delgado-Escueta, A. V., Widelitz, H., Sparkes, R. S., Treiman, L., Maldonado, H. M., Park, M. S., Terasaki, P. I. (1988). Juvenile myoclonic epilepsy (JME) may be linked to the BF and HLA loci on human chromosome 6. *American Journal of Medical Genetics* **31**:185-192.
- Grisar, T., de Nijs, L., Chanas, G., Léon, C., Coumans, B., Foidart, A., Lakaye, B. (2006). Some genetic and biochemical aspects of myoclonus. *Neurophysiologie Clinique* **36**:271-279.
- Gu, W., Brodtkorb, E., Piepoli, T., Finocchiaro, G., Steinlein, O. K. (2005). LGI1: a gene involved in epileptogenesis and glioma progression? *Neurogenetics* **6**:59-66.

- Hannig, G. & Makrides, S. C. (1998). Strategies for optimizing heterologous protein expression in *Escherichia coli*. *Trends in Biotechnology* **16**:54-60.
- Harrison, R. G. (2000). Expression of soluble heterologous proteins via fusion with NusA protein. *inNovations* **11**:4-7.
- Hauser, W. A. & Hesdorffer, D. C. (1990). In: *Epilepsy: Frequency, Causes and Consequences*. New York:Demos.
- Hirose, S., Mitsudome, A., Okada, M., Kaneko, S. (2005). Genetics of idiopathic epilepsies. *Epilepsia* **46** (Suppl. 1):38-43.
- Ibanez-Tallon, I., Pagenstecher, A., Fliegauf, M., Olbrich, H., Kispert, A., Ketelsen, U.-P., North, A., Heintz, N., Omran, H. (2004). Dysfunction of axonemal dynein heavy chain Mdnah5 inhibits ependymal flow and reveals a novel mechanism for hydrocephalus formation. *Human Molecular Genetics* **13**:2133-2141.
- Ikeda, K., Brown, J. A., Yagi, T., Norrander, J. M., Hirono, M., Eccleston, E., Kamiya, R., Linck, R. W. (2003). Rib72, a conserved protein associated with the ribbon compartment of flagellar microtubules and potentially involved in the linkages between outer doublet microtubules. *Journal of Biological Chemistry* **278**:7725-7734.
- Ikeda, T., Ikeda, K., Enomoto, M., Park, M. K., Hirono, M., Kamiya, R. (2005). The mouse ortholog of EFHC1 implicated in juvenile myoclonic epilepsy is an axonemal protein widely conserved among organisms with motile cilia and flagella. *FEBS Letters* **579**:819-822.
- International League Against Epilepsy, ILAE (1989). Commission on classification and terminology of the International League Against Epilepsy. Proposal for revised classification of epilepsies and epileptic syndromes. *Epilepsia* **30**:389-399.
- Janz, D. (1989). Juvenile myoclonic epilepsy. *Cleveland Clinical Journal of Medicine* **56** (suppl):23-33.
- Jones, D. T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology* **292**:195-202.
- Jouveneau, A., Eunson, L. H., Spauschus, A., Ramesh, V., Zuberi, S. M., Kullmann, D. M., Hanna, M. G. (2001). Human epilepsy associated with dysfunction of the brain P/Q-type calcium channel. *Lancet* **358**:801-807.
- Kane, J. F. (1995). Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. *Current Opinion in Biotechnology* **6**:494-500.
- Kapust, R. B. & Waugh, D. S. (2000). Controlled intracellular processing of fusion proteins by TEV protease. *Protein Expression and Purification* **19**:312-318.

- King, S. M. (2006). Axonemal protofilament ribbons, DM10 domains, and the link to juvenile myoclonic epilepsy. *Cell Motility and the Cytoskeleton* **63**:245-253.
- Ko, J. & Kim, E. (2007). Leucine-rich repeat proteins of synapses. *Journal of Neuroscience Research* Apr 30 [Epub ahead of print].
- Kobe, B. & Deisenhofer, J. (1994). The leucine-rich repeat: a versatile binding motif. *Trends in Biochemical Sciences* **19**:415-421.
- Kobe, B. & Deisenhofer J. (1995a). Proteins with leucine-rich repeats. *Current Opinion in Structural Biology* **5**:409-416.
- Kobe, B. & Deisenhofer, J. (1995b). A structural basis of the interactions between leucine-rich repeats and protein ligands. *Nature* **374**:183-186.
- Krex, D., Hauses, M., Appelt, H., Mohr, B., Ehninger, G., Schackert, H. K., Schackert, G. (2002). Physical and functional characterization of the human LGI1 gene and its possible role in glioma development. *Acta Neuropathologica* **103**:255-266.
- Kunapuli, P., Kasyapa, C. S., Hawthorn, L., Cowell, J. K. (2004). LGI1, a putative tumor metastasis suppressor gene, controls in vitro invasiveness and expression of matrix metalloproteinases in glioma cells through the ERK1/2 pathway. *Journal of Biological Chemistry* **279**:23151-23157.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**:680-685.
- LaVallie, E. R., DiBlasio, E. A., Kovacic, S., Grant, K. L., Schendel, P. F., McCoy, J. M. (1993). A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. *Biotechnology* **11**:187-193.
- Levitan, I. B. (1999). It is calmodulin after all! Mediator of the calcium modulation of multiple ion channels. *Neuron* **22**:645-648.
- Lewit-Bentley, A. & Réty, S. (2000). EF-hand calcium-binding proteins. *Current Opinion in Structural Biology* **10**:637-643.
- Lilie, H., Schwarz, E., Rudolph, R. (1998). Advances in refolding of proteins produced in *E. coli*. *Current Opinion in Biotechnology* **9**:497-501.
- Lin, Z. & Rye, H. S. (2006). GroEL-mediated protein folding: making the impossible, possible. *Critical Reviews in Biochemistry and Molecular Biology* **41**:211-239.
- Lipfert, J. & Doniach, S. (2007) Small-angle X-ray scattering from RNA, protein, and protein complexes. *The Annual Review of Biophysics and Biomolecular Structure* **36**:307-327.

- Lobley, A., Whitmore, L., Wallace, B. A. (2002). DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra. *Bioinformatics* **18**:211-212.
- Malakhov, M. P., Mattern, M. R., Malakhov, O. A., Drinker, M., Weeks, S. D., Butt, T. R. (2004). SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. *Journal of Structural and Functional Genomics* **5**:75-86.
- Marblestone, J. G., Edavettal, S. C., Lim, Y., Lim, P., Zuo, X., Butt, T. R. (2006). Comparison of SUMO fusion technology with traditional gene fusion systems: enhanced expression and solubility with SUMO. *Protein Science* **15**:182-189.
- McGuffin, L. J., Bryson, K., Jones, D. T. (2000). The PSIPRED protein structure prediction server. *Bioinformatics* **16**:404-405.
- Mei, Z.-Z., Xia, R., Beech, D. J., Jiang, L.-H. (2006). Intracellular coiled-coil domain engaged in subunit interaction and assembly of melastatin-related transient receptor potential channel 2. *Journal of Biological Chemistry* **281**: 38748-38756.
- Morante-Redolat, J. M., Gorostidi-Pagola, A., Piquer-Sirerol, S., Sáenz, A., Poza, J. J., Galán, J., Gesk, S., Sarafidou, T., Mautner, V. F., Binelli, S., Staub, E., Hinzmann, B., French, L., Prud'homme, J. F., Passarelli, D., Scannapieco, P., Tassinari, C. A., Avanzini, G., Martí-Massó, J. F., Kluwe, L., Deloukas, P., Moschonas, N. K., Michelucci, R., Siebert, R., Nobile, C., Pérez-Tur, J., López de Munain, A. (2002). Mutations in the LGI1/Epitempin gene on 10q24 cause autosomal dominant lateral temporal epilepsy. *Human Molecular Genetics*. **11**:1119-1128.
- Mori, Y., Hara, Y., Aguan, K., Katano, M., Yamamoto, S., Suzuki, T., Yamakawa, K. (2005). Molecular and functional interaction of the juvenile myoclonic epilepsy-related gene EFHC1 with the redox-sensitive TRPM2 channel. In: *Washington: Society for Neuroscience 2005*.
- Munier, A., Serres, C., Kann, M., Boissan, M., Lesaffre, C., Capeau, J., Fouquet, J., Lacombe, M. (2003). Nm23/NDP kinases in human male germ cells: role in spermiogenesis and sperm motility? *Experimental Cell Research* **289**:295-306.
- Murphy, R. M. (1997) Static and dynamic light scattering of biological macromolecules: what can we learn? *Current Opinion in Biotechnology* **8**: 25-30.
- Nabeshima, K., Inoue, T., Shimao, Y., Sameshima, T. (2002). Matrix metalloproteinases in tumor invasion: role for cell migration. *Pathology International* **52**:255-264.
- Nakamura, Y., Gojobori, T., Ikemura, T. (2000). Codon usage tabulated from the international DNA sequence databases: status for the year 2000. *Nucleic Acids Research* **28**:292.
- Nakayama, J., Fu, Y. H., Clark, A. M., Nakahara, S., Hamano, K., Iwasaki, N., Matsui, A., Arinami, T., Ptacek, L. J. (2002). A nonsense mutation of the MASS1 gene in a family with febrile and afebrile seizures. *Annals of Neurology* **52**:654-657.

- Nallamsetty, S. & Waugh, D. S. (2006). Solubility-enhancing proteins MBP and NusA play a passive role in the folding of their fusion partners. *Protein Expression and Purification* **45**:175-82.
- Nicoll, R. A., Tomita, S., Brecht, D. S. (2006). Auxiliary subunits assist AMPA-type glutamate receptors. *Science* **311**:1253-1256.
- Niegowski, D., Hédren, M., Nordlund, P., Eshaghi, S. (2006). A simple strategy towards membrane protein purification and crystallization. *International Journal of Biological Macromolecules* **39**:83-87.
- Nominé, Y., Ristriani, T., Laurent, C., Lefevre, J. F., Weiss, E., Trave, G.(2001). A strategy for optimizing the monodispersity of fusion proteins: application to purification of recombinant HPV E6 oncoprotein. *Protein Engineering* **14**:297-305.
- Novak, U. (2004). ADAM proteins in the brain. *Journal of Clinical Neurosciences* **11**:227-235.
- Novy, R., Drott, D., Yaeger, K., Mierenhof, R. (2001). Overcoming the codon bias of *E. coli* for enhanced protein expression. *inNovations* **12**:1-3.
- Nygren, P. A., Stahl, S., Uhlen, M. (1994). Engineering proteins to facilitate bioprocessing. *Trends in Biotechnology* **12**:184-188.
- Oganesyan, N., Kim, S. -H., Kim, R. (2004). On-column chemical refolding of proteins. *PharmaGenomics* **4**:22-26.
- Oganesyan, N., Kim, S. -H., Kim, R. (2005). On-column protein refolding for crystallization. *Journal of Structural and Functional Genomics* **6**:177-182.
- Ottman, R., Risch, N., Hauser, W. A., Pedley, T. A., Lee, J. H., Barker-Cummings, C., Lustenberger, A., Nagle, K. J., Lee, K. S., Scheuer, M. L., Neystat, M., Susser, M., Wilhelmsen, K. C. (1995). Localization of a gene for partial epilepsy to chromosome 10q. *Nature Genetics* **10**:56-60.
- Pal, D. K., Evgrafov, O. V., Tabares, P., Zhang, F., Durner, M., Greenberg, D. A. (2003). BRD2 (RING3) is a probable major susceptibility gene for common Juvenile Myoclonic Epilepsy. *American Journal of Human Genetics* **73**:261-270.
- Papish, A. L., Tari, L. W., Vogel, H. J. (2002) Dynamic light scattering study of calmodulin-target peptide complexes. *Biophysical Journal* **83**:1455-1464.
- Patel-King, R. S., Benashski, S. E., King, S. M. (2002). A bipartite Ca²⁺-regulated nucleoside-diphosphate kinase system within the *Chlamydomonas* flagellum. The regulatory subunit p72. *Journal of Biological Chemistry* **277**:34271-34279.

