

Ministério da Saúde
FUNDAÇÃO OSWALDO CRUZ
Centro de Pesquisas René Rachou
Programa de Pós-graduação em Ciências da Saúde

**Expressão da proteína recombinante RasGEF1b: um novo
fator de troca de nucleotídeos guanina associado à proteína
Ras induzido pelo *Trypanosoma cruzi***

por

Tatiana Nunes Silveira

**Dissertação apresentada com vistas à obtenção do Título de Mestre em Ciências
da Saúde na área de Biologia Celular e Molecular**

Orientadores:

Ricardo T. Gazzinelli (Laboratório de Imunopatologia)

Alvaro J. Romanha (Laboratório de Parasitologia Celular e Molecular)

**Expressão da proteína recombinante RasGEF1b: um novo
fator de troca de nucleotídeos guanina associado à proteína
Ras induzido pelo *Trypanosoma cruzi***

por

Tatiana Nunes Silveira

**Dissertação submetida à qualificação por banca examinadora em 6 de junho de
2006.**

Examinadores:

Guilherme C. Oliveira (CPqRR-Fiocruz)

Ronaldo Nagem (ICB-UFMG)

"Juro, pela minha fé pela minha honra e de acordo com os princípios éticos do biólogo, exercer as minhas atividades profissionais com honestidade, em defesa da vida, estimulando o desenvolvimento científico, tecnológico e humanístico com justiça e paz."

Juramento Oficial do Biólogo

Agradecimentos

Muito obrigada à...

Ricardo T. Gazzinelli

Rosiane A. Silva

Santuza M. Teixeira

Lucimara Chiari

Fernanda Caldas

Aristóbolo Mendes da Silva

Hélida M. Andrade

Ronaldo Nagem

Warrison Andrade

Henk

Jonas Enrique Perales Aguilar

Alvaro J. Romanha

Guilherme C. Oliveira

Suzana Garcia

Jamil Silvano

Daniela Ribeiro

Eneida Paganini

Laboratório de Imunologia e Bioquímica de Parasitas (ICB-UFMG)

Laboratório de Imunoparasitologia (ICB-UFMG)

Laboratório de Parasitologia Celular e Molecular (CPqRR)

Laboratório de Imunopatologia (CPqRR)

Funcionários do René Rachou, especialmente Jaci

Colegas do Curso

Amigos e familiares

Este trabalho foi realizado com auxílio financeiro da **CAPES** (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior)

Resumo

rasgef1b é um fator de troca de nucleotídeos guanina (GEF) hipotético e altamente conservado. Esse gene contém um domínio RASGEFN com um motivo zíper de leucina e um domínio RASGEF com três sítios de localização nuclear. A expressão do mRNA do *rasgef1b* em macrófagos é induzida por diferentes agonistas de receptores do tipo Toll (TLRs), tais como LPS (TLR4), GPI-mucina (TLR2) e Poli I:C (TLR3). A fim de expressar a proteína recombinante, nós clonamos o cDNA do *rasgef1b* no vetor pQE-30, utilizado para transformar bactérias *E. coli* (XL1-Blue). A His-RasGEF1b expressa em bactérias foi purificada utilizando tampão com alto conteúdo de uréia, seguido por cromatografia de afinidade utilizando resina carregada com níquel. A expressão da proteína foi confirmada em gel 2D, análise por espectrometria de massa e Western blotting utilizando um anticorpo monoclonal anti-His. Além disso, o cDNA do *rasgef1b* foi inserido em um plasmídeo que permite a fusão do epítipo FLAG (pFLAGCMV2) com a proteína RasGEF1b. O pFLAGCMV2 codificando o cDNA do *rasgef1b* foi utilizado para transfectar células HEK293T e a expressão protéica foi avaliada com gel 2D gel e Western blotting utilizando um anticorpo monoclonal anti-FLAG. A proteína reconhecida pelo anti-FLAG mostrou um peso molecular aparente de 56 kDa e ponto isoelétrico de 7,25. Utilizando a técnica de centrifugação diferencial e anticorpos monoclonais para as proteínas específicas das diferentes frações subcelulares, nós demonstramos que a proteína FLAG-RasGEF1b estava presente principalmente nas frações de núcleo e de membranas pesadas. Através do alinhamento entre seqüências da RasGEF1b de diferentes espécies filogeneticamente distintas vimos que esta proteína é altamente conservada com similaridade de 57% a 97% em relação à RasGEF1b murina.

Abstract

rasgef1b is a highly conserved putative Guanine Exchange Factor (GEF) containing a RASGEFN domain with a leucine zipper and a RASGEF domain with three sites of nuclear localization. In macrophages expression of *rasgef1b* mRNA is induced by different TLR agonists, such as LPS (TLR4), GPI-mucin (TLR2) and Poli I:C (TLR3). In order to express the recombinant protein, we cloned the *rasgef1b* cDNA into a pQE30, used to transform *E. coli* (XL1Blue). The recombinant His-RasGEF1b expressed in bacteria was purified using high urea content buffer, followed affinity chromatography using a nickel resin. The expression of the His-RasGEF1b was confirmed in 2D gel, mass fingerprint analysis and Western blot using an anti-HIS mAb. In addition, the *rasgef1b* cDNA was inserted into a plasmid, which allows the insertion of the FLAG epitope (pFLAGCMV2) with the recombinant protein. The pFLAGCMV2 encoding the *rasgef1b* cDNA was used to transfect HEK293T cells and protein expression evaluated with 2D gel and Western blot using an anti-FLAG mAb. The protein recognized by the anti-FLAG showed an apparent molecular weight of 56 kDa and the predicted isoelectric point of 7.25. By using differential centrifugation, and mAbs specific proteins of different cellular fractions, we demonstrated that the protein FLAG-RasGEF1b was mainly present in the nuclear and heavy membrane fractions. Through the alignment among sequences of the RasGEF1b from phylogenetically distinct species we observed that this protein is highly conserved with similarity between 57% and 97% with the murine RasGEF1b.

Abreviaturas

Å - Angstrom

BLAST - *Basic Local Alignment Search Tool*

cDNA- DNA complementar

DDRT-PCR - *differential display reverse transcriptase-polymerase chain reaction*

DEPC - dietil-pirocarbonato

DNA – ácido desoxirribonucléico

dNTPs – 2'-desoxirribonucleotídeo 5'-trifosfato

EDTA - ácido etilenodiaminotetracético

ELISA – *enzyme linked immunosorbent assay*

GPI - glicosilfosfatidilinositol

h - hora

IFN- γ – interferon gamma

IL - interleucina

IPTG - isopropil- β -galactosídeo

kb - quilobase

kDa- quiloDálton

LB - meio de cultura Luria Bertani

LPS – lipopolissacáride derivado de bactéria gram-negativa

M - molar

min - minuto

μ g - micrograma

μ L - microlitro

mM - milimolar

mRNA - RNA mensageiro

NK - natural killer

nm - nanômetro

OD₆₀₀ - densidade óptica a 600 nm

PAMPs - *pathogen-associated molecular patterns*

pb – pares de bases

PCR – *polimerase chain reaction*, reação em cadeia da polimerase

Poli I:C - ácido poli-inosínico-poli-citidílico

PVDF - polyvinylidene fluoride

RIN - radicais intermediários de nitrogênio

RNA - ácido ribonucléico

RNB - *Reverse Northern Blot*

rpm – rotações por minuto

SDS - duodecil sulfato de sódio

SDS- PAGE- gel de poliacrilamida contendo duodecil sulfato de sódio

Strips - tiras de gel de poliacrilamida com gradiente de pH ao longo de sua extensão

Spot - ponto ou mancha correspondente a uma ou mais proteínas em gel bidimensional

Tris - tris-(hidroximetil)-aminometano

TLR - *toll like receptor*, receptor do tipo Toll

TNF – fator de necrose tumoral

U/ μ L - unidades de enzima por microlitro

U V - ultra-violeta

V - volts

Índice

1. Introdução.....	11
1.1. A Doença de Chagas e o <i>Trypanosoma cruzi</i>	11
1.2. Imunopatogênese da Doença de Chagas	12
1.3. Imunidade Inata e o <i>Trypanosoma cruzi</i>	13
1.4. RasGEFs.....	15
2. Justificativa.....	18
3. Objetivos	19
3.1. Objetivo Geral	19
3.2. Objetivos Específicos.....	19
4. Materiais e métodos	20
4.1. Análise da proteína RasGEF1b utilizando ferramentas de bioinformática	20
4.2. Extração do RNA total de células RAW 264.7 estimuladas com RNA de fita-dupla ou endotoxina bacteriana.....	20
4.3. RT-PCR para detecção da expressão do cDNA do <i>rasgef1b</i>	21
4.4. Amplificação da região codificadora do gene <i>rasgef1b</i>	21
4.6. Clonagem no vetor procariótico pQE-30 (Qiagen)	24
4.7. Sequenciamento do cDNA do <i>rasgef1b</i> clonado no pQE-30.....	26
4.8. Expressão da proteína recombinante em bactérias.....	27
4.9. <i>Western Blotting</i> com anticorpo anti-His	28
4.10. Determinação da solubilidade da proteína recombinante.....	28
4.11. Purificação da proteína recombinante His-RasGEF1b.....	29
4.12. Eletroforese bidimensional.....	30
4.13. Microcaracterização dos “spots” retirados de gel bidimensional.....	31
4.14. Imunização dos coelhos e galinhas com a proteína recombinante His-RasGEF1b	31
4.15. Extração dos anticorpos de coelhos e galinhas	32
4.16. ELISA com soro dos coelhos e anticorpos de galinhas.....	33
4.17. Transfecção das células HEK 293T com o plasmídeo pFLAG contendo a região codificadora do <i>rasgef1b</i>	33
4.18. Fracionamento subcelular.....	35

5. Resultados	36
5.1. Análise da proteína RasGEF1b utilizando ferramentas de bioinformática	36
5.2. Expressão do mRNA do <i>rasgef1b</i> em células Raw 264.7 estimuladas com poli I:C ou LPS	55
5.3. Amplificação do cDNA do <i>rasgef1b</i> de macrófagos murinos	56
5.4. Perfil de restrição do cDNA do <i>rasgef1b</i> clonado no vetor TOPO 2.1	57
5.5. Clonagem do gene <i>rasgef1b</i> no vetor pQE-30 e sequenciamento do produto gerado	58
5.6. Expressão da proteína RasGEF1b recombinante em sistema procarioto	61
5.7. Solubilização da proteína His-RasGEF1b recombinante produzida em <i>E. coli</i>	63
5.8. Purificação da His-RasGEF1b recombinante por cromatografia de afinidade	64
5.9. Eletroforese bidimensional e análise por espectrometria de massa da proteína His-RasGEF1b purificada	67
5.10. Produção de anticorpos anti-His-RasGEF1b em coelho e galinha	71
5.11. Localização subcelular da proteína FLAG-RasGEF1b em sistema eucarioto através de centrifugação diferencial e eletroforese bidimensional	72
6. Discussão	75
7. Perspectivas	79
8. Conclusão	80
9. Dados suplementares	81
10. Endereços Eletrônicos	84
11. Referências	85

1. Introdução

1.1. A Doença de Chagas e o *Trypanosoma cruzi*

Tripanosomíase Americana ou Doença de Chagas é uma parasitose encontrada em mamíferos do continente Americano, causado pelo protozoário *Trypanosoma cruzi* (Mastigophora, Kinetoplastidae). Dados sorológicos indicam que entre 16 a 18 milhões de pessoas estão infectadas com o agente causador, *T. cruzi*, e cerca de 100 milhões de pessoas consideradas em risco. A Doença de Chagas é uma das mais sérias parasitoses da América Latina, e, considerando-se que esta doença afeta principalmente indivíduos em idade produtiva, ela tem grande impacto negativo no desenvolvimento econômico deste continente (1).

O ciclo de vida do parasita *T. cruzi* é heteroxênico, e o parasita também é encontrado em triatomíneos hematófagos, os quais são responsáveis pela transmissão natural aos humanos e outros mamíferos (2). Os epimastigotas replicam-se extracelularmente no intestino médio do hospedeiro invertebrado. Ao chegarem ao reto, eles se diferenciam para o estágio infectivo, o de tripomastigotas metacíclicos, os quais são eliminados nas fezes durante a hematofagia do barbeiro no hospedeiro vertebrado. Tripomastigotas metacíclicos circulantes na corrente sanguínea podem infectar qualquer célula nucleada, diferenciando-se em amastigotas e dando início à divisão binária. Quando as células tornam-se cheias de parasitas, os amastigotas diferenciam-se em tripomastigotas sanguíneos e são liberados para o ambiente extracelular e podem infectar outras células nucleadas, transformando-se novamente em amastigotas e continuando a replicação. Uma vez na corrente sanguínea, os tripomastigotas vão disseminar a infecção para outros órgãos do hospedeiro vertebrado. Alternativamente, os tripomastigotas sanguíneos podem infectar o aparelho bucal sugador do vetor, transformarem-se em epimastigotas e, então, perpetuarem o ciclo de vida do *T. cruzi* (2). A forma mais comum de infecção em humanos é pelo contato com fezes contaminadas pelo *T. cruzi*, excretadas pelos insetos triatomíneos (Reduviidae, Hemiptera), popularmente denominados “barbeiros”. Esta doença também pode ser transmitida por transfusão de sangue, transplante de órgãos, via placentária, ou acidente em laboratórios (3).

1.2. Imunopatogênese da Doença de Chagas

No hospedeiro vertebrado, a infecção por *T. cruzi* tem início com uma fase aguda curta, caracterizada por alta parasitemia, seguida de uma fase indeterminada assintomática e uma fase crônica longa mantida por poucos parasitas (4 e 5). A fase aguda inicia-se através das manifestações locais, quando o *T. cruzi* penetra na conjuntiva (sinal de Romana) ou na pele (chagoma de inoculação). Estas lesões aparecem em 50% dos casos agudos dentro de 4 a 10 dias após a picada do barbeiro, regredindo em um ou dois meses. Concomitantemente, os linfonodos satélites são comprometidos e no conjunto forma-se o complexo cutâneo e/ou conjuntivo linfonodal. Esta fase é acompanhada de parasitemia e parasitismo tecidual elevados (6).

A fase indeterminada é definida como o período clinicamente silencioso que segue a fase aguda. Os indivíduos apresentam evidências sorológicas e/ou parasitológicas de infecção, mas permanecem assintomáticos e não exibem sinais eletrocardiográficos de envolvimento do coração e nem anormalidades em raios-X do trato digestivo. Estudos ultraestruturais têm demonstrado que as células inflamatórias presentes são em sua maioria linfócitos e macrófagos, e poucos polimorfonucleares (7).

A miocardite crônica e progressiva e a falência cardíaca, as quais estão associadas com aumento do tamanho do coração, degeneração de miócitos, fibrose intersticial severa e o espessamento das membranas basais caracterizam o quadro patológico da fase crônica cardíaca da Doença de Chagas. Sem dúvida alguma, na fase crônica da doença com suas formas clínicas sintomáticas o fato anatômico mais importante na patogenia da doença é o surgimento de uma fibrose difusa nem sempre relacionada com o foco inflamatório (8).

1.3. Imunidade Inata e o *Trypanosoma cruzi*

Vários estudos demonstram que antes do estabelecimento da imunidade adquirida o sistema imune inato desempenha um papel crucial para a resistência do hospedeiro durante a fase aguda da infecção com patógenos intracelulares (9-12). Dentro de poucas horas ou dias depois da infecção, o sistema imune inato pode ser ativado, através de receptores específicos, por estruturas características de certos patógenos (por exemplo, lipopolissacarídeo e estruturas repetitivas de manose) (13-18). Macrófagos expostos a tais componentes microbianos sintetizam altos níveis de citocinas com múltiplas atividades no sistema imune. Estas citocinas pró-inflamatórias ativam várias funções do sistema imune inato que são essenciais para a resistência do hospedeiro aos patógenos antes do desenvolvimento de uma resposta imune específica (9-12).

Os macrófagos expostos a protozoários parasitas produzem interleucina (IL)-12 que é responsável por iniciar a síntese de interferon- γ (IFN- γ) pelas células “natural killers” (NK) (9-12,16). Uma vez estimulados com IFN- γ , os macrófagos tornam-se altamente responsivos a estimulação por produtos de parasitas, um evento que leva ao aumento da síntese de quimiocinas e citocinas pro-inflamatórias. Quando expostos a parasitas vivos, os macrófagos estimulados com IFN- γ desempenham uma função efetora através da produção de altos níveis de RIN (radicais intermediários de nitrogênio), os quais são altamente tóxicos aos diferentes protozoários, tais como o *T. cruzi* (19).

Nos últimos anos, a identificação e a caracterização de moléculas de protozoários parasitas que desencadeiam a produção de citocinas e RIN pelas células do sistema imune inato têm sido alvo de estudo por vários grupos de pesquisa. Diferentes estudos têm relatado as atividades imunoestimulatórias e regulatórias de âncoras de glicosilfosfatidilinositol (GPI) derivadas das membranas de protozoários, tais como *Plasmodium falciparum*, *Trypanosoma brucei* e *T. cruzi* (20-21).

Em estudos realizados em nosso laboratório, foi demonstrado que a ativação de macrófagos pelo *T. cruzi* é estágio específica. Assim, apenas a forma tripomastigota sanguínea (ou obtida de tecido de cultura) ou a forma amastigota são capazes de induzir a síntese de citocinas e RIN por macrófagos. Baixa atividade foi observada nas formas epimastigotas e tripomastigotas metacíclicas

(22). Além disso, a âncora de GPI foi identificada como a porção ativa das GPI-mucinas dos tripomastigotas responsável por desencadear as diferentes funções dos macrófagos. Mais precisamente, foi demonstrado em nosso laboratório, que as âncoras GPI, altamente purificadas de GPI mucinas, ativam o Toll-Like-Receptor-2 (TLR-2) em concentrações picomolares (23-25). Dentre as citocinas que os macrófagos produzem sob o estímulo das GPI-mucinas, das formas tripomastigotas de cultura, estão TNF- α , IL-12, IL-6, IL-1 (22) e citocinas quimiotáticas como RANTES, KC, MIP2 e IP-10 (26).

Para melhor entendimento sobre a ativação do sistema imune inato pelo *T. cruzi*, nosso laboratório tem estudado a indução e a expressão de genes por macrófagos expostos à forma tripomastigota do parasita. Devido ao fato de as GPI-mucinas serem potentes estimuladoras dos macrófagos, e estas células serem uma das principais componentes no processo inflamatório no coração chagásico (26-29), também foi investigado, no nosso laboratório, o padrão de expressão gênica em macrófagos ativados por glicolipídios de *T. cruzi*. Vários RNA mensageiros de macrófagos inflamatórios cultivados na ausência ou presença de GPI-mucinas e/ou IFN- γ foram comparados pela técnica de *differential display reverse transcriptase-polymerase chain reaction* (DDRT-PCR) (30,31). Utilizando a técnica de Northern blot reverso (RNB) (32), foi confirmada a expressão diferencial de alguns genes em macrófagos estimulados com GPI-mucinas e IFN- γ e no tecido cardíaco de camundongos experimentalmente infectados com *T. cruzi*. Entre os transcritos diferencialmente expressos, foi encontrado um novo gene, de cópia única, altamente conservado nos genomas humano e murino, que codifica uma proteína hipotética com homologia entre genes da família dos Fatores de Troca de Nucleotídeos Guanina (GEFS) (33-35). Este gene, *rasgef1b*, inicialmente denominado *gpi γ 4*, codifica uma proteína hipotética que contém dois domínios relacionados a RasGEFN e RasGEF, domínios comuns à família dos GEFs, envolvidos na ativação de membros da família das Ras-GTPases (36).

1.4. RasGEFs

Os proto-oncogenes *ras* codificam GTPases de membrana que desempenham um papel central na transdução de sinais de fatores de crescimento e diferenciação, antígenos imunes e componentes da matriz extracelular. Proteínas da superfamília Ras e proteínas regulatórias relacionadas têm sido altamente conservadas no curso evolutivo dos eucariotas (37).

As proteínas GTPase pequenas (small GTPase proteins), assim chamadas devido ao seu baixo peso molecular (20–35 kDa), são evolucionariamente conservadas e encontradas entre vários reinos de organismos. A primeira GTPase pequena a ser isolada foi H-Ras e as pequenas GTPases têm sido agrupadas na superfamília Ras. Baseando-se na seqüência de aminoácidos das GTPases pequenas de mamíferos, estas podem ser subdivididas em subfamílias Ras, Rho, Arf, Rab, Ran e Rad/Gem. A subfamília Ras inclui H-Ras, K-Ras (4A e 4B), N-Ras, E-Ras, R-Ras, TC21/R-Ras2, M-Ras/R-Ras3, RalA, RalB, Rap1A, Rap1B, Rap2A, Rap2B, Rap2C, Rit1, Rit2, RheB e RhebL1. Os membros da família Ras melhores caracterizados são H-Ras, K-Ras e N-Ras, os quais regulam uma variedade de processos biológicos que contribuem para a progressão do ciclo celular e, em condições patofisiológicas, ao câncer (38).

Assim como as proteínas G heterodiméricas, as GTPases pequenas agem como “trocadoras moleculares” ciclando entre duas formas interconversíveis: inativa, quando ligadas à guanosina difosfato (GDP) e ativas, ligadas a guanosina trifosfato (GTP). A conversão entre os estados GTP e GDP-ligados é firmemente regulada por dois tipos de proteínas: fatores de troca de nucleotídeos guanina (GEFs) e as proteínas ativadoras de GTPase (GAPs) (39). Um terceiro componente, os inibidores da dissociação de guanina (GDI), pode inibir ou bloquear as ações das GAPs e dos GEFs (40).

Sistemas compostos de GTPases, GAPs e GEFs permitem grande versatilidade na construção de vias de sinalização. Os sinais podem ser amplificados (um GEF pode ativar várias GTPases), integrados (várias vias ativam os mesmos GEF e GAP, e o comportamento de uma GTPase depende do efeito total de todos seus GEFs e GAPs), ou dividido (uma GTPase induz vários efeitos). Essa versatilidade permite que pequenas GTPases mediem diferentes funções biológicas entre diferentes organismos (41). A localização intracelular das Ras-GTPases pode ainda

ser controlada pelas frenal-transferases e outras enzimas que regulam o ancoramento das Ras-GTPases às membranas celulares (40).

A atividade GTPase de cada proteína G pequena é variável mas relativamente muito lenta e é estimulada por GAPs. A maioria das GAPs, tais como a GAP de Ras e a GAP Rab3, são específicas para cada membro ou subfamília das proteínas G pequenas (39, 42, 43), mas algumas GAPs, tais como p190, uma GAP ativa nas proteínas Rho/Rac/Cdc42, mostram uma especificidade mais ampla de substrato (44).

Proteínas GEF causam ativação por catalisarem a troca de GDP por GTP, enquanto que as GAPs inativam a Ras aumentando a taxa de hidrólise de GTP. O GEF interage primeiramente com a forma GDP-ligada para formar um complexo binário de uma proteína G pequena-GEF. Então, o GEF, neste complexo, é substituído pelo GTP para originar a forma GTP-ligada. A maioria dos GEFs, tal como o Son of Sevenless (Sos) e Rab3GEF, são específicos para cada membro ou subfamília das proteínas G pequenas (39, 45, 46), mas alguns GEFs, tal como Dbl, um GEF ativo nas proteínas Rho/Rac/Cdc42, mostram maior especificidade de substrato (47 e 48).

Vários GEFs associados aos membros da subfamília Ras foram isolados e caracterizados em células de mamíferos como Sos, RasGRF, RasGRF2, RasGRP (RasGRP1), C3Gm, Vav, entre outros (49-53). Em células de mamíferos, o primeiro RasGEF a ser isolado foi designado RasGRF ou Cdc25Mn (54-56). Cerca de vinte novos GEFs ativadores dos membros da subfamília Ras (RasGEFs) foram identificados e caracterizados nos últimos anos, com funções de ativação GTPases específicas (57).

Os RasGEFs e as proteínas relacionadas aos RasGEFs são agrupados com base na homologia de domínios conservados presentes nas suas estruturas. O segmento de homologia ao RasGEF Cdc25 de *S. cerevisiae* é o domínio que determina os RasGEFs serem agrupados desta maneira. O domínio GEF é essencial para que os RasGEFs desempenhem a sua função catalítica (58). Além do domínio catalítico Cdc25, alguns RasGEFs possuem ainda outros domínios importantes e funcionais. Por exemplo, o RasGRF1, o Sos1 e o Sos2 possuem os domínios de homologia Dbl (DH, *Dbl homology*) e plequistrina (PH, *pleckstrin homology*) que estão arrançados em conjunto. O domínio DH parece também promover a troca de guaninas em GTPases da subfamília Rho, enquanto o domínio PH participa nas interações proteína-proteína ou proteína-lipídios (59,60).

Pouco se sabe sobre os membros RasGEF1a, RasGEF1b e RasGEF1c. A proteína GEF codificada pelo gene *rasgefa* é essencial para a agregação celular, agindo em nível de ativação da adenil-ciclase em *Dictyostelium*. Células de *Dictyostelium* sem o gene *rasgefb* apresentam desenvolvimento inicial prejudicado, embora elas eventualmente formem corpos de frutificação finos, mas normalmente proporcionais. Além disso, essas células se movem rapidamente, de forma incomum, e interferem no crescimento celular. Em mamíferos e outros organismos, estes membros necessitam de melhor caracterização (61-63).

O *rasgef1b* contém o domínio RasGEFN, um domínio do tipo Cdc25, um padrão de zíper de leucina, que está normalmente envolvido em associações com outras proteínas, e três motivos de localização nuclear. Mais importante, parece ser o primeiro exemplo de um GEF cuja expressão é induzida e mantida durante o processo de ativação de macrófagos/linfócitos (36). Está bem estabelecido que as proteínas cinases ativadas por mitógenos (MAPKs) estão envolvidas na proliferação de linfócitos T durante a estimulação por mitógenos (64,65), como também a síntese de citocinas pró-inflamatórias por macrófagos expostos a agonistas dos receptores do tipo “Toll” (TLR) (66 e 67). Embora Ras seja um passo inicial que desencadeia a fosforilação de algumas MAPKs, os eventos moleculares posteriores que controlam a ativação dessas proteínas são desconhecidos. Além disso, a localização nuclear desse GEF pode indicar um papel alternativo na ativação leucocitária. Então, o papel da proteína RasGEF1b codificada pelo gene *rasgef1b* na ativação de eventos iniciais e funções desempenhadas pelos macrófagos ativados e linfócitos T deve ser investigado (36). Estes achados poderão, eventualmente, fornecer um melhor entendimento nos eventos cruciais da relação parasita-hospedeiro como a ativação do sistema imune inato pelo *T. cruzi* e a patogênese da Doença de Chagas.

2. Justificativa

Os fatores de troca de guanina, RasGEFs, são componentes celulares fundamentais no processo de ativação das Ras-GTPases em resposta à uma diversidade de estímulos extracelulares. Nos últimos cinco anos uma série de RasGEFs foi identificada, porém estudos mais aprofundados merecem ser levados adiante a fim de determinar a real participação dessas proteínas na ativação das diferentes Ras-GTPases em diferentes processos celulares.

Foi identificado, recentemente, em nosso laboratório, um novo RasGEF cuja expressão foi significativamente induzida em macrófagos murinos e humanos expostos a GPI-mucinas de *T. cruzi*. A expressão desse gene, *rasgef1b*, foi também induzida por LPS. Embora a expressão do *rasgef1b* encontrou-se constitutiva em alguns órgãos como cérebro, pulmão e intestino, um aumento significativo nos níveis do seu mRNA foi observado no timo, baço, rim, fígado, coração e em linfonodos após 20 dias de infecção com *T. cruzi*. Esses resultados sugerem que, no decorrer da doença de Chagas, o *rasgef1b* tenha um importante papel na ativação de outros componentes celulares e consequentemente no estabelecimento de uma resposta imune celular.

De nosso conhecimento, este é o primeiro membro da família RasGEF cuja expressão é induzida. Assim, acreditamos que a caracterização molecular e funcional do *rasgef1b* possibilitará o melhor entendimento da participação dos RasGEFs na resposta imune celular desencadeada por GPI-mucinas e também por outros padrões moleculares associados a patógenos (PAMPs) ativadores dos TLRs, como LPS e poli I:C. Além disso, estes estudos iniciais poderão inferir qual seria o papel das Ras-GTPases em processos infecciosos e, mais especificamente, na Doença de Chagas.

3. Objetivos

3.1. Objetivo Geral

Expressar a proteína RasGEF1b recombinante em sistema procarioto e eucarioto.

3.2. Objetivos Específicos

3.2.1. Analisar a sequência de aminoácidos da proteína hipotética RasGEF1b utilizando ferramentas de bioinformática;

3.2.2. Amplificar a região codificadora do gene *rasgef1b*;

3.2.3. Clonar a região codificadora do gene *rasgef1b* no vetor procariótico pQE-30;

3.2.4. Expressar a proteína His-RasGEF1b recombinante em bactérias;

3.2.5. Confirmar a expressão da proteína recombinante por Western Blotting;

3.2.6. Confirmar a expressão da proteína recombinante por espectrometria de massa;

3.2.7. Produzir anticorpos contra a proteína recombinante His-RasGEF1b em coelhos e galinhas;

3.2.8. Transfectar células HEK293T com plasmídeo codificando RasGEF1b associada ao peptídeo FLAG (FLAG-RasGEF1b);

3.2.9. Determinar a localização celular da proteína FLAG-RasGEF1b em células transfectadas através de fracionamento celular e eletroforese bidimensional.

4. Materiais e métodos

4.1. Análise da proteína RasGEF1b utilizando ferramentas de bioinformática

A análise da sequência de aminoácidos da proteína RasGEF1b foi feita com o programa Protparam. Os alinhamentos de sequências de aminoácidos da proteína RasGEF1b entre diferentes espécies, de sequências de RasGEF1a, RasGEF1b e RasGEF1c humanas e de sequências RasGEF1b e RasGEF1c murinas foram realizados com o programa Clustal W. A construção da árvore filogenética foi feita através do programa Phylip, com um bootstrap de 1000 réplicas. A análise de domínios conservados foi feita com o Protein-protein BLAST. A análise de existência de peptídeo sinal foi feita com o SignalP, e a de possíveis sítios de modificações pós-traducionais foi feita através dos seguintes programas: NetAcet1.0 server, NetNglyc, NetOGlyc, NetPhos, YinOYang 1.2. Prediction Server e Prosite. A análise da possível localização subcelular foi feita através dos programas PTarget, NucPred, DbSubLoc e Psort. A análise da estrutura secundária da proteína RasGEF1b foi feita através do programa HNN Secondary Structure Prediction Method e a modelagem da estrutura terciária foi feita utilizando o programa ESyPred3D Web Server 1.0. Os modelos de estrutura terciária foram visualizados e analisados utilizando o programa Pymol. O gráfico de Ramachandran foi construído utilizando o Procheck v.3.5.4.