- Patel-King, R. S., Gorbatyuk, O., Takebe, S., King, S. M. (2004). Flagellar radial spokes contain a Ca^{2+} -stimulated nucleoside diphosphate kinase. *Molecular Biology of the Cell* **15**:3891-3902.
- Peti, W. & Page, R. (2007). Strategies to maximize heterologous protein expression in *Escherichia coli* with minimal cost. *Protein Expression and Purification* **51**:1-10.
- 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**:3113-3125.
- Philo, J. S. (2006). Is any measurement method optimal for all aggregate sizes and types? *The American Association of Pharmaceutical Scientists Journal* **8**:E564-E571.
- Piepoli, T., Jakupoglu, C., Gu, W., Lualdi, E., Suarez-Merino, B., Poliani, P.L., Cattaneo, M. G., Ortino, B., Goplen, D., Wang, J., Mola, R., Inverardi, F., Frassoni, C., Bjerkvig, R., Steinlein, O., Vicentini, L. M., Brustle, O., Finocchiaro, G. (2006). Expression studies in gliomas and glial cells do not support a tumor suppressor role for LGI1. *Neuro-oncology* **8**:96-108.
- Poza, J. J., Saenz, A., Martinez-Gil, A., Cheron, N., Cobo, A.M., Urtasun, M., Marti-Masso, J. F., Grid, D., Beckmann, J. S., Prud'homme, J. F., Lopez de Munain, A. (1999). Autosomal dominant lateral temporal epilepsy: clinical and genetic study of a large Basque pedigree linked to chromosome 10q. *Annals of Neurology* **45**:182-188.
- Pryor, K. D. & Leiting, B. (1997). High-level expression of soluble protein in *Escherichia coli* using a His₆-tag and maltose-binding protein double-affinity fusion system. *Protein Expression and Purification* **10**:309-319.
- Rost, B., Yachdav, G., Liu, J. (2004). The PredictProtein Server. *Nucleic Acids Research* **32** (Web Server issue):W321-W326.
- Routzahn, K. M. & Waugh, D. S. (2002). Differential effects of supplementary affinity tags on the solubility of MBP fusion proteins. *Journal of Structural and Functional Genomics* **2**:83-92.
- Sagane, K., Hayakawa, K., Kai, J., Hirohashi, T., Takahashi, E., Miyamoto, N., Ino, M., Oki, T., Yamazaki, K., Nagasu, T. (2005). Ataxia and peripheral nerve hypomyelination in ADAM22-deficient mice. *BMC Neuroscience* **6**:33.
- Sambrook, J., Fritsch, E. F., Maniatis, T. (1989). *Molecular cloning: a laboratory manual*. 2nd ed. New York: Cold Spring Harbor Laboratory.
- Sander, J. W & Shorvon, S.D. (1996). Epidemiology of the epilepsies. *Journal of Neurology, Neurosurgery, and Psychiatry* **61**:433-443.

- Scheel, H., Tomiuk, S., Hofmann, K. (2002). A common protein interaction domain links two recently identified epilepsy genes. *Human Molecular Genetics* **11**:1757-1762.
- Schrödel, A. & de Marco, A. (2005). Characterisation of the aggregates formed during recombinant protein expression in bacteria. *BMC Biochemistry* **6**:10.
- Schulte, U., Thumfart, J. O., Klocker, N., Sailer, C. A., Bildl, W., Biniossek, M., Dehn, D., Deller, T., Eble, S., Abbass, K., Wangler, T., Knaus, H. G., Fakler, B. (2006). The epilepsy-linked Lgi1 protein assembles into presynaptic Kv1 channels and inhibits inactivation by Kvbeta1. *Neuron* **49**:697-706.
- Scopes, R. K. (1974). Measurement of protein by spectrophotometry at 205 nm. *Analytical Biochemistry* **59**:277-282.
- Senechal, K.R., Thaller, C., Noebels, J.L. (2005). ADPEAF mutations reduce levels of secreted LGI1, a putative tumor suppressor protein linked to epilepsy. *Human Molecular Genetics* **14**:1613-1620.
- Shevchenko, A., Wilm, M., Vorm, O., Mann, M. (1996). Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Analytical Chemistry* **68**:850-858.
- Singh, S. M. & Panda, A. K. (2005). Solubilization and refolding of bacterial inclusion body proteins. *Journal of Bioscience and Bioengineering* **99**:303-310.
- Sirerol-Piquer, M. S., Ayerdi-Izquierdo, A., Morante-Redolat, J. M., Herranz-Pérez, V., Favell, K., Barker, P. A., Pérez-Tur, J. (2006). The epilepsy gene LGI1 encodes a secreted glycoprotein that binds to the cell surface. *Human Molecular Genetics* **15**:3436-3445.
- Skelton, N. J., Kördel, J., Akke, M., Forsén, S., Chazin, W. J. (1994). Signal transduction versus buffering activity in Ca²⁺-binding proteins. *Nature Structural Biology* **1**:239-245.
- Skradski, S. L., Clark, A. M., Jiang, H., White, H. S., Fu, Y. H., Ptacek, L. J. (2001). A novel gene causing a mendelian audiogenic mouse epilepsy. *Neuron* **31**:537-544.
- Smith, D. B. & Johnson, K. S. (1988). Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* **67**:31-40.
- Smyth, D. R., Mrozkiewicz, M. K., McGrath, W. J., Listwan, P., Kobe, B. (2003). Crystal structures of fusion proteins with large affinity tags. *Protein Science* **12**:1313-1322.
- Smolka, M. B. (2002). Análise proteômica de *Xylella fastidiosa* e desenvolvimento de novo método para análise quantitativa de proteomas por espectrometria de massas. Tese (doutorado), Universidade Estadual de Campinas, Instituto de Biologia. pp. 125-126.
- Sorensen, H. P. & Mortensen, K. K. (2005a). Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microbial Cell Factories* **4**:1-8.

- Sorensen, H. P. & Mortensen, K. K. (2005b). Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *Journal of Biotechnology* **115**:113-128.
- Staub, E., Perez-Tur, J., Siebert, R., Nobile, C., Moschonas, N. K., Deloukas, P., Hinzmann, B. (2002). The novel EPTP repeat defines a superfamily of proteins implicated in epileptic disorders. *Trends in Biochemical Sciences* **27**:441-444.
- Steinlein, O. K. (2003). Genes and mutations in human idiopathic epilepsy. *Brain and Development* **26**:213-218.
- Steinlein, O. K. (2004). Genetic mechanisms that underlie epilepsy. *Nature Reviews Neuroscience* **5**:400-408.
- Suissa, M., Place, C., Goillot, E., Berge, B., Freyssingeas, E. (2007). Dynamic light scattering as an investigating tool to study the global internal dynamics of a living cell nucleus. *Europhysics Letters* **78**:38005.
- Suzuki, T., Delgado-Escueta, A. V., Aguan, K., Alonso, M. E., Shi, J., Hara, Y., Nishida, M., Numata, T., Medina, M. T., Takeuchi, T., Morita, R., Bai, D., Ganesh, S., Sugimoto, Y., Inazawa, J., Bailey, J. N., Ochoa, A., Jara-Prado, A., Rasmussen, A., Ramos-Peek, J., Cordova, S., Rubio-Donnadieu, F., Inoue, Y., Osawa, M., Kaneko, S., Oguni, H., Mori, Y., Yamakawa, K. (2004). Mutations in EFHC1 cause juvenile myoclonic epilepsy. *Nature Genetics* **36**:842-849.
- Svergun, D. I. (1992). Determination of the regularization parameters in indirect-transform methods using perceptual criteria. *Journal of Applied Crystallography* **25**:495-503.
- Svergun, D. I. (1999). Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophysical Journal* **76**:2879-2886.
- Svergun, D. I., Petoukhov, M. V., Koch, M. H. J. (1995). CRY SOL – A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *Journal of Applied Crystallography* **28**:768-773.
- Svergun, D. I., Petoukhov, M. V., Koch, M. H. J. (2001). Determination of domain structure of proteins from X-ray solution scattering. *Biophysical Journal* **80**:2946-2953.
- Tang, Y. –C., Chang, H. –C., Roeben, A., Wischnewski, D., Wischnewski, N., Kerner, M. J., Hartl, F. U., Hayer-Hartl, M. (2006). Structural features of the GroEL-GroES nano-cage required for rapid folding of encapsulated protein. *Cell* **125**:903-914.
- Taylor, G. (2003). The phase problem. *Acta Crystallographica* **D59**:1881-1890.
- Terpe, K. (2003). Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Applied Microbiology and Biotechnology* **60**:523-533.