4.2. Extração do RNA total de células RAW 264.7 estimuladas com RNA de fita-dupla ou endotoxina bacteriana

Células RAW 264.7 (macrófagos murinos de linhagem contínua) foram cultivadas em meio DMEM com 10% de soro bovino fetal e 40 µg/mL de gentamicina, à 37° C em 5% CO₂. Estas células foram estimuladas com RNA de fita-dupla (poly- rI:rC, 100µg/ml) ou endotoxina bacteriana (LPS, 100ng/ml). O RNA total das células foi isolado utilizando o reagente Trizol (Life Technologies). O reagente Trizol foi acionado à cultura de células e essas amostras foram incubadas à temperatura ambiente por 5 minutos. Adicionou-se 0,2mL de clorofórmio para cada 1mL de Trizol adicionado. As amostras foram homogeneizadas por 15 segundos e posteriormente incubadas à temperatura ambiente por 2 a 3 minutos. Após essa etapa, foram centrifugadas a 12000 x g por 15

minutos a 4°C e foi observada a separação das fases fenólica e aquosa. A fase aquosa foi transferida para um novo tubo, onde foi adicionado 0,5mL de isopropanol, para cada 1mL de Trizol utilizado na fase de homogeneização inicial. As amostras foram incubadas à temperatura ambiente por 10 minutos e centrifugadas por mais 10 minutos à 4°C. O RNA precipitado foi lavado com etanol 75% e o RNA precipitado no fundo do tubo foi eluído com H₂O_{DEPC}. A quantificação do RNA foi feita através de espectrofotometria. A concentração foi determinada levando-se em conta que uma unidade de O.D. (Abs 260 nm) corresponde a uma solução com 40µg/mL de RNA . As leituras a 260 e 280, respectivamente, serviram como parâmetro na estimativa do grau de contaminação do ácido nucléico com proteínas. Os valores considerados satisfatórios foram aqueles entre 1,7 a 2,0.

4.3. RT-PCR para detecção da expressão do cDNA do *rasgef1b*

Um micrograma de cada RNA total proveniente das células RAW 264.7 foi utilizado para uma reação de transcrição reversa, na presença de 200U da enzima SuperScript II (Gibco, BRL); tampão da enzima; 2,5mM de dNTP; 0,1M de 1,4-dithiothreitol (DTT); inibidor de RNase e um primer oligo d(T). Os iniciadores utilizados na reação de PCR, para amplificarem a região codificadora do RasGEF1b foram HindNH2 (5'-CCCAAGCTTGGGAGTATGCCTCAGACGCCCCCTTC-3') e APACOOH (5'-GAAGGGCCCAACTCTGCCCAAAGGC-3').

4.4. Amplificação da região codificadora do gene *rasgef1b*

Para a Reação em Cadeia da Polimerase (PCR) foram utilizados 2µL de cDNAs moldes obtidos por reação de transcriptase reversa de células de macrófagos murinos: cDNA de macrófagos controle (não estimulados), cDNA de macrófagos estimulados com moléculas tGPI e cDNA de macrófagos estimulados com *T. cruzi* (formas tripomastigotas). Nesta reação, foram utilizados 0,4µL de enzima DNA-polimerase Taq (5U/µL, Phoneutria Biotechnology. & Services); 0,4µL de Pfu (2U/µL, Stratagene) e uma mistura contendo, nas concentrações finais: 2µL dNTPs (2mM

dATP, 2mM dCTP, 2mM dGTP, 2mM dTTP); 2 μ L de Tampão Taq 10X (Phonotria); 1 μ L do iniciador 5'-GCTGTCGACGGGAAAGTATGC-3' (iniciador SalF); 1 μ L do iniciador 5'-TCCAAAGCTTAAACTCTGCCC-3' (iniciador HindR); 11,6 μ L de H₂O Milli-Q autoclavada. O volume final da reação foi de 20 μ L. A reação ocorreu por 35 ciclos de 94°C/45 segundos, 55°C/2 minutos, 72°C/2 minutos e 30 segundos e um último ciclo de extensão a 72°C/10 minutos. Os produtos de amplificação foram submetidos ao fracionamento eletroforético em gel de agarose 1%. Após a adição de tampão 6X de amostra para DNA (0,25% azul de bromofenol; 0,25% etilenoglicol; 30% de glicerol em H₂O), os fragmentos foram fracionados no gel em TAE 1X (0,04 M Tris-acetato; 0,001 M EDTA) contendo brometo de etídio (0,5 μ g/mL). Os DNAs fracionados no gel foram observados em luz ultravioleta (U.V.).

Dois microlitros dos produtos obtidos nas PCRs realizadas com cDNAs de macrófagos murinos estimulados com moléculas GPI e com cDNAs de macrófagos murinos estimulados com *T. cruzi* foram reamplificados nas mesmas condições da reação anterior. Os DNAs foram fracionados em gel de agarose 1% corado com brometo de etídio.

A purificação dos fragmentos do gel e demais purificações de fragmentos de géis foram feitas com o “Kit GFX PCR DNA and GEL Band Purification” (Amersham Biosciences Ltd, Buckinghamshire, England), seguindo o protocolo indicado pelo fabricante. Os fragmentos de DNA do gel correspondentes ao tamanho molecular do cDNA do gene *rasgef1b* foram cortados e pesados em um único tubo de microcentrífuga. Foram adicionados 300 μ L de “Capture buffer”, volume que correspondia ao peso obtido em miligramas dos fragmentos do gel. A mistura foi incubada a 60°C por 15 minutos. A coluna do “Kit” foi colocada no tubo coletor e a agarose dissolvida aplicada na coluna. O sistema foi incubado por 5 minutos à temperatura ambiente. O sistema foi brevemente centrifugado por 30 segundos (14000 \times g) e o tubo coletor esvaziado. Foram adicionados à coluna 500 μ L do tampão de lavagem (10mM Tris-HCl pH 8,0; 0,1mM EDTA) e o sistema foi centrifugado novamente. O tubo coletor foi trocado por um tubo de microcentrífuga novo e 20 μ L de TE (10mM Tris pH 8,0; 1mM EDTA pH 8,0) foram aplicados na coluna. O sistema foi incubado por 5 minutos à temperatura ambiente e centrifugado a 14000 \times g por 1 minuto.

4.5. Clonagem no vetor pCR[®] TOPO 2.1 (Invitrogen)

Do volume obtido na purificação dos fragmentos do gel, 7,8 μL foram incubados a 72°C por 10 minutos em um tubo de microcentrífuga contendo 2mM de dNTPs, Tampão 10X e Taq DNA polimerase. Depois disso, a 4 μL dessa reação, foram adicionados 1 μL de solução salina (1,2M NaCl; 0,06M MgCl₂) e 1 μL do vetor pCR[®] TOPO 2.1. Esses reagentes foram incubados à temperatura ambiente por 30 minutos. A transformação bacteriana foi feita adicionando 2 μL do produto da ligação do pCR[®] TOPO 2.1 com o cDNA do *rasgef1b* e 50 μL de células bacterianas XL1-blue (*E. coli*) quimicamente competentes. A mistura foi incubada no gelo durante 30 minutos. Após este tempo foi feito choque térmico pela incubação por 1 minuto a 42°C, seguido pela incubação em gelo por 1 minuto. As bactérias foram incubadas, sob agitação, a 37°C por uma hora em meio LB sem antibiótico. Aliquotas da suspensão bacteriana foram plaqueadas em meio LB ágar contendo 100 $\mu\text{g}/\text{mL}$ de ampicilina. As placas foram colocadas na estufa, permanecendo a 37°C por 18 horas. Algumas colônias obtidas nessa transformação foram repicadas, utilizando uma ponteira estéril, em uma placa contendo LB ágar e 100 $\mu\text{g}/\text{mL}$ de ampicilina. O restante do conteúdo da ponteira foi dissolvido 30 μL de tampão de lise (10mM Tris pH 8,0; EDTA 0,1mM; 0,1 % Tween[®] 20). As misturas foram aquecidas a 95°C por 10 minutos e centrifugadas a 12000 $\times g$ por 2 minutos. Foram utilizados 4 μL do sobrenadante de cada tubo para PCR. As reações foram realizadas nas mesmas condições citadas anteriormente (4.4), sendo que, alternativamente, foram utilizados dois iniciadores: M13 direto e reverso, os quais flanqueiam o sítio múltiplo de clonagem do vetor pCR[®] TOPO 2.1.

Foi feita extração de DNA plasmidial em baixa escala de duas colônias positivas utilizando o “Kit QIAprep Miniprep” (Qiagen), de acordo com o protocolo indicado pelo fabricante. Cada colônia foi cultivada em 4mL de LB contendo 100 $\mu\text{g}/\text{mL}$ de ampicilina a 37°C, sob agitação, por 18 horas. As culturas bacterianas foram centrifugadas a 10000 $\times g$ por 5 minutos. Os sobrenadantes foram descartados e os sedimentos obtidos solubilizados em 300 μL do tampão de ressuspensão P1 (50mM Tris.Cl pH 8,0; 10mM EDTA; 100 $\mu\text{L}/\text{mL}$ RNase A) cada. Foram adicionados 300 μL de tampão de lise P2 (200mM de NaOH; 1% SDS (p/v)) por tubo e feita a inversão dos mesmos cerca de seis vezes. Os tubos foram incubados à temperatura ambiente por 5 minutos. À mistura, foi

adicionado tampão de neutralização N3 (3M acetato de potássio pH 5,5) gelado, 300 μ L por tubo, seguido de incubação no gelo por 5 minutos. Os conteúdos foram centrifugados a 14000 \times g por 1 minuto. As duas colunas foram equilibradas com 1mL do tampão QBT (750 mM NaCl; 50mM MOPS, pH 7,0; 15% isopropanol (v/v); 0,15% Triton[®]X-100(v/v)) cada. Os sobrenadantes foram aplicados nas colunas e estas foram lavadas quatro vezes com 1mL do tampão de lavagem QC (1,0 M NaCl; 50mM MOPS pH 7,0; 15% isopropanol (v/v)). Os tubos coletores foram descartados e substituídos por tubos de microcentrífuga. Os fragmentos de DNA foram eluídos com 800 μ L de tampão de eluição QF (1,25M NaCl; 50mM Tris.Cl pH 8,5; 15% isopropanol (v/v)) e sedimentados com 560 μ L de isopropanol à temperatura ambiente. Após centrifugação a 14000 \times g por 5 minutos, os sobrenadantes foram removidos e aos sedimentos foram adicionados 700 μ L de etanol 70%, seguindo a mesma centrifugação anterior. Os sobrenadantes foram removidos e os sedimentos secaram em temperatura ambiente por 10 minutos e foram dissolvidos em 30 μ L de TE.

As digestões enzimáticas de 3 μ L do produto de extração plasmidial das colônias do cDNA do *rasgef1b* clonado no pCR[®] TOPO 2.1 e da construção pcDNA3.1/V5-His*rasgef1b* foram feitas com as enzimas *Bam*H I, *Eco*R I, *Hind* III, *Sac* I, *Kpn* I e *Xho* I (Promega Corporation, Madison, WI). Foram utilizados, para cada reação, 1 μ L de cada enzima, tampão 10X específico para cada enzima segundo catálogo do fabricante, BSA10X e H₂O Milli-Q autoclavada suficientes para reações de 20 μ L. As misturas foram incubadas a 37°C por 1 hora e 30 minutos e fracionadas em gel de agarose 1% corado com brometo de etídio. O gel foi observado em luz U.V.

4.6. Clonagem no vetor procariótico pQE-30 (Qiagen)

A digestão enzimática de 20 μ L da construção pCR[®] TOPO 2.1*rasgef1b* foi feita com 2 μ L da enzima de restrição *Hind* III, 3 μ L de BSA 10X, 3 μ L de tampão E 10X e 2 μ L de H₂O Milli-Q autoclavada. A reação foi incubada a 37°C por 6 horas e o resultado verificado em gel de agarose 1% corado com brometo de etídio (0,5 μ /mL), observado em luz U.V. O fragmento digerido foi cortado do gel com bisturi, pesado e purificado. O DNA foi eluído em 20 μ L de TE e digerido com a

enzima de restrição *Sal I* (Promega), utilizando os mesmos reagentes para a digestão anterior, com tampão próprio. O fragmento digerido foi extraído do gel com um bisturi, pesado e purificado. O DNA foi eluído em 20µL de TE.

Quinze microlitros do vetor pQE-30 (Figura 1) foram digeridos utilizando 2µL da enzima *Hind III* e os outros reagentes necessários para uma reação cujo volume final foi de 15µL. A reação foi incubada a 37°C por 6 horas e o fragmento digerido foi cortado do gel e purificado. O DNA foi eluído em 20µL de TE, cuja quantidade foi digerida com 2µL da enzima *Sal I*, mais 3µL de tampão D 10X, 3µL de BSA 10X e 2 µL de H₂O Milli-Q autoclavada, por 6 horas a 37°C. Foram utilizados, para a reação de ligação, 3µL do vetor pQE-30 e 5µL do inserto do cDNA do *rasgef1b*, ambos digeridos com *Sal I* e *Hind III*, 1µL de T4 DNA ligase (Promega) e 1µL de tampão 10X (Promega).

Cinco microlitros do produto de ligação foram transformados em bactérias XL1-blue quimicamente competentes, de acordo com os protocolos de transformação descritos anteriormente. Das colônias obtidas da transformação, dez foram escolhidas para diagnóstico por PCR utilizando os pares de iniciadores usados na amplificação inicial e o SalF juntamente com um iniciador flanqueando a região de clonagem do pQE-30 (pQE-30R). Os procedimentos foram os mesmos utilizados para o diagnóstico das colônias obtidas da transformação pCR[®] TOPO 2.1*rasgef1b*. Os produtos de PCR foram fracionados em gel de agarose corado com brometo de etídio.

Foi feita extração de DNA plasmidial em baixa escala de três colônias, escolhidas entre as últimas 10, utilizando o “Kit QIAprep Miniprep” (Qiagen, Hilden, Germany). A extração ocorreu conforme as instruções indicadas pelo fabricante, como descrito anteriormente. Um novo diagnóstico por PCR das três colônias foi realizado, desta vez utilizando uma combinação entre os iniciadores SalF e HindR; SalF e pQE-30R; pQE-30F (outro iniciador flanqueador do sítio múltiplo de clonagem do pQE-30) e pQE-30R; pQE-30F e HindR. Os produtos amplificados foram fracionados em gel de agarose 1% corado com brometo de etídio.

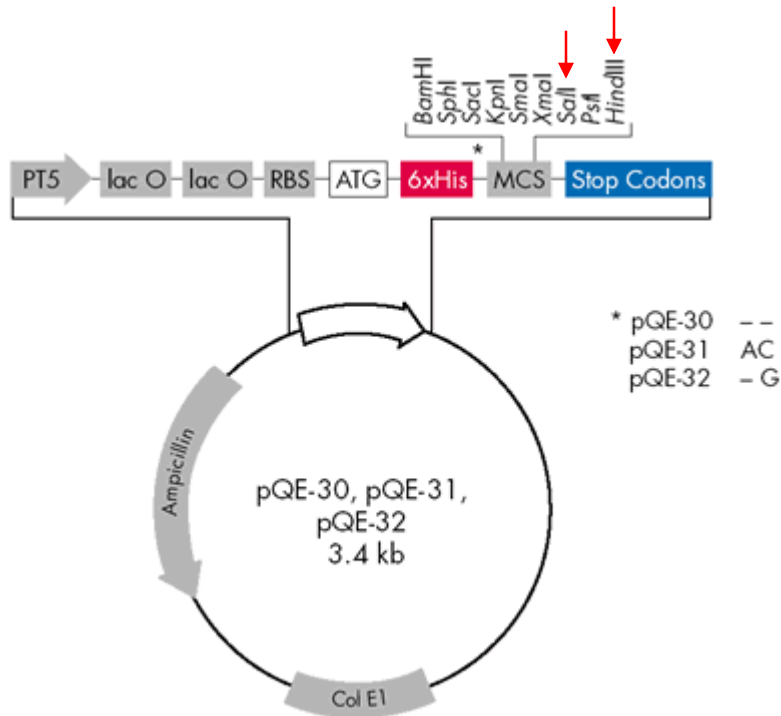


Figura 1. Mapa do vetor pQE-30 (Qiagen). As setas apontam para os sítios de restrição das enzimas *Sal* I e *Hind* III presentes no sítio múltiplo de clonagem do vetor. Esses sítios estão também presentes nos iniciadores utilizados para a clonagem do *rasgef1b* no pQE-30. PT5, promotor T5 de fago reconhecido pela RNA polimerase de *E. coli*; RBSII, sítio sintético de ligação do ribossomo; MCS, sítio múltiplo de clonagem; Ampicilin, gene de resistência à ampicilina.