- Tomita, S., Adesnik, H., Sekiguchi, M., Zhang, W., Wada, K., Howe, J. R., Nicoll, R. A., Bredt, D. S. (2005). Stargazin modulates AMPA receptor gating and trafficking by distinct domain. *Nature* **435**:1052-1058.
- Tropea, J. E., Cherry, S., Nallamsetty, S., Bignon, C., Waugh, D. S. (2007). A generic method for the production of recombinant proteins in *Escherichia coli* using a dual hexahistidine-maltose-binding protein affinity tag. *Methods in Molecular Biology* **363**:1-19.
- Tsumoto, K., Umetsu, M., Kumagai, I., Ejima, D., Philo, J. S., Arakawa, T. (2004). Role of arginine in protein refolding, solubilization, and purification. *Biotechnology Progress* **20**:1301-1308.
- Tsumoto, K., Ejima, D., Kita, Y., Arakawa, T. (2005). Review: why is arginine effective in suppressing aggregation? *Protein and Peptide Letters* **12**:613-619.
- Turnbull, J., Lohi, H., Kearney, J.A., Rouleau, G. A., Delgado-Escueta, A. V., Meisler, M. H., Cossette, P., Minassian, B. A. (2005). Sacred disease secrets revealed: the genetics of human epilepsy. *Human Molecular Genetics* **14**:2491-2500.
- Yasukawa, T., Kanei-Ishii, C., Maekawa, T., Fujimoto, J., Yamamoto, T., Ishii, S. (1995). Increase of solubility of foreign proteins in *Escherichia coli* by coproduction of the bacterial thioredoxin. *Journal of Biological Chemistry* **270**:25328-25331.
- Walther, D., Cohen, F. E., Doniach, S. (2000). Reconstruction of low resolution three-dimensional density maps from one-dimensional small angle X-ray scattering data for biomolecules in solution. *Journal of Applied Crystallography* **33**:350-363.
- Waugh, D. S. (2005). Making the most of affinity tags. *Trends in Biotechnology* **23**:316-320.
- Wennerstrand, P., Johannsson, I., Ugochukwu, E., Kavanagh, K., Edwards, A., Arrowsmith, C., Williams, L., Sundstrom, M., Guo, K., von Delft, F. & Oppermann, U. (2005). Crystal structure of human potassium channel Kv beta-subunit (KCNA2). To be Published, (<http://www.rcsb.org/pdb/explore/explore.do?structureId=1ZSX>).
- Whitmore, L. & Wallace, B. A. (2004). DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Research* **32**:W668-W673.
- Wilkinson, D. L. & Harrison, R. G. (1991). Predicting the solubility of recombinant proteins in *Escherichia coli*. *Biotechnology* **9**:443-448.
- Wilson, W. W. (2003). Light scattering as a diagnostic for protein crystal growth – A practical approach. *Journal of Structural Biology* **142**:56-65.
- Winawer, M. R., Ottman, R., Hauser, A. W., Pedley, T. A. (2000). Autosomal dominant partial epilepsy with auditory features: defining the phenotype. *Neurology* **54**:2173-2176.

- Wolf, P. (1992). Juvenile myoclonic epilepsy. Epilepsy syndrome in infancy, childhood and adolescence. In: Company JL, editor. pp. 313-327.
- Zielinski, J. J. (1988). Epidemiology of epilepsy. In: Laidlaw, J., Richens, A., Oxley, J. (eds) *A Textbook of Epilepsy*, Third edition, Churchill Livingstone, New York, pp. 21-48.
- Zhang, S. P., Zubay, G., Goldman, E. (1991). Low-usage codons in *Escherichia coli*, yeast, fruit fly and primates. *Gene* **105**:61-72.
- Zhang, X., Beuron, F., Freemont, P. S. (2002). Machinery of protein folding and unfolding. *Current Opinion in Structural Biology* **12**:231-238.
- Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* **31**:3406-3415.

6. ENDEREÇOS ELETRÔNICOS

Abnova:

www.abnova.com.tw/product_search/PS_detail_protein.asp?GS_searchbox=LGI1&catalog_id=H00009211-P01&geneid=9211&family_id= (acessado em fevereiro de 2005)

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

ClustalW: <http://align.genome.jp/>

DICHROWEB: <http://public-1.cryst.bbk.ac.uk/cdweb>

European Molecular Biology Laboratory (EMBL; Protein Expression and Purification Core Facility): www.pepcore.embl.de/index.html

Expasy: www.expasy.org

Kazusa DNA Research Institute: www.kazusa.or.jp/eng/index.html

Mfold: www.bioinfo.rpi.edu/applications/mfold/old/dna/

NCBI/Pubmed: www.ncbi.nlm.nih.gov/entrez/

Rare Codon Caltor: www.doe-mbi.ucla.edu/~sumchan/caltor.html

SMART: <http://smart.embl-heidelberg.de/>

PDB: <http://www.pdb.org/pdb/home/home.do>

Pfam: <http://pfam.janelia.org/hmmsearch.shtml>

PHD: www.predictprotein.org/

PROSITE: <http://expasy.org/prosite/>

ProtParam Tool: <http://ca.expasy.org/toos/protparam.html>

PSIPRED: <http://bioinf.cs.ucl.ac.uk/psipred/psiform.html>

Recombinant Protein Solubility Prediction: <http://biotech.ou.edu/>

Vanderbilt University (Center for Structural Biology): <http://structbio.vanderbilt.edu/wetlab/>

Webcutter 2.0: <http://rna.lundberg.gu.se/cutter2/>

7. ANEXOS

ANEXO 1

LG11 codon usage

ATG GAA TCA GAA AGA AGC AAA AGG ATG GGA AAT GCC TGC ATT CCC CTG AAA AGA ATT GCT TAT TTC CUA TGT
 CTC TTA TCT GCG CTT TTG CTG ACT GAG GGG AAG AAA CCA GCG AAG CCA AAA TGC CCT GCC GTG TGT ACT TGT
 ACC AAA GAT AAT GCT TTA TGT GAG AAT GCC AGA TCC ATT CCA CGC ACC GTT CCT CCT GAT GTT ATC TCA TTA
 TCC TTT GTG AGA TCT GGT TTT ACT GAA ATC TCA GAA GGG AGT TTT TTA TTC ACG CCA TCG CTG CAG CTC TTG
 TTA TTC ACA TCG AAC TCC TTT GAT GTG ATC AGT GAT GCT TTT ATT GGT CTT CCA CAT CUA GAG TAT TTA
 TTC AUA GAA AAC AAC ATC AAT AAG TCA ATT TCA AGA CAT ACT TTC CCG GGA CUA AAG TCA TTA ATT CAC TTG
 AGC CTT GCA AAC AAC AAT CTC CAG ACA CTC CCA AAA GAT ATT TTC AAA GGC CTG GAT TCT TTA ACA AAT GTG
 GAC CTG AGG GGT AAT TCA TTT AAT TGT GAC TGT AAA CTG AAA TGG CUA GTG GAA TGG CTT GGC CAC ACC AAT
 GCA ACT GTT GAA GAC ATC TAC TGC GAA GGC TCA CCA GAA TAC AAG AAG CGC AAA ATC AAT AGT CTC TCC TCG
 AAG GAT TTC GAT TGC ATC ATT ACA GAA TTT GCA AAG TCT CAA GAC CTG CCT TAT CAA TCA TTG TCC AUA GAC
 ACT TTT TCT TAT TTG AAT GAT GAG TAT GTA GTC ATC GCT CAG CCT TTT ACT GGA AAA TGC ATT TTC CTT GAA
 TGG GAC CAT GTG GAA AAG ACC TTC CGG AAT TAT GAC AAC ATT ACA GGC ACA TCC ACT GTA GTA TGC AAG CCT
 AUA GTC ATT GAA ACT CAG CTC TAT GTT ATT GTG GCC CAG CTG TTT GGT GGC TCT CAC ATC TAT AAG CGA GAC
 AGT TTT GCA AAT AAA TTC AUA AAA ATC CAG GAT ATT GAA ATT CTC AAA ATC CGA AAA CCC AAT GAC ATT GAA
 ACA TTC AAG ATT GAA AAC AAG TGG TAC TTT GTT GTT GCT GAC AGT TCA AAA GCT GGT TTT ACT ACC ATT TAC
 AAA TGG AAC GGA AAC GGA TTC TAC TCC CAT CAA TCC TTA CAC GCG TGG TAC AGG GAC ACT GAT GTG GAA TAT
 CUA GAA AUA GTC AGA ACA CCT CAG ACA CTC AGA ACG CCT CAT TTA ATT CTG TCT AGT AGT TCC CAG CGT CCT
 GTA ATT TAT CAG TGG AAC AAA GCA ACA CAA TTA TTC ACT AAC CAA ACT GAC ATT CCT AAC ATG GAG GAT GTG
 TAC GCA GTG AAG CAC TTC TCA GTG AAA GGG GAC GTG TAC ATT TGC TTG ACA AGA TTC ATT GGT GAT TCC AAA
 GTC ATG AAA TGG GGA GGC TCC TCG TTC CAG GAT ATT CAG AGG ATG CCA TCG CGA GGA TCC ATG GTG TTC CAG
 CCT CTT CAA AUA AAT AAT TAC CAA TAT GCA ATT CTT GGA AGT GAT TAC TCC TTT ACT CAA GTG TAT AAC TGG
 GAT GCA GAG AAA GCC AAA TTT GTG AAA TTT CAG GAA TTA AAT GTT CAG GCA CCA AGA TCA TTC ACA CAT GTG
 TCC ATT AAT AAG CGT AAT TTT CTT TTT GCT TCC AGT TTT AAG GGA AAT ACA CAG ATT TAC AAA CAT GTC AUA
 GTT GAC TTA AGC GCA TGA

The Number of Bases in the above Sequence = 1674

The Number of Codons in the above Sequence = 558

Amino Acid	Rare Codon	Frequency of Occurrence
Arginine	CGA	3
	CGG	2
	AGG	4
	AGA	9
Glycine	GGA	9
	GGG	3
Isoleucine	AUA	7
Leucine	CUA	5
Proline	CCC	3
Threonine	ACG	2