4.7. Sequenciamento do cDNA do *rasgef1b* clonado no pQE-30

Um micrograma dos plasmídeos obtidos na clonagem contendo o inserto de interesse foi sequenciado utilizando os iniciadores SalF e HindR e o “DYEnamic™ ET Terminator Kit” (Amersham, Biosciences) no sequenciador automático Mega BACE 1000 DNA Analysis System no Núcleo de Análise de Genoma e Expressão Gênica (NAGE, ICB- UFMG). A análise da procura por homologia em banco de dados de seqüências nucleotídicas do Centro Nacional de Biotecnologia e Informação (NCBI) foi realizada utilizando-se o algoritmo BLAST (<http://www.ncbi.nlm.nih.gov/BLAST/>) (54).

4.8. Expressão da proteína recombinante em bactérias

Três colônias do pQE30*rasgef1b*, uma colônia do vetor pQE-30 e uma colônia do vetor pQE-40 com o cDNA do gene da enzima dihidrofolato redutase clonado (pQE-40DHFR) foram inoculadas em 4mL de meio LB com 100µg/mL de ampicilina, cada. As misturas foram incubadas, sob agitação, a 37°C, por 18 horas. Foi feito o inóculo 1:200 de cada colônia em 20mL de LB com ampicilina. A absorbância a 600nm em espectrofotômetro foi verificada com 1mL de cada inóculo até a O.D. desejada (entre 0,4 e 0,6). A indução foi feita com IPTG (isopropiltio-β-galactosidase) a 1mM concentração final. No momento em que o IPTG foi adicionado (0h), retirou-se 1,5mL de cada inóculo, que foi centrifugado a 10000 \times g por 2 minutos. O sobrenadante foi descartado e o sedimento solubilizado em 50µL de H₂O Milli-Q autoclavada e tampão de amostra 6X para proteína. As amostras foram estocadas a -70°C. Após 4 horas (4h) e 18 horas (18h) de indução, foram realizados os mesmos procedimentos, sendo retirada a mesma quantidade de bactérias. Os extratos protéicos totais foram fracionados em SDS-PAGE 12% e corados com Coomassie blue. Os géis foram feitos a partir de uma solução estoque (p/v) de acrilamida 29,2 % e bis-acrilamida 0,8 % diluída para uma concentração final de 12 % em Tris-HCl 375mM, pH 8,8; e SDS 0,1% (p/v) para preparação do gel de separação e, diluída a 4 % em Tris-HCl 125mM, pH 6,8; SDS 0,1% (p/v) para preparação do gel de aplicação e concentração das amostras. A polimerização dos géis de concentração e separação ocorreu na presença de persulfato de amônio e TEMED (Sigma-Aldrich, Inc.) nas concentrações finais de 0,1 % e 1:2000, v/v, respectivamente. As amostras de proteínas foram preparadas em tampão de amostra (Tris-HCl 0,0625M, pH 6,8 contendo SDS 2%, glicerol 10%; azul de bromofenol 0,0005%; β-mercaptoetanol 2,5%), homogeneizadas em “vortex” e fervidas por 5 minutos a 90 – 100°C. A eletroforese foi realizada a 200V por cerca de 45 minutos, em tampão Tris-glicina (Tris base 25mM pH 8,3, glicina 250mM, SDS 0,1%). O gel foi corado com azul de coomassie (Coomassie blue R250 0,25%, metanol 50%, ácido acético 10 %) durante 2 horas em temperatura ambiente e lavado várias vezes com solução descorante total (metanol 30%, ácido acético 10%). O peso molecular da proteína foi comparado com padrão de peso molecular de proteínas Bench Mark TM Protein Ladder (Invitrogen).

4.9. Western Blotting com anticorpo anti-His

Alternativamente, as amostras foram fracionadas em SDS-PAGE 12,5% e transferidas eletroforeticamente para uma membrana de PVDF. A membrana foi incubada em solução de bloqueio (PBS 1X 0,01M de NaHPO₄; 0,15M de NaCl, pH 7,2; 5% leite em pó desnatado; 0,1% Tween® 20), por 1 hora, sob agitação, à temperatura ambiente. Nessa mesma solução, foi feita uma diluição de 1:4000 do anticorpo monoclonal anti-histidina (RGS.HisAntibody (100) Isotype mouse IgG₁, Qiagen). A membrana foi incubada na solução com anticorpo, nas mesmas condições, por 1 hora e, posteriormente, lavada com solução de bloqueio por três vezes, 10 minutos de duração cada lavagem. A membrana foi então incubada na solução de bloqueio com o segundo anticorpo, IgG de cabra anti-IgG de camundongo conjugado com peroxidase (Sigma), diluído 3000 vezes. Foram feitas três lavagens, 15 minutos cada, em solução de bloqueio, solução PBS 1X com 0,1% de Tween® 20 e solução PBS 1X. A revelação foi feita em câmara escura, utilizando a solução substrato “ECL plus Western Blotting Detection System” (Amersham Biosciences) em filme Raio-X (Kodak).

4.10. Determinação da solubilidade da proteína recombinante

Uma colônia selecionada foi induzida por 4 horas com IPTG, num total de 50mL de cultura bacteriana, conforme descrito anteriormente. Esse volume foi dividido igualmente em dois tubos (1 e 2) e ambos foram centrifugados a 6000 \times g por 15 minutos a 4°C. O sedimento celular do tubo 1 foi solubilizado em 2,5mL de tampão de lise (50mM NaH₂PO₄; 300mM NaCl; 10mM Imidazol pH 8,0). O sedimento obtido do tubo 2 foi solubilizado em 2,5mL de tampão de lise contendo 0,25% de Tween® 20. Os tubos foram mantidos por 30 minutos no gelo. Em seguida, ambos os lisados foram sonicados seis vezes com pulsos de 10 segundos em potência 30 e intervalo de 10 segundos entre cada pulso. Os sonicados foram centrifugados a 10000 \times g por 30 minutos a 4°C. Os sobrenadantes foram reservados e denominados 1A e 2A originados dos tubos 1 e 2, respectivamente. O sedimento

do tubo 1 foi solubilizado em 2,5mL de tampão de lise e reservado como amostra 1B. O sedimento do tubo 2 (2B) foi solubilizado em 2,5mL de tampão de lise contendo 0,25% de Tween[®] 20 e re-extraído por sonicação conforme descrito anteriormente. O sobrenadante obtido da re-extração foi denominado 2B₁ e o sedimento denominado 2B₂. Frações das amostras 1A, 1B, 2A, 2B, 2B₁ e 2B₂ foram fracionadas em SDS-PAGE 12,5%. O gel de poliacrilamida foi posteriormente corado com “coomassie blue”.

4.11. Purificação da proteína recombinante His-RasGEF1b

Uma colônia (1) foi inoculada em 500mL de meio LB com ampicilina (100mg/mL) e a cultura bacteriana cresceu conforme condições descritas anteriormente. Na OD₆₀₀ 0,5 foi adicionado 1mM de IPTG. Após 4 horas, a cultura foi centrifugada a 7000 *x g* por 10 minutos, a 4°C. O sobrenadante foi descartado e o sedimento solubilizado em 25 mL de tampão de lise. As bactérias foram lisadas através de três ciclos de congelamento (- 70°C) e descongelamento (banho-maria a 42°C) e depois sonicadas em três pulsos de 30 segundos, potência de 30% em sonicador (Fisher Scientific), intercalados por 30 segundos no gelo, para liberação total das proteínas citoplasmáticas. O lisado foi centrifugado a 7000 *x g* por 20 minutos a 4°C. O sobrenadante foi descartado e o sedimento solubilizado em 25mL de tampão de lise contendo 8M de Uréia. Esse volume foi filtrado em filtro 0,45µm imediatamente antes da purificação no sistema Äkta Explorer (Amersham Biosciences).

O método cromatográfico utilizado para a purificação da His-RasGEF1b foi afinidade pelo níquel utilizando a coluna “HiTrap[™] Chelating” (Amersham Biosciences) de 1mL, carregada com 0,5mL de NiSO₄ (100mM). A coluna foi equilibrada com 5 CV (volumes de coluna) com o tampão de corrida (50mM NaH₂PO₄, 300 mM NaCl, 8M Uréia), sendo o fluxo utilizado de 1mL/min. Após a injeção dos 25mL do extrato protéico a coluna foi lavada com 10mL do tampão de corrida para lavar as proteínas não ligadas. A proteína His-RasGEF1b foi então eluída da coluna com o tampão de eluição (50mM NaH₂PO₄, 300mM NaCl, 8M Uréia, 500mM de imidazol) e coletada em 10

frações de 1mL cada. Para finalizar, a coluna foi reequilibrada com 5 CV com o tampão de corrida. As frações foram fracionadas em SDS-PAGE, posteriormente corado com Coomassie blue.

A troca de tampão da proteína purificada foi feita com a coluna “HiTrap™ Desalting Column” (Amersham Biosciences) de 5mL. A coluna foi equilibrada com 5 CV de PBS 1X e 1mL da amostra foi injetado na coluna. O total de 20 frações de 0,5mL cada foi coletado. O fluxo utilizado foi de 2mL/min. Foram fracionados, em SDS-PAGE 12,5%, 10µL de cada fração obtida nesse processo.

4.12. Eletroforese bidimensional

Para a realização da Isoeletrofocalização (IEF), a amostra protéica foi incubada à temperatura ambiente, a 20 rpm durante 30 minutos. Ao tampão IEF (8M Uréia, 4% CHAPS, 2M Thiouréia, traço de azul de bromofenol) foram acrescentados 10mg/mL de DTT, 0,2% de anfólitos específicos para a “strip” utilizada e o volume da amostra protéica. A mistura foi agitada por 1 hora e centrifugada por 30 minutos a 11500 x g. O sobrenadante foi retirado e aplicado na bandeja IEF. A “strip” (Bio-Rad Laboratories, Hercules, CA, USA) foi colocada na bandeja com amostra e após 10 minutos foi adicionado óleo mineral. A rehidratação da “strip” foi feita de forma passiva durante 4 horas, sem corrente, e de forma ativa, durante 12 horas, a 50V (Volts) no equipamento “PROTEAN IEF Cell” (Bio-Rad). A Isoeletrofocalização foi realizada a 500V por uma hora, 1000V por uma hora, 8000V por 2 horas e 8000V até serem acumulados 40000V/hora. A “strip” foi retirada da máquina e o excesso de óleo mineral drenado. A “strip” foi equilibrada em tampão de equilíbrio (6M Uréia; glicerol 30%; SDS 2%; 50mM Tris-HCl pH8,8; traço de azul de bromofenol) contendo 10mg/ml de DTT, por 15 minutos, sob agitação em bandeja descartável. O tampão foi trocado por outro contendo 25mg/ml de iodoacetamida e a amostra incubada por 15 minutos, sob agitação, à temperatura ambiente. A “strip” foi retirada da bandeja, lavada com tampão de corrida SDS-PAGE, colocada sobre o gel de poliacrilamida 12%, e selada com agarose 0,5% em tampão de corrida. A eletroforese foi realizada sob 150V. Após a corrida, o gel foi corado com Coomassie blue G250 2%, compatível com análise por espectrometria de massa.

Alternativamente, a amostra protéica foi transferida para uma membrana de nitrocelulose, para a técnica de Western Blotting, já descrita. O gel corado e o filme revelado foram digitalizados e normalizados. As imagens foram analisadas pelo programa “PDQUEST-7.3.0” (BioRad) para identificar a posição das proteínas nos géis utilizando como referência os marcadores de pI e de peso molecular. A utilização do programa PDQUEST-7.3.0 permite considerar vários critérios para análise das proteínas presentes nos géis: abundância, mudanças induzidas ou desenvolvidas na expressão, modificações e presença de uma fração particular.

4.13. Microcaracterização dos “spots” retirados de gel bidimensional

Os “spots” com maior probabilidade de serem a proteína RasGEF1b foram retirados do gel com bisturi estéril, eluídos e digeridos com tripsina. Primeiramente, cada spot foi lavado com 400µL de acetonitrila 50% / bicarbonato de amônio 25mM por 15 minutos, no “vortex”. O processo foi feito três vezes. Os “spots” foram desidratados com 200µL de acetonitrila por 5 minutos e secados rapidamente em “speed vac”. Dez microlitros de tripsina (20ng, Promega) foram adicionados a cada “spot” e os tubos foram mantidos no gelo por dez minutos. Em seguida foram adicionados 20µL de bicarbonato de amônio (50mM) e os tubos foram incubados a 37°C por 16 horas. A tripsina foi então transferida para um tubo novo e ao tubo que continha o spot foram acrescentados 30µL de ácido fórmico 5% / Acetonitrila 50%, por 30 minutos. Esse processo foi feito duas vezes. As amostras foram concentradas até o volume de 10µL, passadas em “ZipTip” coluna C8 (Eppendorf AG, Germany) e eluídas em 8µL de acetonitrila 50% / TFA0,1%. As amostras foram analisadas no Maldi/Tof-Tof (Applied Biosystems, Foster City, CA, USA), no Laboratório de Toxinologia do Departamento de Fisiologia e Farmacodinâmica do Instituto Oswaldo Cruz, Rio de Janeiro-RJ.

4.14. Imunização dos coelhos e galinhas com a proteína recombinante His-RasGEF1b

Após a purificação da proteína recombinante His-RasGEF1b expressa em bactérias, através de cromatografia de afinidade por coluna de níquel, em HPLC, iniciamos os experimentos de

imunização. As amostras contendo a proteína His-RasGEF1b foram fracionadas em SDS-PAGE 12% e os géis corados com Coomassie blue. Foram aplicados em três géis 50µg, e em outros três géis 150µg de proteína recombinante. A banda relativa ao peso molecular da proteína His-RasGEF1b foi cortada do gel e triturada em nitrogênio líquido com um pistilo. A proteína triturada foi colocada em um tubo de microcentrífuga de 1,5mL; e a este foi adicionado adjuvante de Freund Sigma) em quantidade igual ao volume da proteína triturada. Esta etapa foi realizada um dia antes das três imunizações em galinhas e coelhos. Foram utilizados três coelhos (*Oryctolagus cuniculus*) e três galinhas (*Gallus gallus*) para os experimentos de imunização. O coelho e a galinha 1, controles, foram imunizados apenas com o adjuvante de Freund, completo ou incompleto. Os coelhos e as galinhas 2 e 3, foram imunizados com 50µg e 150µg da proteína His-RasGEF1b respectivamente, juntamente com a quantidade de adjuvante de Freund correspondente ao volume de proteína triturada. Na primeira imunização foi utilizado o adjuvante de Freund completo, e nas posteriores foi utilizado o adjuvante incompleto. A imunização foi feita através de injeção intramuscular na coxa dos coelhos e no músculo peitoral das galinhas, com um intervalo de 15 dias entre cada imunização. Foi coletado 5mL de sangue dos coelhos antes da primeira imunização, e 15 dias após a terceira imunização. Os ovos das galinhas foram coletados no mesmo período.

4.15. Extração dos anticorpos de coelhos e galinhas

Após a retirada do sangue dos coelhos, este foi colocado em um tubo de 15mL, e centrifugado a 3000 rpm a 4°C por 15 minutos. O sobrenadante foi retirado e o soro utilizado para os ensaios de ELISA. A extração dos anticorpos IgY de galinha foi feita diretamente da gema do ovo. Após coletar os ovos, a gema foi separada da clara e solubilizada em 25ml de PBS 100mM. À esta mistura, adicionou-se 25mL de clorofórmio e agitou-se até ser obtida uma fase semi-sólida. Logo após, centrifugou-se a 1200 x g por 30 minutos. O sobrenadante foi retirado, e adicionou-se PEG 6000 sólido até atingir uma concentração final de 12% (p/v). Agitou-se novamente até obter uma mistura homogênea a qual é centrifugada a 15700 x g por 10 minutos. O sobrenadante foi descartado e o sedimento solubilizado em 2mL de PBS 100mM.

4.16. ELISA com soro dos coelhos e anticorpos de galinhas

Inicialmente uma placa com 48 poços e uma placa com 96 poços foram sensibilizadas com 100µL/poço de proteína His-RasGEF1b a 5µg/ml em tampão carbonato (15mM Na₂CO₃, 35mM NaHCO₃, pH 9,6). As placas foram incubadas “overnight” a 4°C. O volume foi desprezado por inversão das placas e em seguida adicionado 110µL de tampão de bloqueio (PBS1x, 2% leite em pó desnatado) e incubado por 1 hora e 30 minutos a 37°C. Em seguida, o soro dos coelhos e anticorpos das galinhas foram diluídos em tampão de incubação (PBS1x, 2% leite em pó desnatado, 0,05% Tween®20) numa concentração variando de 1:20 à 1:1280. Nos 48 poços da placa, foram adicionados sucessivamente 100µL dos soros dos coelhos antes da primeira, e após a terceira imunização, nas diferentes concentrações. O mesmo foi realizado nos 96 poços com os anticorpos de galinha. Foi feita uma incubação por 1 hora e 30 minutos a 37°C e em seguida lavado seis vezes com tampão de lavagem (0,23M NaCl, 0,05% Tween®20). O anticorpo secundário foi diluído em tampão de incubação numa concentração de 1:10000. Foi utilizado o anticorpo anti-IgG coelho (Promega) e o anticorpo anti-IgY (Sigma), ambos conjugados com peroxidase. A incubação foi feita com 100µL em cada poço, inclusive no branco (poços controles sem anticorpos primários), por 1 hora a 37°C e em seguida lavado por seis vezes com tampão de lavagem. A revelação foi feita com 10ml de tampão citrato-fosfato (0,05M Na₃C₆H₅O₇ . 2H₂O; 0,05M Na₂H₂PO₄ .H₂O) contendo 10µL de TMB e 3µL de H₂O₂ 40 volumes. A reação foi parada com H₂SO₄ 1:20 e a leitura feita a 450nm em espectrofotômetro.

4.17. Transfecção das células HEK 293T com o plasmídeo pFLAG contendo a região codificadora do *rasgef1b*

Células HEK 293T (fibroblastos de rim de embrião humano) foram cultivadas em meio DMEM (Dulbecco’s MEM) com 10% de soro bovino fetal (Gibco) e 100unidades/mL de penicilina/streptomicina, à 37°C em 5% CO₂. A transfecção foi feita pelo método de fosfato de

cálcio, com 2µg do plasmídeo pFLAGCMV2 (Sigma) (Figura 2) contendo a região codificadora do gene *rasgef1b* e, como controle, com 2µg do vetor vazio pcDNA3.1/V5-His (Invitrogen). A passagem das células foi realizada em placas de 10 cm de diâmetro, de forma que no dia da transfecção a confluência estava 70%. Para a transfecção, o meio foi trocado e após 4 horas o complexo de DNA foi adicionado em cada placa. O complexo é formado adicionando-se CaCl₂ 2M ao DNA a ser transfectado e, à essa mistura, HeBs 2X (0,28M NaCl; 0,05M Hepes; 1,5mM NaH₂PO₄). As células foram incubadas em estufa a 37°C, durante a noite, o meio foi retirado e as placas lavadas com PBS. Novo meio foi repostado e as células permaneceram na estufa por cerca de 16 horas.

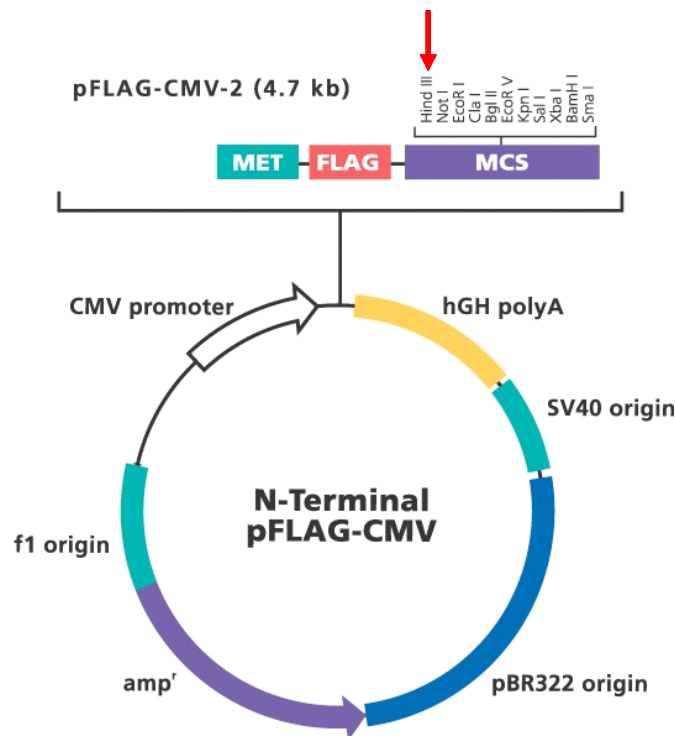


Figura 2. Mapa do vetor pFLAG-CMV2 (Sigma). A seta aponta para o sítio de restrição da enzima *Hind* III presente no sítio múltiplo de clonagem do vetor. Esse sítio está também presente nos iniciadores utilizados para a clonagem do *rasgef1b* no pFLAG-CMV2. CMV promoter: promotor do citomegalovírus humano; SV40, f1 origin, pBR2 origin: origens de replicação; amp': gene de resistência à ampicilina. A clonagem do cDNA do *rasgef1b* neste vetor foi feita e cedida pelo Dr. Aristóboles Mendes da Silva.

4.18. Fracionamento subcelular

Para o fracionamento subcelular, 10^7 células foram lavadas em PBS gelado e ressuspensas em 1mL de tampão de homogeneização (20mM Hepes-KOH, pH 7.5; 10mM KCl; 1,5 mM MgCl₂; 1mM Na-EDTA; 1mM DTT; 250mM sacarose; 0.1mM PMSF e 8µg/ml de aprotinina). Após incubação no gelo por 10 minutos, as células foram homogeneizadas por 20 passagens através de uma seringa de insulina de 1mL e centrifugadas a 700 x g, por 10 minutos, à 4°C, para remover núcleos e debris celulares. O sobrenadante foi coletado e centrifugado a 10000 x g, por 25 minutos, a 4°C. O sedimento resultante foi utilizado como fração de membrana pesada. O sobrenadante foi coletado e centrifugado a 100000 x g, por 60 minutos, a 4 °C, para originar a fração de membrana leve (sedimento) e o citosol (sobrenadante). Os sedimentos foram normalizados em tampão de homogeneização (68). Para identificar cada fração subcelular, foi feito “Western Blotting” com anticorpos contra proteínas específicas de cada fração. Para o núcleo, foi utilizado o anti-histona 4 (Santa Cruz) na diluição de 1:2000, para a fração de membranas pesadas o anti-calreticulina diluído 1:1000 vezes, para a fração de membranas leves o anti-rab 4, 1:200 vezes, e para o citosol o anti-actina (Sigma) 1:3000 vezes. Foi utilizado o anti-FLAG (Sigma), diluído 6000 vezes, para a localização subcelular da proteína RasGEF1b expressa em fusão com o peptídeo FLAG. Anticorpos secundários anti-coelho diluídos 5000 vezes foram utilizados para a revelação dos anticorpos anti-histona, anti-calreticulina e anti-Rab4. Anticorpos secundários anti-camundongo diluídos 2000 vezes foram utilizados para a revelação dos anticorpos anti-actina e anti-FLAG.

As amostras correspondentes às frações de membrana pesada das células transfectadas foram submetidas à eletroforese bidimensional, coradas com Coomassie coloidal e/ou transferidas para membrana de nitrocelulose para a realização de Western Blotting com anti-FLAG.

5. Resultados

5.1. Análise da proteína RasGEF1b utilizando ferramentas de bioinformática

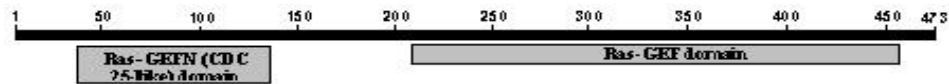
A utilização de programas de bioinformática que fazem predições *in silico* sobre uma dada seqüência fornecida é ideal para obter informações, mesmo que hipotéticas, sobre proteínas ainda não caracterizadas. A partir da seqüência de aminoácidos de uma proteína, pode-se fazer predições a cerca de modificações pós-traducionais na seqüência protéica, estruturas secundárias, estrutura terciária e até mesmo de interações com outras proteínas. Os resultados gerados guiam o trabalho experimental e fornecem subsídios para perguntas biológicas pertinentes à proteína analisada. Uma vez que a proteína RasGEF1b está em processo de caracterização, foi feita uma análise preliminar utilizando ferramentas de bioinformática para serem obtidas predições sobre sua estrutura.

Segundo o Prosite (69), a proteína RasGEF1b possui um domínio RasGEF, na porção N-terminal, entre os aminoácidos 34 a 161, e um domínio catalítico, de troca de nucleotídeos guanina, situado entre os aminoácidos 205 a 453. A assinatura de domínio de fatores de troca de nucleotídeos guanina da proteína Ras está no domínio catalítico, entre os aminoácidos 371 a 401. O domínio RasGEFN contém um zíper de leucina codificado pelos aminoácidos 72-93 (LFMHPYELMAKVCHLCVEHQRL), e o domínio RasGEF contém três motivos de localização nuclear PVKKKHR, PVSRLKK, e PFERDRK codificados pelos aminoácidos 276 a 282, 316 a 322 e 421 a 427, respectivamente. Além disso, o programa detectou um sinal de direcionamento para microcorpúsculos entre os aminoácidos 471 a 473 (Figura 3). Verificou-se que a seqüência da proteína RasGEF1b humana apresenta alta similaridade com a proteína murina. As únicas diferenças são substituições conservadas de 12 aminoácidos, com 97,6% de homologia.

O programa pTarget fornece predições de localização subcelular de uma seqüência protéica (70). Para a proteína RasGEF1b a predição de localização é de 81,4% no citoplasma celular. O programa DBsubloc (Database of Protein Subcellular Localization) prediz localização subcelular para seqüências eucariotas (71). Segundo este programa, a localização da RasGEF1b é nuclear, com 74% de exatidão. O programa Psort (72) também atribuiu o maior score (0,6) para localização nuclear. Entretanto, estes dois últimos resultados são contrapostos pela predição fornecida pelo programa NucPred, cujo score para a RasGEF1b é 0,64. Neste programa, somente

seqüências com “score” acima ou igual a 0,8 são preditas para terem localização subcelular nuclear, com 93% de exatidão (73).

RasGEF1b



RasGEF1b-Mouse and Human homologues

Mouse:	1	MPQTPPFSAMFDSSGYNRMLYQS AEDSCGGLYYHDMNLLSGSLEALIQHLVFNVDYYPDR	60
Human:	1	MPQTPPFSAMFDSSGYNRMLYQS AEDSCGGLYYHDMNLLSGSLEALIQHLVFNVDYYPDR	60
Mouse:	61	TYIFTFLSSRLFMHPYELMAKVCHLCVEHQRLSEGDGDKNQMRKIAPKILQLLTEWTET	120
Human:	61	TYIFTFLSSRLFMHPYELMAKVCHLCVEHQRLSDPDSDKNQMRKIAPKILQLLTEWTET	120
Mouse:	121	FPYDFRDERMMRNLKDLAHRMASGEEQTYRKNVQQMMQCLIRKLAALSQYEEVLAKLSST	180
Human:	121	FPYDFRDERMMRNLKDLAHRMASGEEQTYRKNVQQMMQCLIRKLAALSQYEEVLAKLSST	180
Mouse:	181	ATDRLTULKTKPQS IQRD IMTVCSDPYTLAQLLTHIELERLNYIGPEEFVQAIFVQKDFLD	240
Human:	181	ATDRLTULKTKPQS IQRD IITVCMDPYTLAQLLTHIELERLNYIGPEEFVQAIFVQKDFLD	240
Mouse:	241	NDKSCYSEKPKTRNLEAYVEWFMRLSYLVATEICMPVKKKHRARMIEYFIDVARECFNIG	300
Human:	241	NDKSCYSEKPKTRNLEAYVEWFMRLSYLVATEICMPVKKKHRARMIEYFIDVARECFNIG	300
Mouse:	301	NFNSLMAIISGMNMS PVSRLKKTWAKVKTAKFDILEHQMDPSSMFFNYRTALRGAAQRSL	360
Human:	301	NFNSLMAIISGMNMS PVSRLKKTWAKVKTAKFDILEHQMDPSSMFFNYRTALRGAAQRSL	360
Mouse:	361	TAKSSREKIVIPFSSLLIKDIYFLNEGCVNRLPNGHVNFKEKFWELAKQVSEFMTWKQVEC	420
Human:	361	TAKSSREKIVIPFSSLLIKDIYFLNEGCVNRLPNGHVNFKEKFWELAKQVSEFMTWKQVEC	420
Mouse:	421	PFERDRKILQYLLSVPVFS DALYLASYESEGPENNI EKDRWKSLS SLLGRV	473
Human:	421	PFERDRKILQYLLSVPVFS DALYLASYESEGPENNI EKDRWKSLS SLLGRV	473

Figura 3. Análise comparativa de similaridade entre as proteínas hipotéticas humana e murina codificadas pelo gene *rasgef1b*. Na parte superior da figura está esquematizada a localização dos domínios RasGEFN (domínio CD25-like) e RasGEF presente nos homólogos humano e murino. Na parte inferior da figura há o alinhamento de aminoácidos da proteína hipotética RasGEF1b murina e humana. Os domínios RasGEFN e RasGEF estão em cinza, itálico e sublinhados. O zíper de leucina presente no domínio RasGEFN está negrito sublinhado. Há três motivos de localização nuclear no domínio RasGEF: PVKKHR, PVSRLKK e PFERDRK, codificados pelos aminoácidos 276-282, 316-322 e 421-427, respectivamente, os quais estão no interior da caixa azul. Os aminoácidos GRV dentro da caixa vermelha indicam direcionamento para microcorpúsculos (36, modificada).

O programa Protparam computa várias propriedades físico-químicas que podem ser deduzidas de uma dada seqüência protéica (74). Segundo este programa, a proteína RasGEF1b possui 473 aminoácidos, apresenta peso molecular de 55273,4 Dáltons e ponto isoelétrico teórico 8.19. O ponto isoelétrico de uma proteína é o pH na qual o polipeptídeo tem carga elétrica nula e então não se move em um campo elétrico (86). A determinação do ponto isoelétrico é importante para os experimentos de isoeletrofocalização, que baseiam-se no ponto isoelétrico das moléculas para separá-las eletroforeticamente. Essa informação fornecida pelo programa foi utilizada para os experimentos de eletroforese bidimensional.

A meia-vida de uma proteína é uma predição do tempo gasto para que metade da quantidade da proteína na célula desapareça depois da sua síntese. O ProtPram relaciona a meia-vida de uma proteína com o aminoácido presente no seu N-terminal e fornece a predição de meia-vida para três organismos modelos (humano, *E.coli* e levedura). O papel do N-terminal originou-se de observações em que a identidade do resíduo N-terminal de uma proteína desempenha um papel importante na determinação da estabilidade *in vivo* (74). Resultados relacionados à meia-vida da RasGEF1b mostram que, em reticulócitos de mamífero, *in vitro*, a meia-vida é de 30 horas. Em leveduras seria superior a 20 horas, *in vivo*, e, em *Escherichia coli*, *in vivo*, superior a 10 horas. O índice de instabilidade foi computado como 47.38, o que classifica a proteína como instável.

O grau de hidropaticidade de um peptídeo ou proteína é calculado como a soma dos valores de hidropaticidade de todos os aminoácidos dividida pelo número de resíduos na seqüência. O valor fornecido de hidropaticidade para a seqüência da RasGEF1b é -0.439. Este valor atribui uma característica hidrofílica à RasGEF1b, sugerindo uma localização celular citoplasmática. Todos os resultados deste programa não levam em conta qualquer modificação pós-traducional anotada.