Repeated and/or Consecutive Rare Codons

CGG GGA CUA = 1

AGA ACG = 1

CGA GGA = 1

ATC trinca de bases correspondente ao resíduo 222

ACA trinca de bases correspondente ao resíduo 224

GGA TCC sítio de restrição interno BamHI

ANEXO 3

Proteína	bp	AA	MW	pI teórico	negatively charged residues	positively charged residues	Vetor
Caudas							
SUMO	357	119	13,4	5,9	21	16	pETSUMO
MBP	1176	392	43,0	5,5	52	47	pETMBP
NusA	1557	518	57,3	4,6	99	59	pETNusA
GST	690	229	26,7	5,9	35	31	pGEX-4T1
Trx	327	109	17,9	5,2	26	16	pET32a
His ₆	57	19	2,0	9,6	0	1	pET28a
Proteínas-alvo							
LGI1 (AA 1-557)	1674	557	63,8	8,1	56	62	–
LGI1N (AA 1-222)	669	222	24,8	8,3	24	27	–
LGI1C (AA 224-557)	1005	335	38,9	8,5	32	35	–
LGI1 isoforma	892	291	33,1	6,6	34	33	–
LGI1 tr5	748	243	28,4	9,3	21	28	–
LGI1 tr6	442	141	16,3	9,2	11	15	–
ED-ADAM22	1473	487	53,8	5,7	63	54	–
hKvβ1.3	1203	401	44,7	9,4	38	52	–
Proteínas de fusão							
GST-LGI1	–	786	90,4	7,7	91	93	pGEX-4T1
GST-LGI1N	–	450	51,4	6,8	59	58	pGEX 4T1
GST-LGI1C	–	565	65,7	6,9	67	66	pGEX-4T1
Trx-LGI1	–	724	81,7	6,6	82	78	pET32a
Trx-LGI1N	–	387	42,6	6,0	24	27	pET32a
Trx-LGI1C	–	503	57,0	6,2	58	51	pET32a
SUMO-LGI1	–	676	77,2	7,6	77	78	pETSUMO
SUMO-LGI1C	–	455	52,4	6,8	53	51	pETSUMO
MBP-LGI1C	–	729	82,2	6,2	84	77	pETMBP
NusA-LGI1C	–	855	96,5	5,0	131	94	pETNusA
SUMO-hKvβ1.3	–	520	58,1	9,0	59	68	pETSUMO
SUMO-ADAM22	–	606	67,2	5,8	84	70	pETSUMO
Controles e proteases							
MBP-GFP	–	633	70,1	5,6	86	69	pETMBP
NusA-GFP	–	759	84,4	4,8	133	86	pETNusA
His-TEV	–	249	28,4	8,6	21	24	pS219V
His-Ulp1	–	232	26,8	6,8	30	29	pH6Ulp1

ANEXO 4

EFHC1N codon usage

ACT TGG TCT **CCC** ATG AAA CAA GCC CCA CCT GCG GAT TTT ATT CCT GCG CAT GTG GCC TTT GAC AAA AAG GTA
 CTG AAA TTT GAT GCC TAT TTC CAA GAA GAT GTT CCT ATG TCA ACT GAG GAA CAG TAT **AGG** ATC CGT CAG GTG
 AAC ATT TAC TAT TAT **CUA** GAA GAT GAC AGC ATG TCT GTC **AUA** GAG CCT GTT GTA GAA AAT TCT **GGA** ATC CTT
 CAA GGC AAG TTA **AUA** AAA CGC CAG **CGG** **CUA** GCC AAG AAT GAC **CGG** GGT GAC CAT TAC CAT TGG AAA GAC **CUA**
 AAT **CGA** **GGA** **AUA** AAC ATC ACA ATT TAT GGC AAA ACT TTC CGC GTT GTT GAC TGT GAC CAA TTC ACA CAG GTA
 TTT TTA GAA AGC CAA **GGA** ATT GAG TTA AAT CCA CCA GAG AAG ATG GCT CTT GAT CCT TAC ACT GAA CTC **CGA**
 AAA CAG CCT CTT CGT AAG TAT GTC ACC CCA TCA GAC TTT GAT CAA CTC AAG CAA TTT **CTC** ACC TTT GAC AAA
 CAG GTC CTT **CGA** TTC TAT GCA ATC TGG GAT GAT ACA GAC AGC ATG TAT GGT GAA TGT **CGG** ACC TAC ATC ATT
 CAT TAC TAT CTT ATG GAT GAT **ACG** GTG GAA ATT **CGA** GAG GTC CAC GAA **CGG** AAT GAT **GGG** **AGA** GAT CCT TTC
 CCA CTC **CUA** ATG AAC CGC CAG CGT GTG **CCC** AAA GTT TTG GTG GAA AAT GCA AAG AAC TTC CCT CAG TGT GTG
CUA GAA ATC TCT GAC CAA GAA GTG TTG GAA TGG TAT ACT GCT AAA GAC TTC ATT GTT **GGG** AAG TCA CTC ACT
 ATC CTT **GGG** **AGA** ACT TTC ATT TAT TAT GAT TGT GAT CCA TTT ACT **CGA** **CGG** TAT TAC AAA GAG AAG TTT **GGA**
 ATC ACT GAT TTA TGA CGC GAG GAT

The Number of Bases in the above Sequence = 888

The Number of Codons in the above Sequence = 296

Amino Acid	Rare Codon	Frequency of Occurrence
Arginine	CGA	5
	CGG	5
	AGG	1
	AGA	2
Glycine	GGA	4
	GGG	3
Isoleucine	AUA	3
Leucine	CUA	5
Proline	CCC	2
Threonine	ACG	1

Repeated and/or Consecutive Rare Codons

CGG **CUA** = 1

CGA **GGA** **AUA** = 1

GGG **AGA** = 2

CGA **CGG** = 1

ANEXO 5

PSIPRED PREDICTION RESULTS EFHC1N

Key

Conf: Confidence (0=low, 9=high)

Pred: Predicted secondary structure (H=helix, E=strand, C=coil)

AA: Target sequence

PSIPRED HFORMAT (PSIPRED V2.5 by David Jones)

```
Conf: 988889887533431467637907877545641465552000102279999982288099
Pred: CCCCCCCCCCCCCCHHHHHHCCEEEEEEEEECCCCCCHHCCEEEEEEEEECCCCCE
AA: TWSPMKQAPPADFIPAHVAFDKKVLKFDAYFQEDVPMSTEEQYRIRQVNIYYLEDDSMS
      10          20          30          40          50          60

Conf: 973473116898752557754578888861359990259748997489999828678999
Pred: EEECCCCCCCCCEEEEECCCCCCCCCHHHCCCCCEEEEEEEEECCCCCHHHH
AA: VIEPVVENSILQGKLIKRLAKNDRGDHYHWDLNRGINITIYGKTRFRVVDCDQFTQV
      70          80          90         100         110         120

Conf: 999869637887667888368885101247787556799999987289278999997077
Pred: HHHHCCCCCCCCCCCCCHHHHHHCCCCCCHHHHHHHHHHHHCCCCCEEEEEEECCC
AA: FLESQGIELNPPEKMALDPYTELRKQPLRKVYVTPSDFDQLKQFLTFDKQVLRFYAIWDDT
      130         140         150         160         170         180

Conf: 877670569999999857919999736765887775300253114554455432233212
Pred: CCCCCCEEEEEEEEECEEEEEEEEECCCCCCCCCEEEEECECCCCCCCCCCCCC
AA: DSMYGECDRTYIIHYLMDDTVEIREVHERNDGRDPFPLLMNRQVRPKVLVENAKNFPQCDV
      190         200         210         220         230         240

Conf: 3456754320059988265988997893799978968999999986499879
Pred: CCCCCCCCCCCHHCCCCCEEECCCCCEEECCCHHHHHHHHHHCCCCC
AA: LEISDQEVLEWYTAKDFIVGKSLTILGRFFIYDCDPFTRRYKKEKFGITDL
      250         260         270         280         290
```

ANEXO 6

EFHC1C codon usage

TTT GGG TCT ACC ATG AAA GCT CCA AAA AAA GAC GTT ATT AAA ATG CTG GTG AAT GAT AAC AAG GTG CTT CGT
 TAT TTG GCT GTA CTG GAA TCC ACC ATC CCA GAA GAC AAA GAC CGC AGA TTT GTC TTC TCT TAC TTT CUA GCT
 ACC GAC ATG ATC AGT ATC TTT GAG CCT CCT GTT CGC AAT TCT GGT ATC ATT GGG GGC AAG TAC CTT GGC AGG
 ACT AAA GTT GTT AAA CCA TAC TCT ACA GTG GAC AAC CCT GTC TAC TAT GGC ACC AGT GAC TTC TTC ATT GGT
 GCT GTG ATT GAA GTG TTT GGT CAC CGG TTC ATC ATC CTT GAT ACA GAC GAG TAT GTT TTG AAA TAC ATG GAG
 AGC AAC GCT GCC CAG TAT TCA CCA GAA GCA CTC GCG TCA ATT CAG AAC CAT GTC CGA AAG CGA GAA GCG CCT
 GCT CCA GAA GCA GAA AGC AAG CAA ACT GAA AAG GAT CCA GGC GTG CAG GAA TTG GAA GCA TTA AUA GAC ACA
 ATT CAG AAG CAA CTG AAA GAT CAC TCA TGC AAA GAC AAC ATT CGT GAG GCA TTT CAA ATT TAT GAC AAG GAA
 GCT TCA GGA TAT GTG GAC AGA GAC ATG TTC TTT AAA ATC TGT GAA TCG CTT AAC GTC CCA GTG GAT GAC TCC
 TTG GTT AAG GAG TTA ATC AGG ATG TGC TCT CAT GGA GAA GGC AAA ATT AAC TAC TAT AAC TTT GTT CGT GCT
 TTC TCA AAC TGA CTC GAG GAT

The Number of Bases in the above Sequence = 741

The Number of Codons in the above Sequence = 247

Amino Acid	Rare Codon	Frequency of Occurrence
Arginine	CGA	2
	CGG	1
	AGG	2
	AGA	2
Glycine	GGA	2
	GGG	2
Isoleucine	AUA	1
Leucine	CUA	1
Proline	CCC	3
Threonine	ACC	0

ANEXO 7

PSIPRED PREDICTION RESULTS EFHC1C

Key

Conf: Confidence (0=low, 9=high)

Pred: Predicted secondary structure (H=helix, E=strand, C=coil)

AA: Target sequence

PSIPRED HFORMAT (PSIPRED V2.5 by David Jones)

```
Conf: 946246999998659669999999457876773589999874798299886376127782
Pred: CCCHHHHHHHHHHCCCEEEEEEEEECCCCCCCCCEEEEEEEEECCCCCCCCC
AA: KAPKKDVIKMLVNDNKVLRYLAVLESPIPEDKDRRFVFSYFLATDMISIFEPPVRNSGII
      10          20          30          40          50          60
```

```
Conf: 200122003468888756776538645302317898533999970008999998402266
Pred: CHHHHHHCCCCCCCCCCCCCHHHCCCCCEEEECCEEEEEEECHHHHHHHHHHCCCC
AA: GGKYLGRITKVVPYSTVDNPFVYGPSPDFFIGAVIEVFGRFIILDTDEYVLKYMESNAAQ
      70          80          90         100         110         120
```

```
Conf: 68777898998612221100113445432389801789999999997411437899999
Pred: CCHHHHHHHHHHCCCCCHHHHHHHHHHCCCCCHHHHHHHHHHHHCCCCCHHHHHHH
AA: YSPEALASIQNHVRKREAPAPEAESKQTEKDPGVQELEALIDTIQQLKDHSLKDNIREA
      130         140         150         160         170         180
```

```
Conf: 843287579985799999999981899998999999655889982979999998609
Pred: HHHCCCCCCCCCHHHHHHHHHHHHCCCCCHHHHHHHHHHCCCCCCCCCHHHHHHHHCC
AA: FQIYDKEASGYVDRDMFFKICESLNVPVDDSLVKELIRMSHGEKINYYNFVRAFSN
      190         200         210         220         230
```

ARTIGO PUBLICADO

“Identification and characterization of a proteolysis resistant fragment containing the PCI domain in the *Arabidopsis thaliana* INT6/eIF3e translation factor”

Marcelo Jun Murai, Carneiro, F. R. G., Gozzo, F. C., Ierardi, D. F.,
Pertinhez, T. A., and Zanchin, N. I. T.