O peptídeo sinal está relacionado com o transporte e sinalização de proteínas para diferentes compartimentos celulares, além de discriminar entre proteínas solúveis e de membrana. A seqüência sinal está na porção amino-terminal da proteína e a marca para ser translocada para o lúmen do retículo endoplasmático. No lúmen do retículo endoplasmático, as proteínas recém-sintetizadas são modificadas de diversas formas. O programa SignalP utiliza redes neurais (neural networks, NN) e modelos Hidden Markov (HMM) em eucariotos (75). O programa fornece predições sobre presença de peptídeo sinal em uma dada seqüência protéica. Todos os “scores” SignalP3-NN fornecidos para a RasGEF1b foram muito baixos, característicos de proteínas não-secretadas. O modelo Hidden Markov calcula a probabilidade de uma seqüência submetida conter um peptídeo sinal ou não. O

modelo eucariótico HMM também relata a probabilidade de um sinal de âncora. A predição dada por este modelo para a RasGEF1b foi de uma proteína não-secretada, com probabilidade zero de conter peptídeo sinal e zero de conter sinal para âncora.

Após a remoção da sequência sinal, os peptídeos são enovelados, as pontes dissulfeto são formadas e muitas proteínas são glicosiladas. As glicoproteínas são geralmente ligadas aos oligossacarídeos através dos resíduos de asparagina e esses oligossacarídeos N-ligados são diversos. Depois de modificadas, as proteínas podem se direcionar para uma variedade de destinos intracelulares. No complexo de Golgi, oligossacarídeos O-ligados são adicionados e os N-oligossacarídeos são então modificados. Os programas utilizados forneceram várias predições de modificações pós-traducionais como glicosilação, fosforilação, sulfatação e miristilação. Estas modificações podem alterar a localização da proteína, influenciar na estabilidade e na sua conformação, na sua capacidade de interação com membranas ou domínios hidrofóbicos de outras proteínas e até mesmo na sua função (86).

Não há probabilidade de proteínas sem peptídeo sinal serem expostas à maquinaria de N-glicosilação e então não devem ser glicosiladas *in vivo*, embora haja motivos potenciais para este tipo de modificação. Segundo os programas Net N Glyc 1.0 server (76) e Net O Glyc 3.1 server (77), não há sítios potenciais para N ou O- glicosilação. O programa Yin O Yang 1.2 (78) forneceu predições para 5 sítios de modificações O- GlcNac nas posições 4, 190, 243, 278, 283 (na sequência da RasGEF1b de 473 aminoácidos). O programa NetAcet1.0 Server fornece predições para substratos de N-acetiltransferase A (NetA) (79). Não foram encontrados resíduos de serina, treonina, alanina ou glicina nas posições 1-3, necessárias para este tipo de modificação. Segundo o programa NetPhos 2.0 server, que fornece predições para sítios de fosforilação (80), há 11 sítios para fosforilação em serinas, 5 em treoninas e 9 em tirosinas. O Prosite forneceu sete possíveis sítios de fosforilação pela caseína cinase II (entre os aminoácidos 23-26, 94-97, 143-146, 168-171, 180-183, 214-217, 364-367) dois prováveis sítios de N-miristilação (41-46, 300-305), um sítio de sulfatação por tirosina (116-130) e um sítio de fosforilação por proteína cinase dependente de cGMP e cAMP (249-252) (Tabela 1 e Figura 4). As fosforilações não dependem da presença de peptídeo sinal e podem ainda ocorrer, dependendo da exposição espacial ao solvente dos prováveis sítios a serem fosforilados.

Tabela 1 – Possíveis modificações pós-traducionais da proteína RasGEF1b

Tipo de Modificação pós-traducional	Aminoácidos envolvidos	Posição dos aminoácidos na seqüência protéica
O-GlcNa	T T k k a	4 190 243 288 283
Fosforilação por Caseína cinase II	SaeD SegD SgeE SqyE TatD ThiE Ssre	23-26 94-97 143-146 168-171 180-183 214-217 364-367
Proteína cinase dependente de cAMP e GMP	RKkT	249-252
Fosforilação por Proteína cinase C	SsR TyR TdR SeR TaK SsR Twk SIR	69-71 148-150 182-184 247-249 329-331 364-366 414-416 464-466
Sulfatação	ewtftfpYdfrderm	116-130
N- miristilação	GSleAL GNfnSL	41-46 300-305

1 mpqtppfsam fdssgynrl yq**saed**scgg lyyhdnlls **gslealiqhl** vprnvdyypdr
 O-GlcNac CKII Phospho N-myristoylation
 61 tyiftilss rlfmhpylem akvchlcveh qrl**segd**gdk nqmrkiapki lqll**ewtet**
 121 fpydfrderm mmlkdlahr mas**gee**qtyr knvqqmmqcl irklaals**sqy** eevlaklss**t**
 Tyrosine sulfation site
 181 **atd**rltvk**t** kpqsiqrdim tvcsdpytla qql**thi**eler lnyigpeefv qafvqkdpld
 241 nd**k**scyse**rk** **kt**rnlleayve wfnrlsylv teicmpv**kkk** hrarmieyfi dvarecfni**g**
 CAMP_PHOSPHO_SITE
 301 **fnsl**maais gmnmspvslr kktwakvhta kfdilehqmd pssnfynyrtr alrgaaqrsi
 361 tah**ssre**kiv ipffsllikd iyflnegcvn rlpnghvnfe kfvelakqvs efmtwkqvec
 421 pferdrvlq yllsvpvfse dalylasyes egpenniekd rwklsrssl grv

Figura 4. Esquema da distribuição das modificações pós-traducionais preditas por programas de bioinformática para a proteína RasGEF1b. Em verde, O-glicosilação (O-GlcNac); em azul, fosforilação por caseína cinase II (CKII Phospho); em vermelho, N-miristilação (N-myristoylation) em laranja, sítio para fosforilação por proteína cinase dependente de cAMP e GMP e em cinza sítio para sulfatação (Tyrosine sulfation site).

O Clustal W é um programa com a finalidade geral de alinhamento múltiplo de sequências de DNA ou proteínas (81). Ele produz alinhamentos múltiplos biologicamente significativos de sequências divergentes. Este programa calcula a melhor correspondência para as sequências selecionadas e as alinha de forma que as identidades, as similaridades e as diferenças possam ser vistas. Utilizando este software, foi feito o alinhamento da sequência da proteína RasGEF1b entre diferentes espécies filogeneticamente distantes, para analisar a homologia entre elas (Figura 5). As sequências utilizadas foram derivadas de: camundongo, *Mus musculus*; homem, *Homo sapiens*; galinha, *Gallus gallus*; rã, *Xenopus tropicalis*; peixe-zebra, *Danio rerio*; ouriço-púrpura, *Strongylocentrotus purpuratus* e abelha, *Apis mellifera*. O alinhamento começa da sequência maior, de abelha, seguido das demais proteínas, todas alinhadas de acordo com a similaridade entre a sequência da RasGEF1b de *Mus musculus*. O alinhamento gerado foi utilizado pelo programa Phylip (58) para gerar uma árvore filogenética consenso. Como esperado, as sequências humana e de camundongo estão agrupadas em um ramo com um valor de bootstrap de 996 (Figura 6).

Foi feita, então, uma análise de domínios conservados através de domínios depositados no banco de dados do NCBI (82). As sequências das proteínas são comparadas entre si (utilizando blastp) e então são combinadas em grupos guiadas por um processo de similaridade. Todas as proteínas analisadas contém o domínio RasGEF, relacionado com a catálise da troca de nucleotídeos guanina da proteína Ras (Tabela 2).

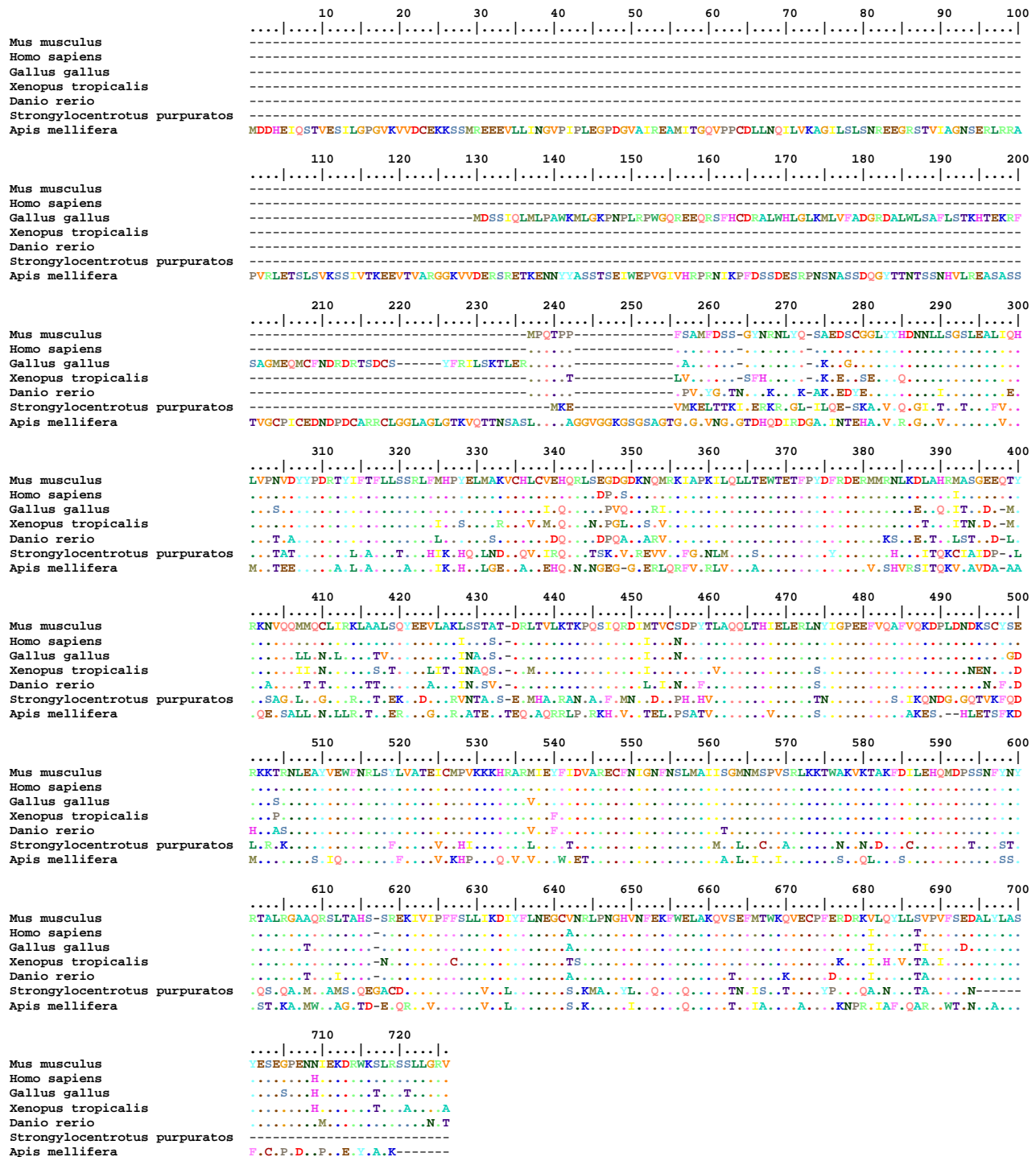


Figura 5. Alinhamento entre seqüências da proteína RasGEF1b derivadas de várias espécies filogeneticamente distintas utilizando o Clustal W. Por meio deste alinhamento é possível observar os aminoácidos conservados (pontos) e distintos (letras) desta proteína entre os diversos organismos. Cores diferentes representam aminoácidos diferentes.

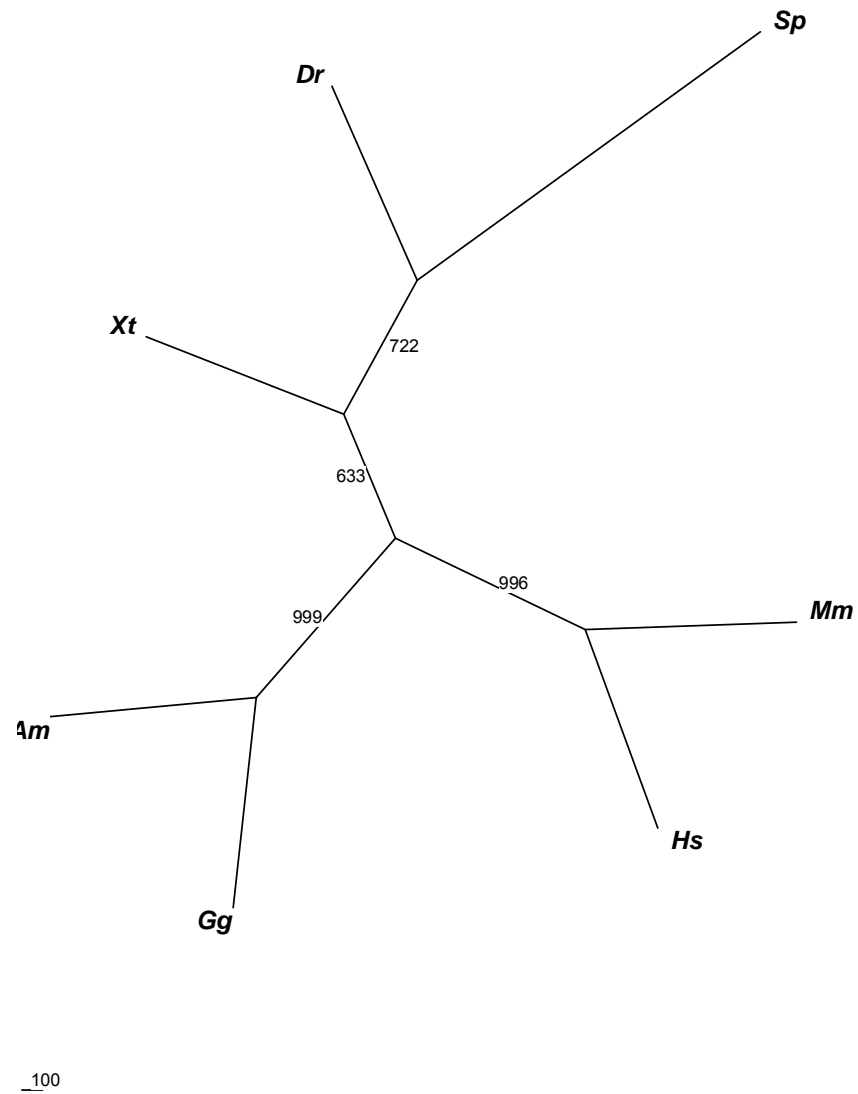


Figura 6. Filogenia da proteína RasGEF1b de espécies representativas evolutivamente distantes. Os valores numéricos nos ramos são valores de “bootstrap”. Am, *Apis mellifera*; Dr, *Danio rerio*; Gg, *Gallus gallus*; Hs, *Homo sapiens*; Mm, *Mus musculus*; Sp, *Strongylocentrotus purpuratus*; Xt, *Xenopus tropicalis*. -100, escala.

Tabela 2 – Estrutura de domínios conservados da proteína RasGEF1b entre diversas espécies

Espécie	Domínios conservados *
<i>Mus musculus</i>	
<i>Homo sapiens</i>	
<i>Gallus gallus</i>	
<i>Xenopus tropicalis</i>	
<i>Danio rerio</i>	
<i>Strongylocentrotus purpuratus</i>	
<i>Apis mellifera</i>	

* RasGEFN (domínio N-terminal fator de troca de nucleotídeos guanina associado à proteína Ras), RasGEF (domínio fator de troca de nucleotídeos associado à Ras).

Outros membros da família RasGEF têm suas seqüências de aminoácidos depositadas no banco de dados do NCBI. Uma vez que o estudo da proteína RasGEF1b iniciou-se com o estudo de cDNAs de macrófagos murinos, foi feito o alinhamento entre as seqüências de RasGEF1b e RasGEF1c murinas (Figura 7a). Apesar de a seqüência da RasGEF1a murina ainda não estar depositada no banco de dados, há a seqüência humana. Foi feito então, o alinhamento entre as

proteínas dos três membros RasGEF1a, RasGEF1b e RasGEF1c humanos para ser verificado o grau de homologia entre eles (Figura 7b). Este último alinhamento começa na maior proteína, RasGEF1c, seguida das demais. Todos os membros RasGEF analisados apresentam domínios RasGEFN e RasGEF altamente conservados (Tabela 3).