Publicado na Revista
Cell Biochemistry and Biophysics

Dados produzidos no início da pós-graduação, sob orientação do Prof. Dr. Nilson Ivo Tonin Zanchin, do Laboratório Nacional de Luz Síncrotron (LNLS), antes de mudança de orientação. Posteriormente foi publicado na revista *Cell Biochemistry and Biophysics*.

ORIGINAL ARTICLE

Identification and Characterization of a Proteolysis-Resistant Fragment Containing the PCI Domain in the *Arabidopsis thaliana* INT6/eIF3e Translation Factor

Marcelo J. Murai, Flávia R. G. Carneiro, Fabio C. Gozzo, Daniela F. Ierardi, Thelma A. Pertinhez, and Nilson I. T. Zanchin*

Center for Structural Molecular Biology, Brazilian Synchrotron Light Laboratory-LNLS, Campinas-SP, Brazil

Abstract

The PCI domain comprises approx 200 amino acids and is found in subunits of the eukaryotic translation initiation factor 3 (eIF3), the 26S proteasome and the COP9/signalosome complexes. The PCI domain is involved in protein-protein interaction, and mouse INT6 truncated proteins lacking the PCI domain show cell malignant-transforming activity. In this work, the *Arabidopsis thaliana* INT6/eIF3e (AtINT6) protein was dissected using limited proteolysis, and a protease-resistant fragment containing the PCI domain was identified. Based on mass spectrometry analyses of the protease-resistant fragments and on secondary structure prediction, AtINT6-truncated proteins were cloned and expressed in *Escherichia coli*. Stability studies using thermal unfolding followed by circular dichroism revealed a midpoint transition temperature of 44°C for the full-length AtINT6 protein, whereas the truncated proteins comprising residues 125-415 (AtINT6TR2) and 172-415 (AtINT6TR3) showed transition temperatures of 49 and 58°C, respectively. AtINT6TR3 contains the PCI domain with additional amino acids at the N and C termini. It shows high solubility, and together with the high thermal stability, should facilitate further characterization of the PCI domain structure, which is important to understand its function in protein-protein interaction.

Index Entries: PCI domain; eIF3; INT6; limited proteolysis; circular dichroism.

INTRODUCTION

Sequence alignment identified the PCI domain in the subunits of three large protein complexes: the 26S proteasome, the plant photomorphogenic regulatory complex (COP9), and the translation initiation factor 3 complex (eIF3) (1). It has been predicted to be a 200-amino acid α -helical domain involved in protein-protein interaction (1,2).

INT6 is a PCI domain-containing protein first identified in mice, being encoded by a gene located in a frequent genome integration site of the mouse mammary tumor virus (MMTV) (3). Studies with truncated proteins generated by MMTV integration in the mouse

genome indicated that INT6 might be involved in the control of cell proliferation (4,5). Interestingly, in this case, these proteins did not contain the PCI domain. INT6 was subsequently shown to copurify with the eIF3 complex isolated from both reticulocyte lysates and human HeLa cells (6). In addition, it has been suggested to act in translational repression via interaction with the P56 interferon-induced protein (7,8). INT6 interacts also with the human type I T-cell leukemia virus Tax oncoprotein (9) and with the Ret finger protein 1 (10). INT6 interaction with the Tax leads to subcellular redistribution (9) that probably affects its function. Indeed, under normal conditions INT6 shows a cell cycle-dependent localization with a preference for the nuclear compartment (11). In *Arabidopsis thaliana*, the INT6 ortholog (AtINT6) has been found preferentially in the nucleus and interacts directly with the *c* subunit of eIF3 and

* Author to whom all correspondence and reprint requests should be addressed. E-mail: zanchin@lnls.br

with the CSN7 subunit of the COP9 signalosome (12). The COP9 complex is a suppressor of light-dependent development; nonetheless, equivalent complexes are found in eukaryotic organisms with one exception, the yeast *Saccharomyces cerevisiae*. Therefore, protein interaction studies in *A. thaliana* provided evidence that the nuclear role of INT6 is associated with the function of the COP9 signalosome.

INT6 has many interacting partners and its interaction activity has been assigned to the PCI domain that is found in the C-terminal half of the protein (2,12). However, it is unclear how this domain can recognize multiple proteins, and the characterization of its structure should provide relevant information about its mechanism of action. Recently, two nuclear magnetic resonance structures of partial PCI domains have been deposited in the Protein Data Bank (PDB) by S. Suzuki and coworkers for the mouse hypothetical protein Aah51541 and the COP9 complex subunit 4 (PDB accession no. 1WI9 and 1UFM, respectively). These partial domains correspond to approximately one-third of the domain identified by sequence alignment (1,2) and do not provide significant information about the PCI domain function. Characterization of the mammalian INT6 is hampered by the insolubility shown by the recombinant protein. These facts have prompted us to select the soluble *A. thaliana* INT6 protein, which shows 62% amino acid similarity to the human counterpart, for limited proteolysis and conformational studies. The protease-resistant fragments were identified by mass spectrometry (MS), and some truncated forms of AtINT6 related to these fragments and containing the PCI domain were constructed. They showed different expression efficiency, solubility, and thermal stability. Most importantly, the results presented in this work evidenced that we have been able to identify a soluble and highly stable AtINT6-truncated protein containing the PCI domain, thus opening the way to the characterization of the PCI structure.

MATERIALS AND METHODS

Plasmid Construction, Protein Expression, and Purification

The cDNA encoding *A. thaliana* INT6 (AtINT6) (National Center for Biotechnology Information accession no. AF285832) was amplified by reverse transcription followed by polymerase chain reaction (PCR) from total RNA by using oligonucleotides ONZ39 (5'-AAAAGGATCCCTAGCGAGTTGCTTGCGCCTG-3') and ONZ41 (5'-GAGAGGATCCCATATGGAGGAAAGCAAACAGAAC-3'). The PCR product was inserted into the *Nde*I and *Bam*HI sites of the expression vector pET28a (Novagen, Madison, WI), producing plasmid

pET-AtINT6. Three truncated forms of AtINT6 were subcloned into the *Nde*I and *Bam*HI sites of pET28a plasmid to express the PCI-containing region of AtINT6. The cDNA region containing amino acids 124–441 was amplified by PCR using oligonucleotides ONZ127 (5'-CCAGCATATGTACCAGATTGGTCCAGACC-3') and ONZ39 to produce vector pET-AtINT6TR1. Vector pET-AtINT6TR2 contains the cDNA region from amino acids 124 to 415, which was PCR-amplified using oligonucleotides ONZ127 and ONZ128 (5'-CTGGATCCTAGTTTATCAACTGCTCATGCAG-3'), and vector pET-AtINT6TR3 contains the cDNA region encoding amino acids 172–415, amplified by PCR using oligonucleotides ONZ235 (5'-GAGTCATATGTGGGGAAAGCTCGCATCTGA-3') and ONZ128. All vectors express proteins with a histidine-tag at the N terminus, and the inserts were verified by DNA sequencing.

Proteins were expressed using *E. coli* strain BL21(DE3) *slyD*⁻, kindly provided by Dr. Ryland Young (Texas A&M University, College Station, TX). Protein expression assays and cell extract preparation were performed as described previously (13). Full-length AtINT6 and AtINT6TR2 proteins were purified by metal-chelating chromatography followed by ion-exchange chromatography on Q-Sepharose (GE Healthcare, Little Chalfont, Buckinghamshire, UK). AtINT6TR3 was purified by a single metal-chelating affinity chromatography step. For affinity purification, the extract from 2 L of LB cultures was incubated with 400 μ L of Ni²⁺-NTA (QIAGEN, Valencia, CA) for 1 h at 4°C; transferred to a spin column, which was washed with 20 column volumes (*cv*) of buffer A (50 mM sodium phosphate, pH 7.2, 100 mM NaCl, 5% glycerol, 1.5 mM imidazole, and 0.5 mM phenylmethylsulfonyl fluoride [PMSF]); and eluted with a linear 0–200 mM imidazole gradient in 45 *cv* of buffer A. Ion-exchange chromatography was performed on a 5-mL Q-Sepharose column (GE Healthcare). The column was washed with 10 mL of a buffer 1 (10 mM sodium phosphate, pH 7.2, 20 mM NaCl, 5% glycerol, and 3.5 mM 2-mercaptoethanol) and eluted with 10 *cv* by using a 0–1000 mM NaCl gradient. Protein samples were concentrated using Ultrafree-4 centrifugal filter devices (Millipore, Billerica, MA), and protein concentration was determined by extinction coefficient at A₂₈₀.

Circular Dichroism Spectroscopy

Circular dichroism (CD) experiments were performed using a Jasco J-810 spectropolarimeter (Jasco, Tokyo, Japan) coupled to a Peltier Jasco PFD-425 system for temperature control. Protein samples were prepared at 4 μ M in 10 mM sodium phosphate buffer at pH 7.2. The spectra were collected using a 1-mm optical path cell in the spectral range of 190–260 nm, and a 1-mm bandwidth with step size of 0.5 nm and scan speed of 50 nm/min. For each measurement, the mean values of four spectra were taken to improve the signal-to-noise ratio. The ellip-

ticity was expressed as the mean residue molar ellipticity (θ) ($\text{deg}\cdot\text{cm}^2\cdot\text{dmol}^{-1}$). Thermal unfolding experiments were performed by following the loss of secondary structure at 222 nm in the range from 10 to 90°C.

Limited Proteolysis of AtINT6

Sequencing-grade proteases (trypsin, thermolysin, chymotrypsin, papain, and subtilysin) were purchased from Sigma (St. Louis, MO). One hundred-microliter limited proteolysis reactions were performed in 10 mM Tris, pH 7.4, 10 mM NaCl, and 3.5 mM 2-mercaptoethanol at room temperature for up to 4 h. The concentration of AtINT6 was 1 mg/mL, and the enzyme-to-substrate ratio used was 1:100 (w/w). At time intervals, aliquots were removed and boiled in sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) sample buffer to inactivate the proteases. For MS analyses, trypsin was inactivated with 1 mM PMSF. Time-course digestions were analyzed by SDS-PAGE, and the protease-resistant fragments were analyzed by MS.

Mass Spectrometry

For matrix-assisted laser desorption ionization/time of flight (MALDI-TOF) peptide fingerprinting, proteolysis reactions were fractionated on SDS-PAGE, and bands were subjected to in-gel trypsin digestion for 16 h for complete digestion. Samples were desalted with ZipTipC18 (Millipore), according to manufacturer's instructions. Subsequently, 0.5 μL of the sample was spotted with 0.5 μL of matrix solution on a stainless steel sample plate and allowed to dry at room temperature. α -Cyano-4-hydroxycinnamic acid (Sigma) at 10 mg/mL in water/acetonitrile (1:1, v/v) and 0.1% formic acid were used as matrix for ionization. Mass spectra were acquired on a Voyager-DE PRO MALDI-TOF (Applied Biosystems, Foster City, CA) spectrometer. MS using electrospray ionization was performed on a Waters/Micromass Q-ToF Ultima API spectrometer (Waters, Milford, MA) with a nanoflow interface. The digested sample (10 μL) was desalted using a Waters Opti-Pak C18 trap column for 5 min. The sample was eluted from the C18 trap column at a 250 nL/min flow with a mobile phase containing water/acetonitrile (1:1, v/v) and 0.1% formic acid. The instrument conditions were 3 kV for the spray voltage, 100 V for the cone voltage, cone gas at 30 L/h, and source temperature of 100°C. The final spectrum was processed using the MaxEnt1 program (Waters/Micromass).

RESULTS AND DISCUSSION

Recombinant AtINT6 Protein Contains a Protease-Resistant Domain

The *A. thaliana* INT6/eIF3e ortholog (AtINT6) is a 441-amino acid protein sharing 62% amino acid similar-

ity with the human INT6 protein (9). The human INT6 is insoluble when expressed in bacterial systems (data not shown); nonetheless, soluble AtINT6 can be obtained using *E. coli*. Histidine-tagged AtINT6 was purified by metal-chelating affinity and ion-exchange chromatography (Fig. 1, lane I). To determine whether the PCI domain defined by primary structure alignment (1,2) also corresponds to a structural domain, we conducted limited proteolysis experiments by using proteases showing different specificities (trypsin, thermolysin, chymotrypsin, papain, and subtilysin) to digest recombinant AtINT6 (Fig. 1). SDS-PAGE analyses of the protease-resistant products showed that chymotrypsin produced a major fragment of approx 30 kDa (Fig. 1A). Thermolysin and trypsin produced two well-defined bands of approx 25 and 30 kDa and 18 and 30 kDa, respectively (Fig. 1A,B). Although there is accumulation of an approx 30-kDa band, the sample digested with papain occurs as a smear of digested peptides (Fig. 1C). For subtilysin, only a band of approx 18 kDa was detectable at intermediate incubation times (Fig. 1D). That an approx 30-kDa band was generated by limited proteolysis with four of the five proteases tested (chymotrypsin, thermolysin, trypsin, and papain) indicated that the proteolysis-resistant fragment should correspond to the same region of AtINT6.

PCI Domain Is Part of the AtINT6 Protease-Resistant Product

To identify the amino acid sequence of the trypsin- and thermolysin-resistant fragments, we initially performed peptide fingerprinting by using MALDI-TOFMS. The peptides identified for the approx 30-kDa trypsin-resistant domain encompass the region from amino acid 167 to 418 (Fig. 2); however, several of the internal peptides could not be detected. Nevertheless, the predicted molecular mass of this region is 29,623.96 Da and is consistent with the size of the product observed on SDS-PAGE. The peptides identified for the approx 30- and 25-kDa thermolysin fragments contain amino acids 124–385 (predicted molecular mass of 31,056 Da) and amino acids 175–385 (predicted molecular mass of 25,136 Da), respectively. Although the two fragments showed sequence overlapping in the region from 175 to 385 and molecular masses consistent with the bands observed by SDS-PAGE, the peptide fingerprinting strategy turned out to be inconclusive, because several internal peptides could not be detected, thus raising doubts whether the N- and C-terminal ends of the proteolysis products were accurately determined. Therefore, the AtINT6 trypsin-resistant products also were analyzed by MS under electrospray ionization. Analysis with this technique identified a major product of 29,623 Da (Fig. 3), which includes amino acid residues from 167 to 418 because the predicted molecular mass for this region of AtINT6 is

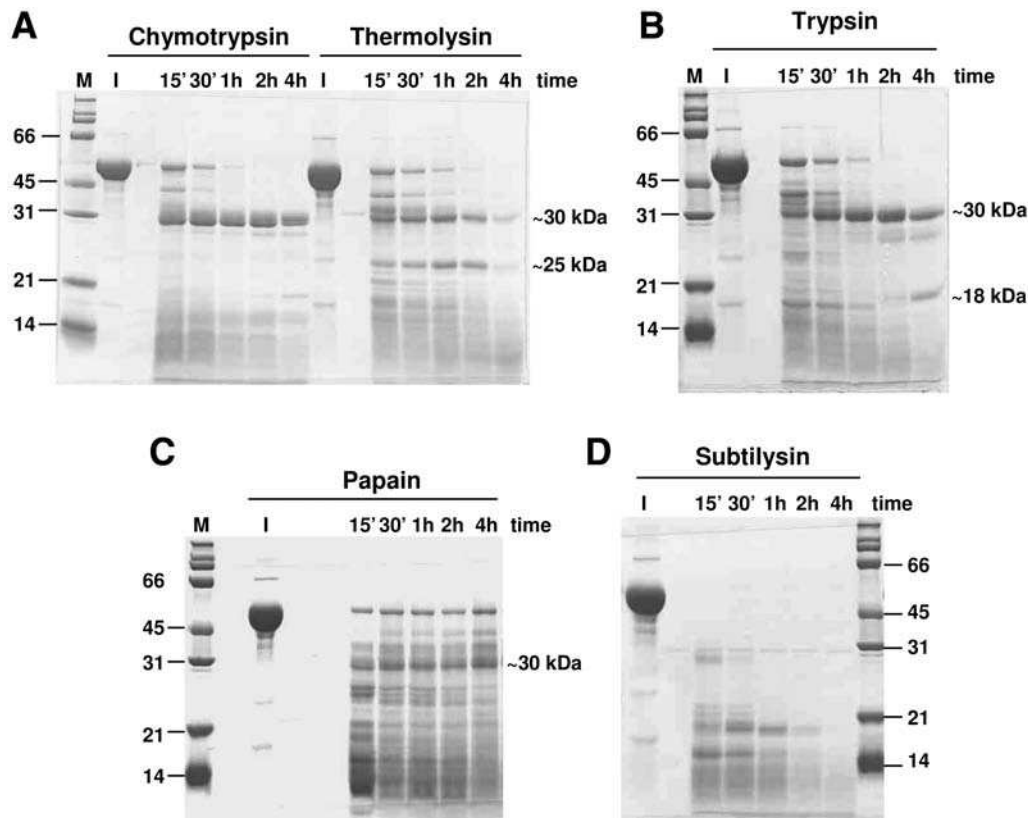


Fig. 1. SDS-PAGE analyses of AtINT6 partial proteolysis. The proteases used in the assays are indicated above each panel, and the incubation time is indicated above each lane. I, untreated AtINT6 protein. The molecular weight marker and the protease resistant products also are indicated.

29,623.96 Da and matches exactly the trypsin-resistant fragment identified by MALDI-TOF. The approx 30-kDa thermolysin product (encompassing residues 124–385) does not correspond exactly to the approx 30-kDa trypsin (encompassing residues 167–418), suggesting that the protease-resistant core fragment is located in the region between amino acids 167–175 and 385–418 (Fig. 2). Note that this region overlaps but does not match exactly the region identified by sequence alignment in the C-terminal half that contains approx 200 amino acids of the PCI domain and characterized by the higher sequence conservation (1,2). For AtINT6, the conserved C-terminal half includes amino acids 288–404 (Conserved Domain Database, www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=pfam01399). In conclusion, however, limited proteolysis analyses revealed that the resistant fragments include the PCI domain identified by sequence alignment with additional amino acids at the N terminus and possibly also at the C terminus.

Construction and Expression of Truncated Proteins Containing the PCI Domain

The finding that the trypsin and thermolysin proteolysis products of AtINT6 show similar molecular mass

but differ at the N- and C-terminal regions led us to construct vectors to express two truncated proteins with N and C terminus tentatively determined by using both the MS information and secondary structure prediction. Amino acid 124 corresponds to the N terminus of the thermolysin largest resistant product, and amino acid 418 corresponds to the C terminus of the trypsin largest fragment. The truncated protein AtINT6TR2 contains this region (Fig. 4D), although its C terminus was set at amino acid 415 instead of 418 to better fit the scores of secondary structure prediction (14), which indicate low scores for amino acids 417–418 to form α -helix (data not shown). Indeed, by removing these amino acids, we intended to generate a truncated protein without unstructured regions at the C terminus. In parallel, a second truncated protein, AtINT6TR1, containing amino acids 124–441 was constructed to obtain a protein with the native C terminus of AtINT6 (Fig. 4D), which we expected to positively affect its stability. Surprisingly, AtINT6TR1 was totally insoluble after expression in *E. coli* (Fig. 4A,B), suggesting that the 26 C-terminal amino acids negatively interfere with the folding of this truncated protein. Relatively high levels of soluble AtINT6TR2 were obtained at lower incuba-