A)

```

      10      20      30      40      50      60      70      80      90     100
rasGEF1b  .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
rasGEF1c  ..R.LTA.D.VTPGSLSPPTTE.T.GEQA.QPLL.GAPS.A..DT.....TA.....EKA.....IE.R..L.R.....I.Q.Q.DKPVLL..

      110     120     130     140     150     160     170     180     190     200
rasGEF1b  NQMRKIAPKILQLLIEWTETFPYDFRDERMMRNLDLAHRMASGEEQTYRKNVQQMMQCLIRKLAALSQYEEVLAKLSSTATDRLTVLTKTPQSIQRDIM
rasGEF1c  ARV..FGA.L...A.....R..EE.STIGH.T.VVG.ISPCD.--GSR.H.LL.T.HQ...S.G.GP.S.VG----DKPISYRTPK.A..H.ELL

      210     220     230     240     250     260     270     280     290     300
rasGEF1b  TVCSPPYTLAQQLTHIELERLNIGPEEFVQAFVQKDPDLDNKSCYSEKKTNRLEAVVWVNRLSYLVATEICMPVKKKHRARMIEYFIDVARECFNIG
rasGEF1c  G.....V.....RH.....N.....AGT.PRP.D---N.V...K...C.....A..Q..QV..F.....

      310     320     330     340     350     360     370     380     390     400
rasGEF1b  NNSLMAIISGMNMSPVSRLLKKTWAKVKTKAKFDILEHQMDPSSNFYNYRTALRGAQRSLTAHSSREKIVIPFSLIKDIYFLNEGCNRLPNGHVNE
rasGEF1c  .....F.....TG..C.....H.....A.....

      410     420     430     440     450     460     470
rasGEF1b  KFWELAKQVSEFMTWKQVECPFERDKVILQVLLVVPVSEDALYLASVESEGPENNIKDRWKSLSRSLGRV
rasGEF1c  ..L.....G..I.....Q.PSITH..YTA.I...G.....S..SQT..E.....I..KT

```

B)

```

      10      20      30      40      50      60      70      80      90     100
rasGEF1a  .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
rasGEF1b  .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
rasGEF1c  WEPGGRAAAAAGGAPVRGAEARTGSCPWAGPGVPARRRPAD...LSA.D.VTP--GSLSPPTTEPTDGEQA---QPLL.GAPS.A..T.....

      110     120     130     140     150     160     170     180     190     200
rasGEF1a  T.....V...P.HD.L.R.GQI...KQOLEAGPE.AKLKSFSA..V..K...A...Q..KA.AE..AIT..VTQCD.ENG.V
rasGEF1b  PNVDVYDRTYIFTFLLSSRLFMHPV.ELMAKVCHLCVEHQRLSDPDSKQMRKIAPKILQLLIEWTETFPYDFRDERMMRNLDLAHRMASGEEQ--TV
rasGEF1c  .TA.....EKA.....IE.R..L.R.....I.Q.Q.DK.VL..ARV..FG..L...A.....R..QE.STIGH..VVG...P-CDE-A..

      210     220     230     240     250     260     270     280     290     300
rasGEF1a  K.ATA..T.S.LLS...R..LQ.LRE.LRPPAV.KGPI...PAA.K..LG..C..LV.....D.VSS.Y..DLM.IVSHM.S...HR..RGDL
rasGEF1b  RKNVQQMMQCLIRKLAALSQYEEVLAKISSTSDRLTVLTKTPQSIQRDIITVCNDPYTLAQQLTHIELERLNIGPEEFVQAFVQKDPDLDNKSCYSEK
rasGEF1c  .RMH.LL.A.HQ.....R.GP.G.VGADKP---ISYRTPK.A..H.ELLG..S.....V.....RH.....N.....AST.P.F.D..

      310     320     330     340     350     360     370     380     390     400
rasGEF1a  T..YS...DN...C..M...V.RV...T..L.F.....M.....L..A.....S.....V..H...C...
rasGEF1b  KKTNRLEAVVWVNRLSYLVATEICMPVKKKHRARMIEYFIDVARECFNIGNNSLMAIISGMNMSPVSRLLKKTWAKVKTKAKFDILEHQMDPSSNFYNYR
rasGEF1c  ..S.....K...C.....A..Q..QV..F.....R.....F.....TG..C.....

      410     420     430     440     450     460     470     480     490     500
rasGEF1a  ...Q..T...QM.N...V.N.FV...HKIHT.H...I..K...ISR.IH...T.....K.K..QS...A.IY..E..FV..F..
rasGEF1b  TALRGAQRSLTAHSSREKIVIPFSLIKDIYFLNEGCNRLPNGHVNEKFWELAKQVSEFMTWKQVECPFERDKVILQVLLVVPVSEDALYLASVE
rasGEF1c  .....H.....L.....G..I.....Q.AS TH..Y.A.I...G.....

      510     520
rasGEF1a  .....V...S..T..TT..N.A
rasGEF1b  SEGPENHIEKDRWKSLSRSLGRV
rasGEF1c  ..S...QT..E...A...I..KT

```

Figura 7. Alinhamento das seqüências de aminoácidos dos membros RasGEFa, b e c e esquema de domínios conservados. Em A, alinhamento entre RasGEFb e c murinos. Em B, alinhamento entre RasGEFa, b e c humanos.

Tabela 3 - Estrutura de domínios conservados entre RasGEFs A, B e C.

Proteína	Espécie	Representação em domínios conservados
RasGEF1b	<i>Mus musculus</i>	
RasGEF1c	<i>Mus musculus</i>	
RasGEF1b	<i>Homo sapiens</i>	
RasGEF1c	<i>Homo sapiens</i>	
RasGEF1a	<i>Homo sapiens</i>	

O esqueleto covalente de uma proteína típica contém centenas de ligações individuais. Devido à livre rotação possível em volta dessas ligações, a proteína pode assumir várias conformações. Entretanto, cada proteína tem uma função química ou estrutural específica, sugerindo fortemente que cada proteína tem uma única estrutura tridimensional.

O termo estrutura secundária se refere à conformação local de alguma parte de um polipeptídeo. Alguns poucos tipos de estrutura secundária são particularmente estáveis e ocorrem em uma grande variedade de proteínas. As mais proeminentes são as conformações α e β .

Na estrutura de α hélice o esqueleto polipeptídico é firmemente girado em torno de um eixo imaginário desenhado longitudinalmente no meio da hélice, e as cadeias laterais dos aminoácidos são dipostas do lado de fora da hélice. Na conformação β , o esqueleto da cadeia polipeptídica é estendido em forma de zigzag, e, quando as cadeias são dispostas desta forma lado a lado, podem formar várias pregas, originando o arranjo folha β pregueada (86).

A provável estrutura secundária da proteína RasGEF1b foi analisada com o programa HNN Secondary Structure Prediction Method (83). A predição fornecida mostra que a seqüência está organizada em α -hélices, folhas- β e voltas (coils), sendo que, a maioria dos aminoácidos está organizada estruturalmente em α -hélices (Figura 8).

O arranjo tridimensional de todos os átomos de uma proteína se refere à estrutura terciária desta proteína. Considerando altos níveis estruturais, as proteínas são classificadas em dois grandes grupos: proteínas fibrosas, as quais têm as cadeias polipeptídicas arranjadas em grandes folhas, consistindo de um único tipo de estrutura secundária; e proteínas globulares, que têm as cadeias polipeptídicas laterais enoveladas em uma forma esférica ou globular, frequentemente contendo vários tipos de estruturas secundárias. Utilizando o programa de modelagem Esypred, foi feita uma modelagem da proteína RasGEF1b baseando-se em um modelo de uma sequência de alta similaridade do Protein Data Bank (PDB). A proteína modelo utilizada para a modelagem foi o Sos (gi|4389274|pdb|1BKD|S), um RasGEF bem caracterizado, cuja estrutura cristalográfica foi feita junto com a proteína Ras, mostrando a interação entre as duas proteínas.

O Esypred constrói modelagens protéicas em três dimensões a partir de estruturas tridimensionais depositadas no Protein Data Bank, baseando-se no alinhamento entre sequências com maiores similaridades. A estrutura tridimensional final é construída usando o pacote de modelagem do Modeller (83). Utilizando este programa, construímos um modelo tridimensional hipotético para a RasGEF1b, destacando a organização das estruturas secundárias nesta proteína (Figura 9a). O modelo 1bkd cadeia S foi utilizado como molde para a modelagem. Este modelo apresenta identidade de 23% (similaridade 46%) sobre os 473 aminoácidos da RasGEF1b murina (Figura suplementar 1).

Também construímos um modelo de interação da RasGEF1b com a proteína Ras, baseado em um modelo existente criado a partir de dados cristalográficos (Figura 9b). Este modelo disponível representa a interação da proteína Ras com o GEF Sos (Figura 9c). Quando é feita a sobreposição do modelo da RasGEF1b com o modelo Sos ligado à Ras, os mesmos domínios da RasGEF1b parecem estar envolvidos na interação (Figura 9d). Para visualizarmos as modelagens, utilizamos o programa Pymol. O Pymol é um programa molecular de gráficos projetado para a visualização em tempo real e geração rápida de imagens moleculares de alta qualidade, gráficos e animações. O programa permite a visualização em 3D das proteínas, destacando densidade, superfícies, e rotações. Inclui também a edição, renderização, e produção de filmes moleculares. Editamos as modelagens construídas destacando a superfície das proteínas, particularmente os aminoácidos hidrofóbicos. Acreditamos que a interação RasGEF-Ras ocorra através destes aminoácidos e, pela visualização dos modelos observamos que há uma cavidade altamente

hidrofóbica (Figura 10a) onde a Ras provavelmente se liga (Figura 10b). Com um giro de cento e oitenta graus na molécula Ras “ligada”, é possível observar a presença de aminoácidos hidrofóbicos (Figura 10c), provavelmente envolvidos na interação.

A qualidade do modelo foi avaliada pelo gráfico de Ramachandran gerado pelo programa PROCHECK v.3.5.4. Este programa verifica a qualidade estereoquímica de uma estrutura protéica produzindo gráficos de plotagem, analisando a geometria do modelo como um todo e resíduo por resíduo (84). O gráfico de Ramachandram apresentou 87,5% de resíduos em regiões mais favoráveis, e apenas 0,2% dos resíduos em regiões não permitidas (ASN 455) (Figura 11), sendo assim considerado um bom modelo.

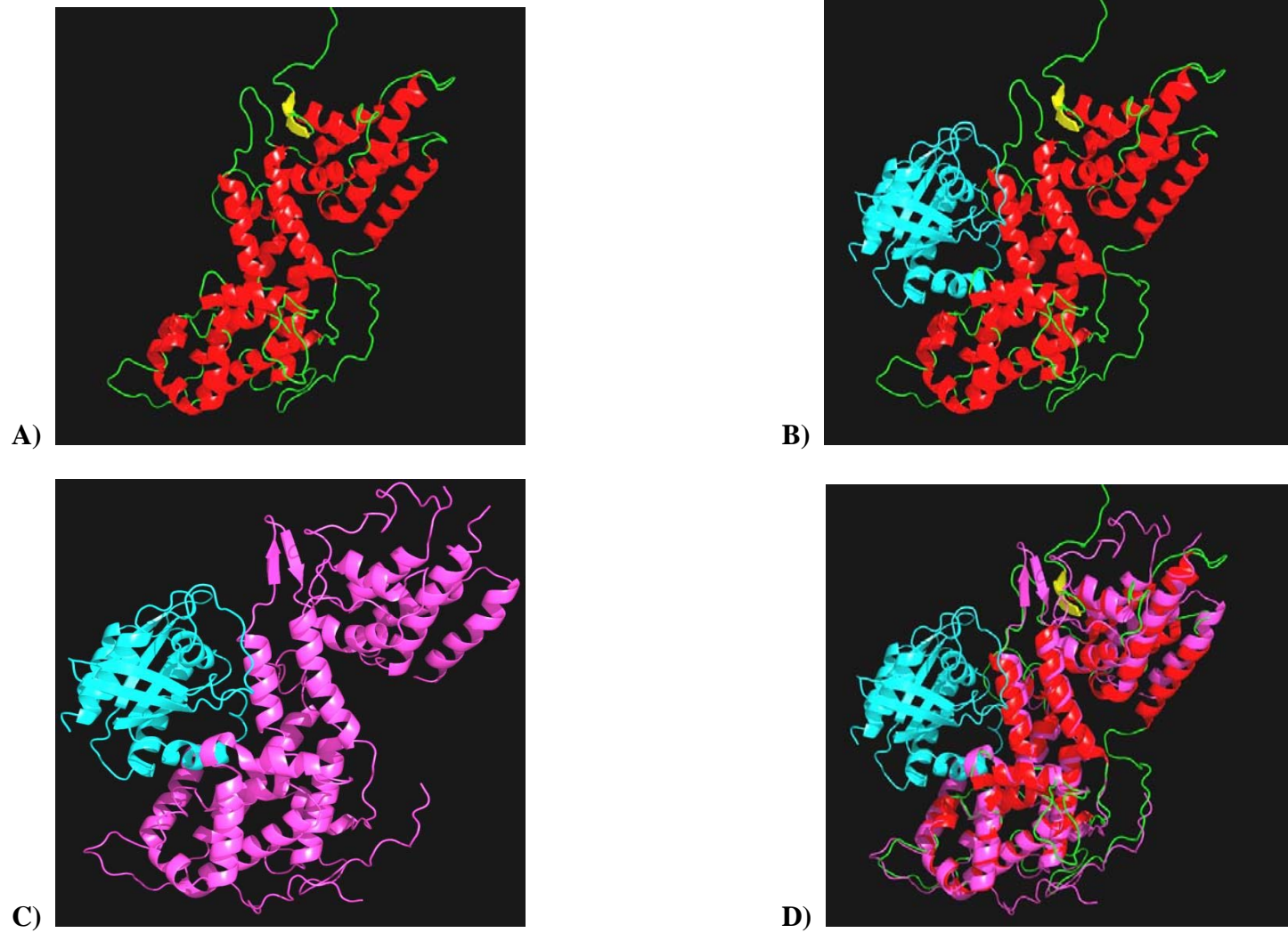


Figura 9. Modelos tridimensionais de estruturas terciárias destacando a organização das estruturas secundárias nas proteínas. (A) RasGEF1b; (B) RasGEF1b ligada à Ras; (C) Sos ligada à Ras; (D) Sobreposição de RasGEF1b à Sos ligada à Ras.