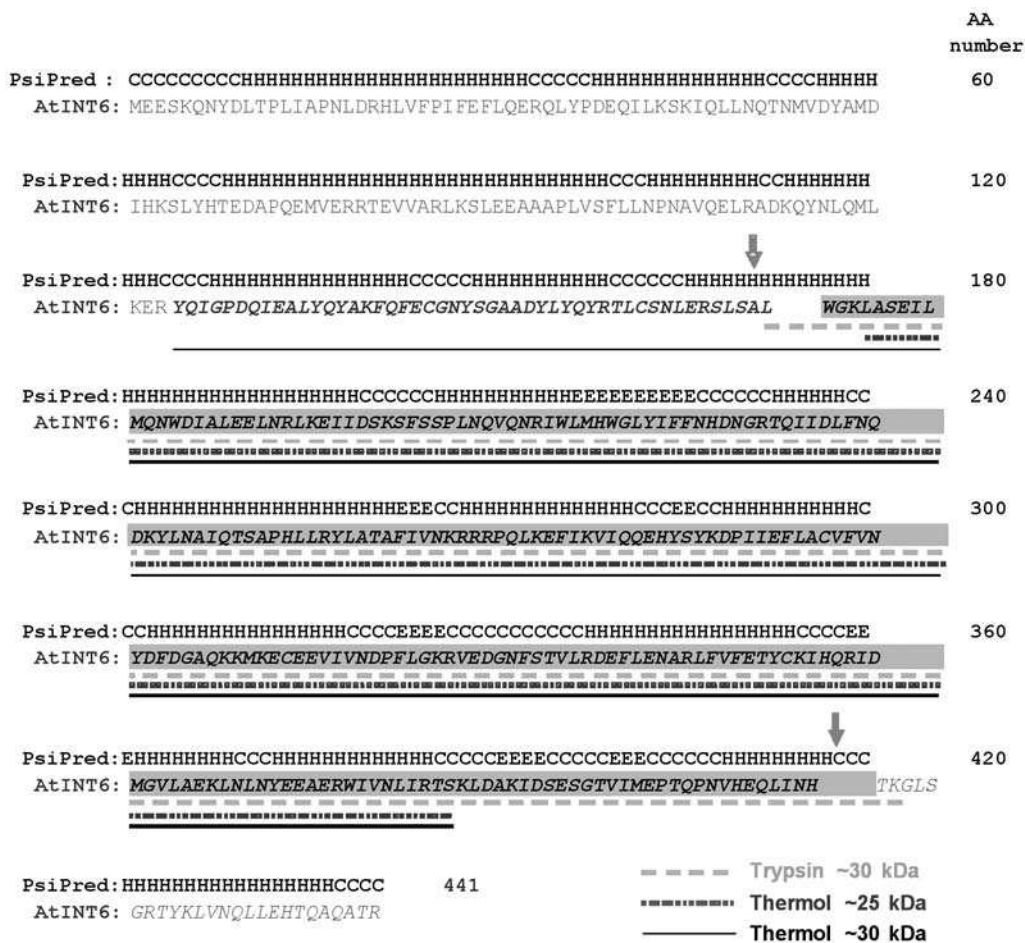


Fig. 2. AtINT6 amino acid sequence showing the protease-resistant fragments identified by MALDI-TOF fingerprinting and electrospray ionization MS. Lines underneath the amino acid sequence indicate the approx 30-kDa thermolysin-resistant fragment (solid line), the approx 25-kDa thermolysin-resistant fragment (dotted line), and the approx 30-kDa trypsin-resistant fragment (dashed line). Arrows indicate the fragment from residue 167 to 418 identified using electrospray ionization. The secondary structure (predicted using PSIPRED) (11) is shown above the amino acid sequence. H, α -helix; E, β -sheet; C, coil. The amino acid sequence corresponding to AtINT6TR1 is shown in italics, AtINT6TR2 is in bold, and AtINT6TR3 is marked with a gray box.

tion temperatures, such as 25°C (Fig. 4A), whereas incubation of *E. coli* cells above this temperature resulted in production of insoluble protein.

A third truncated form, AtINT6TR3, was constructed to adjust the N terminus more closely to the trypsin proteolytic product. This product contains amino acids 167–418 and was identified by both MALDI-TOF and electrospray ionization MS. The C terminus of AtINT6TR3 corresponds to the C terminus of AtINT6TR2, and the N terminus was set at amino acid 172 instead of 167 because, according to the secondary structure prediction using PSIPRED (14), amino acids 167–171 show a low score for α -helix formation (data not shown). Thus, as for the C terminus, we intended to remove all unstructured amino acids from the N terminus of AtINT6TR3 (Fig. 4D). As shown in Figure 4, this strategy was very effective to

generate a truncated protein showing high level of expression in *E. coli*. AtINT6TR3 also showed high solubility, being soluble even in cells incubated at higher temperatures, thus suggesting it to be more stable than AtINT6TR2. AtINT6TR3, made up of amino acids 172–415, contains the predicted PCI domain identified by bioinformatics and should serve as a model protein for the structural characterization of the PCI domain.

Conformational and Stability Analyses of AtINT6 and of Its Truncated Forms

The purified recombinant proteins (Fig. 4C) showed CD spectra with high helical content (Fig. 5), as predicted for the PCI domain and for the AtINT6 full-length protein (1,2). The CD spectra obtained for AtINT6 and AtINT6TR2 are very similar. Interestingly,

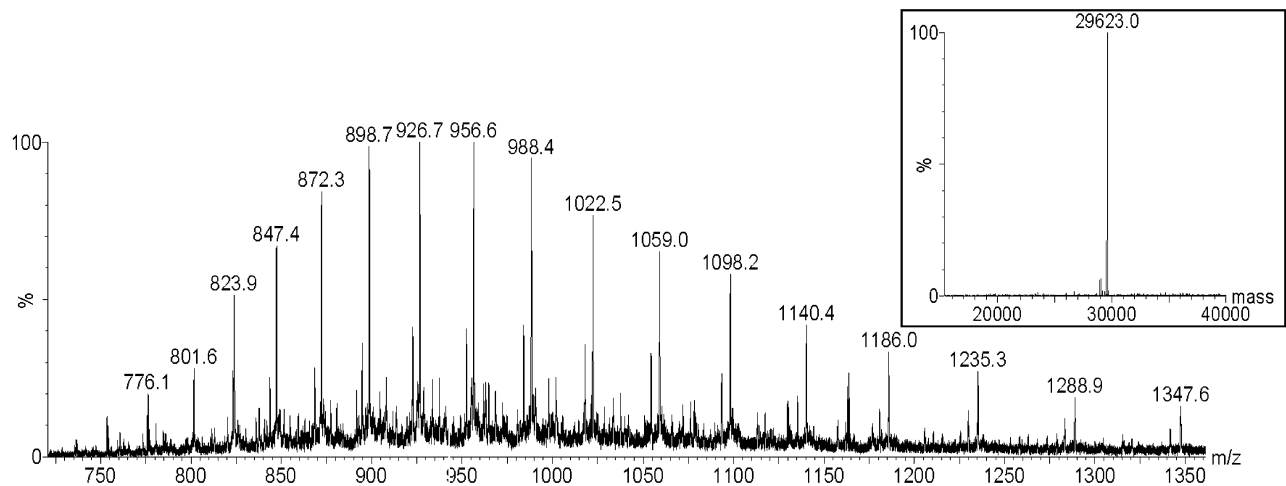


Fig. 3. Electrospray MS spectrum of AtINT6 trypsin-resistant domain. The inset shows the deconvoluted spectrum and the actual mass of the domain.

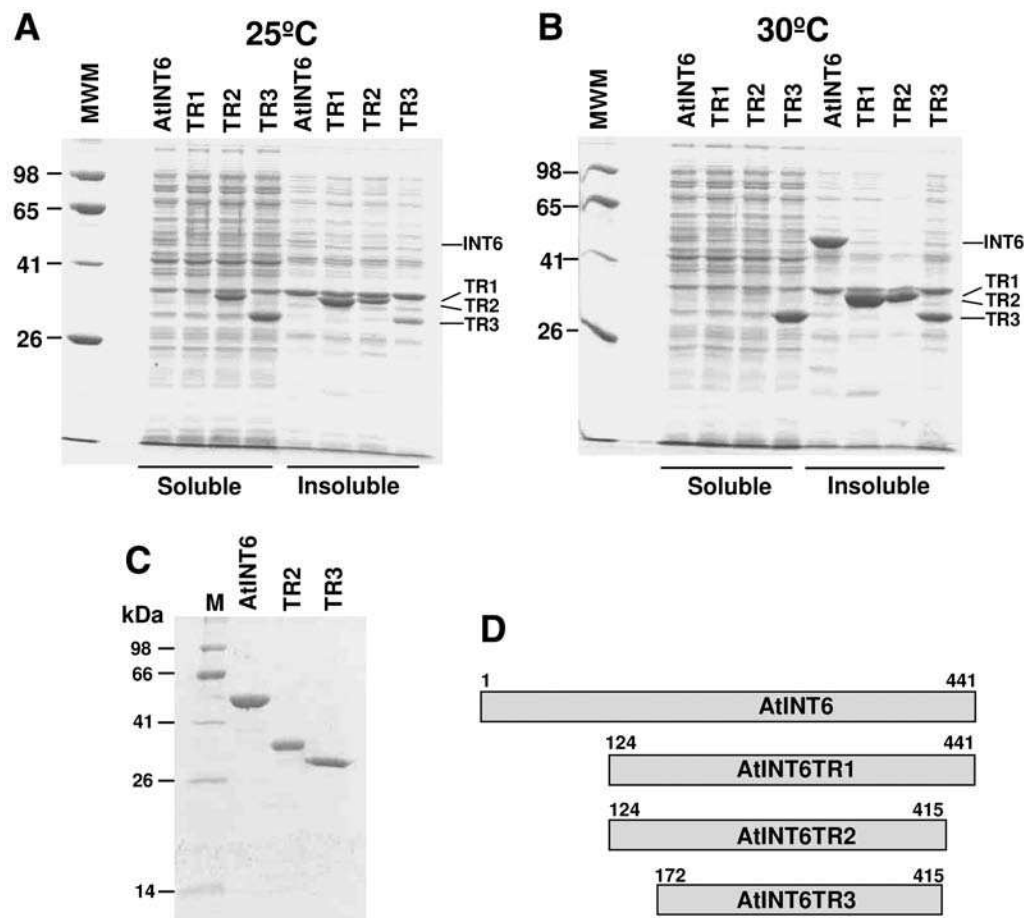


Fig. 4. Expression of the AtINT6 truncated forms in *E. coli*. (A and B) SDS-PAGE analyses of the recombinant proteins expressed in *E. coli* at 25 and 30°C, respectively. Soluble and insoluble fractions are indicated at the bottom of each panel. (C) SDS-PAGE gel showing the purified recombinant proteins. (D) Scheme representing the AtINT6 deletion mutants.

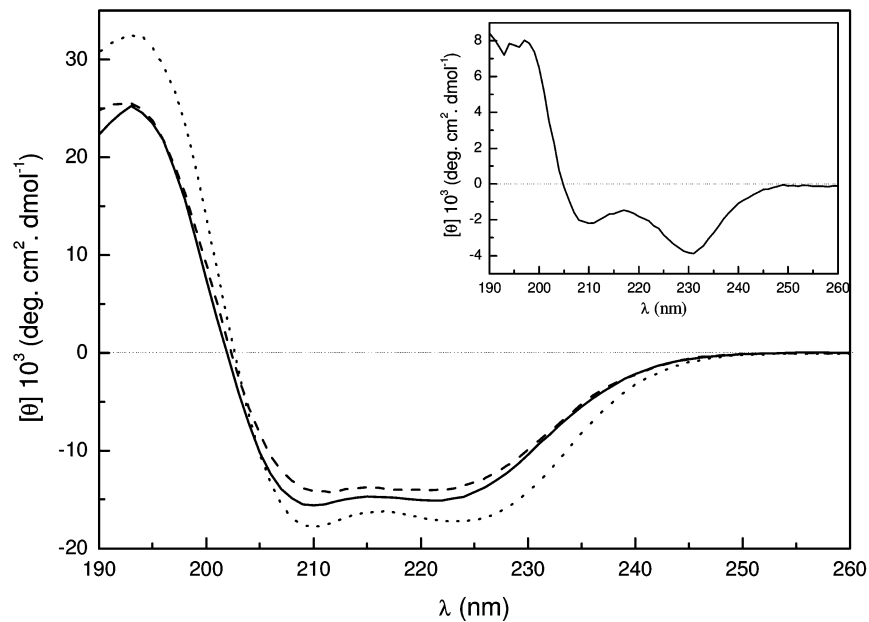


Fig. 5. Circular dichroism spectra of AtINT6 and of its truncated forms. Spectra of the full-length AtINT6 (solid line) and the truncated proteins AtINT6TR2 (dashed line) and AtINT6TR3 (dotted line). Inset, difference spectrum between AtINT6TR3 and AtINT6.

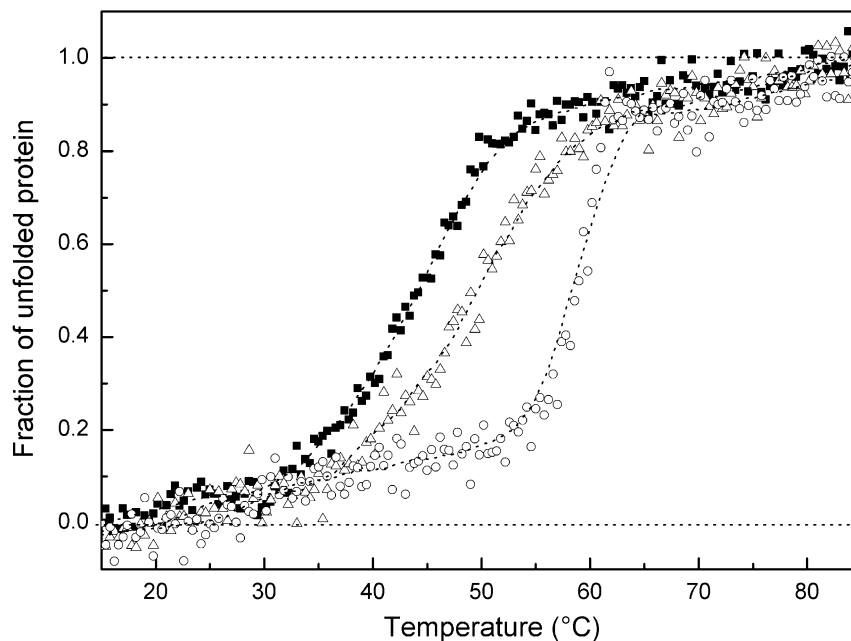


Fig. 6. Thermal denaturation of AtINT6 (■), AtINT6TR2 (△), and AtINT6TR3 (○) measured at 222 nm. The curves show the fraction of unfolded protein relative to increasing temperatures.

AtINT6TR3 exhibits a hyperoptical activity and the band at 220 nm shifted to 225 nm. Indeed, the AtINT6TR3–AtINT6 difference spectrum (Fig. 5, inset) reveals a negative band at 230 nm that can be assigned to the lower content of aromatic residues in AtINT6TR3

(28) compared with AtINT6 (46) and AtINT6TR2 (37). Indeed, aromatic residues show a positive contribution in this region of the CD spectrum (15). In addition, the positive and negative peaks at 195 and 208 nm, respectively, of the difference spectrum indicate that

AtINT6TR3 is characterized by a higher content of α -helix (Fig. 5, inset). However, the overlapping of the 222-nm contribution because of the α -helix with the aromatic contribution at 230 nm does not allow quantitative estimation.

The differences observed in the expression levels of the recombinant proteins indicate that folding and stability of AtINT6TR3 should be favored compared with AtINT6TR2 and full-length AtINT6. Thermal unfolding was followed by CD and showed that AtINT6 and AtINT6TR2 possess midpoint transition temperatures of 44 and 49°C (Fig. 6), respectively. AtINT6TR3, in contrast, exhibited a significantly higher midpoint transition temperature of 58°C, with a sharper transition from the folded to the unfolded state (Fig. 6), confirming the higher structure stability and a stronger cooperativity of the transition. After thermal denaturation at 90°C, refolding was found to be irreversible for the three proteins (data not shown). It is intriguing that the 26 amino acids of the C terminus have such a strong negative effect on the solubility of AtINT6TR1, and we have no explanation so far. The effect of deleting the sequence between amino acids 124 and 172 was striking, because AtINT6TR3 shows higher expression levels, solubility, and thermal stability than AtINT6TR2. In this case, however, we feel that these differences are due to the high content of hydrophobic residues of the 124–172 region that also includes two of the five cysteine residues of the protease resistant product (Fig. 2).

Characterizing the PCI domain structure is essential to understanding its function because the mouse INT6-truncated proteins lacking the PCI domain show cell-transforming activity (4,5). Furthermore, INT6 protein–protein interaction activity is mediated by the PCI domain region (2,12), and no clues are available to explain how the PCI domain can specifically recognize different binding partners. In conclusion, the combination of limited proteolysis, MS, and CD spectroscopy used in this work allowed us to identify a truncated protein showing high level of expression, and, more importantly, high solubility and stability. This truncated protein, AtINT6TR3, contains the PCI domain identified by bioinformatics with some additional amino acids at the N- and C-termini and has the potential to serve as a model protein for the three-dimensional structure determination of the PCI domain.

ACKNOWLEDGMENTS

We are grateful to Celso Eduardo Benedetti for providing *A. thaliana* mRNA; to Markus B. Smolka and José Camillo Novello for the peptide fingerprinting analysis; and to Patrícia R. Marques, Tereza C. Lima Silva,

Adriana C. Alves, Luciana Rodrigues Camillo, and Zildene Gonçalves Correa for technical support.

REFERENCES

- Hofmann, K. and Bucher, P. (1998) The PCI domain: a common theme in three multiprotein complexes. *Trends Biochem. Sci.* **23**, 204–205.
- Kim, T., Hofmann, K., von Arnim, A. G., and Chamovitz, D. A. (2001) PCI complexes: pretty complex interactions in diverse signaling pathways. *Trends Plant Sci.* **6**, 379–386.
- Marchetti, A., Buttitta, F., Miyazaki, S., Gallahan, D., Smith, G. H., and Callahan, R. (1995) Int-6, a highly conserved, widely expressed gene, is mutated by mouse mammary tumor virus in mammary preneoplasia. *J. Virol.* **69**, 1932–1938.
- Rasmussen, S. B., Kordon, E., Callahan, R., and Smith, G. H. (2001) Evidence for the transforming activity of a truncated Int6 gene, in vitro. *Oncogene* **20**, 5291–5301.
- Mayeur, G. L. and Hershey, J. W. (2002) Malignant transformation by the eukaryotic translation initiation factor 3 subunit p48 (eIF3e). *FEBS Lett.* **514**, 49–54.
- Asano, K., Merrick, W. C., and Hershey, J. W. (1997) The translation initiation factor eIF3-p48 subunit is encoded by int-6, a site of frequent integration by the mouse mammary tumor virus genome. *J. Biol. Chem.* **272**, 23,477–23,480.
- Guo, J. and Sen, G. C. (2000) Characterization of the interaction between the interferon-induced protein P56 and the Int6 protein encoded by a locus of insertion of the mouse mammary tumor virus. *J. Virol.* **74**, 1892–1899.
- Guo, J., Hui, D. J., Merrick, W. C., and Sen, G. C. (2000) A new pathway of translational regulation mediated by eukaryotic initiation factor 3. *EMBO J.* **19**, 6891–6899.
- Desbois, C., Rousset, R., Bantignies, F., and Jalinot, P. (1996) Exclusion of Int-6 from PML nuclear bodies by binding to the HTLV-I Tax oncoprotein. *Science* **273**, 951–953.
- Moris-Desbois, C., Bochar, V., Reynaud, C., and Jalinot P. (1999) Interaction between the Ret finger protein and the int-6 gene product and co-localization into nuclear bodies. *J. Cell Sci.* **112**, 3331–3342.
- Watkins S. J. and Norbury, C. J. (2004) Cell cycle-related variation in subcellular localization of eIF3e/INT6 in human fibroblasts. *Cell Prolif.* **37**, 149–160.
- Yahalom, A., Kim, T. H., Winter, E., Karniol, B., von Arnim, A. G., and Chamovitz, D. A. (2001) Arabidopsis eIF3e (INT-6) associates with both eIF3c and the COP9 signalosome subunit CSN7. *J. Biol. Chem.* **276**, 334–340.
- Ausubel F. M., Brent R, Kingston R. (1998) *Current Protocols in Molecular Biology*, New York: Wiley.
- McGuffin L. J., Bryson K, and Jones D. T. (2000) The PSIPRED protein structure prediction server. *Bioinformatics* **16**, 404–405.
- Sreerama, N., Manning, M. C., Poweres, M. E., Zhang, J.-X., Goldenberg, D. P., and Woody, R. W. (1999) Tyrosine, phenylalanine, and disulfide contributions to the circular dichroism of proteins: circular dichroism spectra of wild-type and mutant bovine pancreatic trypsin inhibitor. *Biochemistry* **38**, 10,814–10,822.

DECLARAÇÃO

Declaro para os devidos fins que o conteúdo de minha dissertação/ tese de mestrado/doutorado intitulada “Expressão e purificação de proteínas relacionadas à epilepsia”

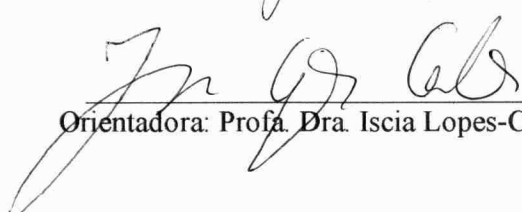
() não se enquadra no Artigo 1º, § 3º da Informação CCPG 002/06, referente a bioética e biossegurança.

() está inserido no Projeto CIBio (Protocolo nº _____), intitulado _____

() tem autorização da Comissão de Ética em Experimentação Animal (Protocolo nº _____).

(X) tem autorização do Comitê de Ética para Pesquisa com Seres Humanos (?) (Protocolo nº 470/03).


Aluno: Marcelo Jun Murai


Orientadora: Prof.ª Dra. Iscia Lopes-Cendes

Para uso da Comissão ou Comitê pertinente:

(X) Deferido () Indeferido

Nome:

Função:


Prof.ª Dra. Carmen Silvia Bertuzzo
Matrícula 17484-0
Presidente do Comitê de Ética em Pesquisa
Faculdade de Ciências Médicas - UNICAMP