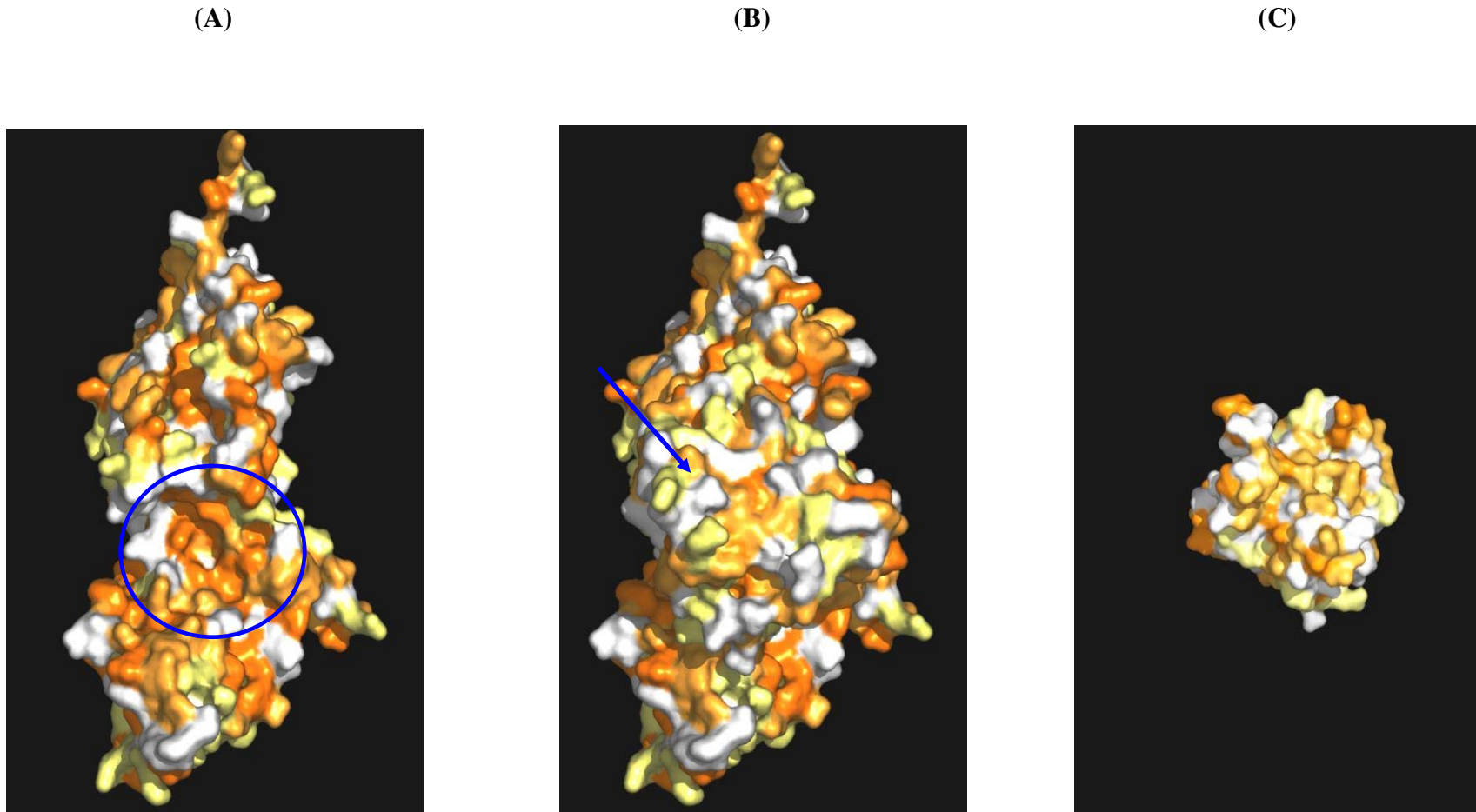


Figura 10. Modelagens tridimensionais de superfície destacando hidrofobicidade. Os aminoácidos estão divididos em altamente hidrofóbicos (laranja), hidrofóbicos (laranja claro), não hidrofóbicos (amarelo) e restantes (branco). (A) RasGEF1b e em destaque no círculo azul o sulco hidrofóbico onde provavelmente a Ras se encaixa; (B) RasGEF1b ligada à Ras (seta); (C) Vista da porção da proteína Ras que provavelmente interage com o sulco da RasGEF1b.

A região do Sos (massa molecular de cerca de 150 kDa) que é funcional para a troca de nucleotídeos guanina na Ras abrange cerca de 500 resíduos e contém blocos de sequências que são conservados em outros fatores de troca de nucleotídeos específicos para Ras, tais como o CDc25, Sdc25 e o fator de liberação de nucleotídeos guanina em Ras (GRF) (39). A estrutura do modelo refere-se ao complexo da H-Ras humana (Human Harvey-Ras, resíduos 1-166) com a região de fator de troca de nucleotídeos guanina do Sos1 humano (resíduos 568 a 1044) cristalizado. O domínio catalítico da Sos apresenta um sulco hidrofóbico e ao redor dele há um arranjo de interações polares e carregadas entre Sos e Ras (99).

A partir dos modelos existentes e gerados, foi feita uma análise sobre os possíveis aminoácidos da proteína Sos que interagem com a Ras e o seu equivalente estrutural da RasGEF1b baseada no alinhamento das seqüências de aminoácidos das proteínas Sos e Ras. Por se tratar de uma análise inicial, restringimos a busca para interações entre átomos de nitrogênio e oxigênio, com as distâncias entre os átomos variando de 0.0 a 3.0 Å e de 0.0 a 4.0 Å (Tabela 4). As interações que podem ocorrer nesta distância são pontes de hidrogênio entre as moléculas. Como não se sabe com qual Ras a RasGEF1b interage, uma conclusão sobre os aminoácidos responsáveis pela interação entre as duas moléculas poderia ser errônea.

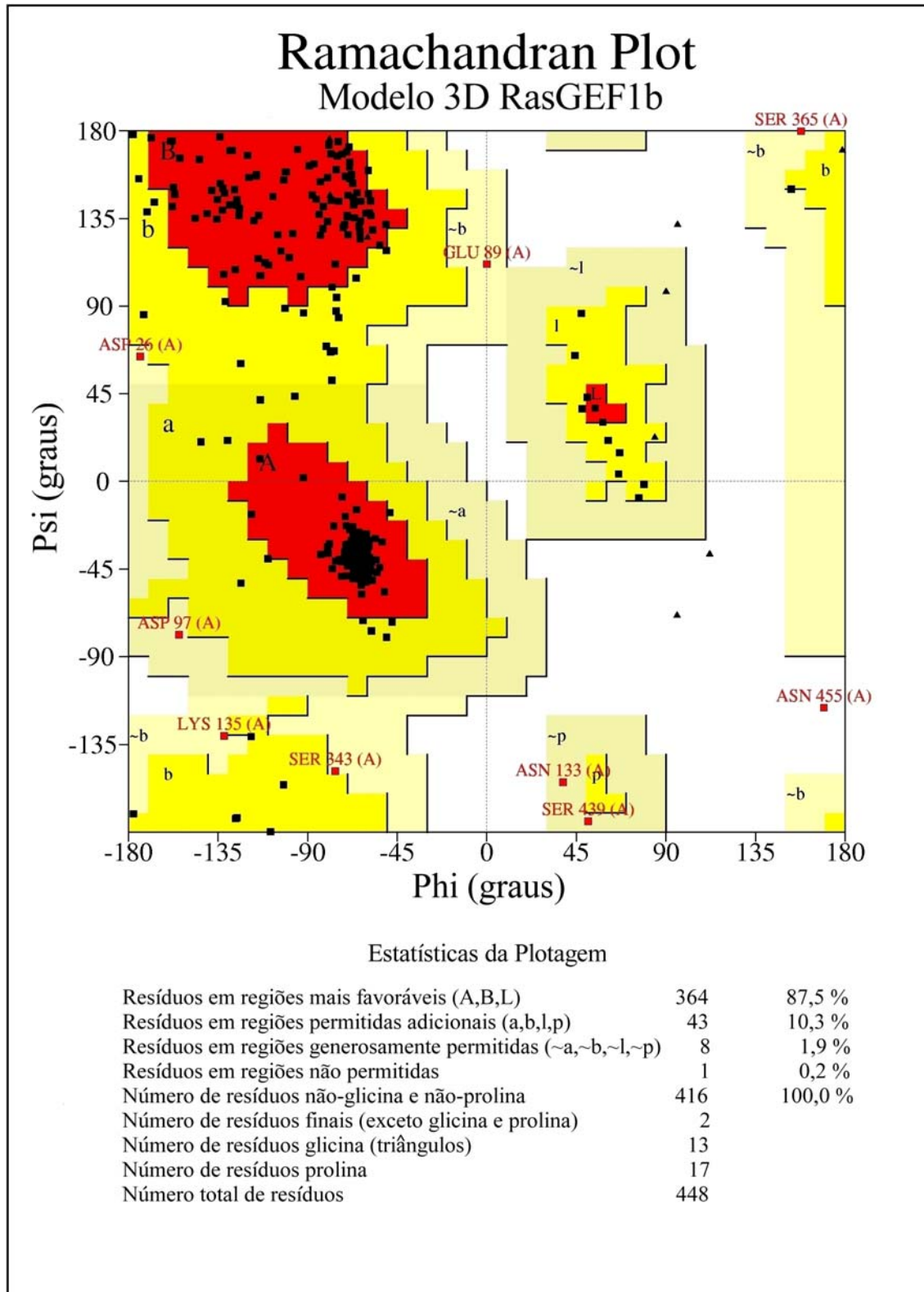


Figura 11. Ramachandran Plot do modelo tridimensional da proteína RasGEF1b. O gráfico representa a disposição dos aminoácidos ao longo de um gradiente que varia de regiões favoráveis (vermelho) a regiões não permitidas (branco). O único resíduo da região não permitida é a Asparagina de número 455.

Tabela 4 – Lista de contatos entre átomos de oxigênio e hidrogênio entre a proteína Ras e a Sos e o equivalente estrutural em RasGEF1b

Lista de contatos - átomos de N e O (modelo 1BKD: Sos(S) + Ras(R))			Equivalente estrutural em RasGEF1b (Alinhamento da estrutura primária)
distância: 0.0 - 3.0 Angstroms			
Sos	Ras	Distância*	
Trp 809S NE1	Gly 60R O	2.86 ***	Val 234
Asn 879S O	Arg 73R NH1	2.41 ***	Asn 274
His 911S NE2	Asp 54R OD1	2.90 ***	Phe 345
Tyr 912S OH	Tyr 71R OH	2.79 ***	Tyr 346
Thr 935S OG1	Ala 59R O	2.84 ***	Lys 379
	Gln 61R NE2	2.70 ***	Lys 379
Glu 942S OE1	Ser 17R OG	2.61 ***	Glu 386
Asn 944S OD1	Tyr 32R N	2.75 ***	-
Glu 1002S OE2	Arg 68R NH2	2.83 ***	-
Asp 1007S OD1	Arg 102R NH2	2.70 ***	Asp 440
distância: 3.0 - 4.0 Angstroms			
Sos	Ras	Distância	
Arg 826S NE	Glu 63R OE1	3.16 ***	Glu 260
Arg 826S NH2	Glu 63R OE1	3.74 *	Glu 260
	Glu 63R OE2	3.73 *	Glu 260
Thr 829S OG1	Glu 63R O	3.34 *	Asn 263
Ser 876S N	Gln 70R OE1	3.57 *	Ser 271
Ser 876S OG	Gln 70R OE1	3.02 ***	Ser 271
Ser 881S N	Asp 69R OD2	3.08 ***	Ser 276
Ser 881S OG	Asp 69R OD1	3.83 *	Ser 276
	Asp 69R OD2	3.46 *	Ser 276
	Val 103R N	3.99 *	Ser 276
Ser 908S OG	Gln 70R NE2	3.68 *	Ser 342
Asp 910S O	Tyr 40R OH	3.68 *	Asn 344
Lys 913S NZ	Glu 37R OE2	3.08 ***	Gln 357
Phe 930S N	Tyr 64R OH	3.99 *	Phe 374
Gly 931S N	Tyr 64R OH	3.40 *	Phe 375
Gly 931S O	Gln 61R NE2	3.48 *	Phe 375
Thr 935S N	Ala 59R O	3.97 *	Lys 379
Thr 935S OG1	Thr 58R O	3.58 *	Lys 379
Asn 936S ND2	Pro 34R O	3.02 ***	Asp 380
Lys 939S NZ	Asp 57R O	3.01 ***	Phe 383
Gly 943S O	Asp 30R O	3.01 ***	Gly 387
	Tyr 32R N	3.98 *	Gly 387
	Gln 25R NE2	3.52 *	Gly 387
Asn 944S OD1	Asp 30R O	3.65 *	-
	Tyr 32R O	3.46 *	-
Asn 944S ND2	Tyr 32R O	3.11 ***	-
Glu 1002S OE2	Arg 68R NH1	3.69 *	-
Lys 1003S NZ	Gln 95R OE1	3.40 *	-

* Distância calculada em Angstroms (Å)

Utilizando o modelo tridimensional da RasGEF1b construído verificamos a localização espacial e a área de acessibilidade ao solvente dos aminoácidos sujeitos à fosforilação por cinases preditos por programas de bioinformática (Tabela 5). Alguns destes aminoácidos estão no interior da molécula, e mesmo sendo sítios potenciais de fosforilação não têm acesso ao solvente e acessibilidade à maquinaria de fosforilação (Figura 12a). Outros aminoácidos apresentam a área de acessibilidade ao solvente baixa, porém, dependendo da orientação da cadeia lateral, o aminoácido pode adquirir uma conformação na qual a hidroxila (OH) esteja mais exposta ao solvente (12 b). Os aminoácidos que apresentam área de acessibilidade ao solvente alta estão expostos na superfície da molécula e podem facilmente ser fosforilados (Figura 12c). Segundo estes critérios de acessibilidade os aminoácidos receberam diferentes probabilidades de serem de fosforilados (tabela 5).

Tabela 5 - Probabilidades de fosforilações inferidas a partir do modelo tridimensional

Cinase	Aminoácido a ser fosforilado	Área acessível ao solvente*	Probabilidade de ocorrer a fosforilação**
Caseína cinase II	23 (Serina)	24,9	Alta
	94 (Serina)	41,5	Alta
	143 (Serina)	1,2	Indefinida
	168 (Serina)	42,8	Alta
	180 (Treonina)	37,4	Alta
	214 (Treonina)	0,2	Muito baixa
	364 (Serina)	23,0	Indefinida
Cinase dependente de cAMP e cGMP	252 (Treonina)	0	Muito baixa
Cinase C	69 (Serina)	0,5	Muito baixa
	148 (Treonina)	25,9	Indefinida
	182 (Treonina)	0,2	Indefinida
	247 (Serina)	3,5	Muita baixa
	329 (Treonina)	42,0	Alta
	364 (Serina)	23,0	Indefinida
	414 (Treonina)	40,7	Alta
	464 (Serina- não há no modelo construído)	---	Indefinida

* Área calculada em Angstrom ao quadrado (Å^2).

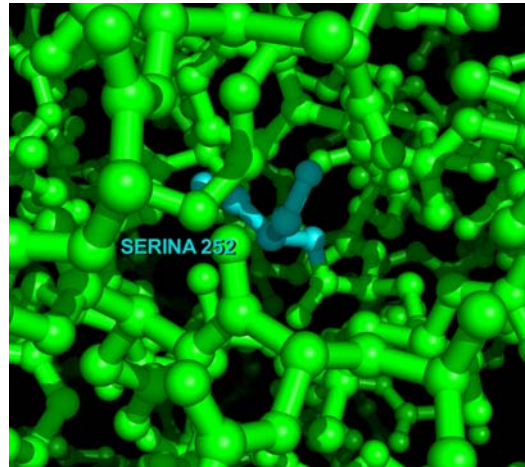
** Probabilidade inferida considerando a exposição do aminoácido na superfície da molécula e sua acessibilidade à maquinaria de fosforilação. O tamanho da enzima não é considerado.

Probabilidade Alta: a área acessível ao solvente é considerável.

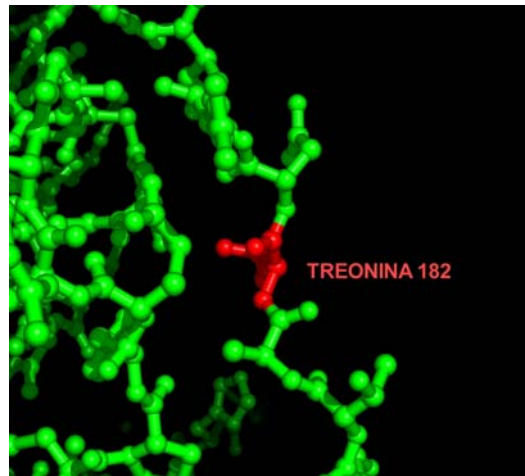
Indefinida: dependendo da rotação que o aminoácido adquirir na molécula ele pode adquirir uma conformação na qual a hidroxila (OH) esteja mais exposta e disponível ao solvente.

Muito baixa: o aminoácido está enterrado na molécula (a hidroxila não está exposta), sendo pouquíssimo provável ocorrer a fosforilação.

A)



B)



C)

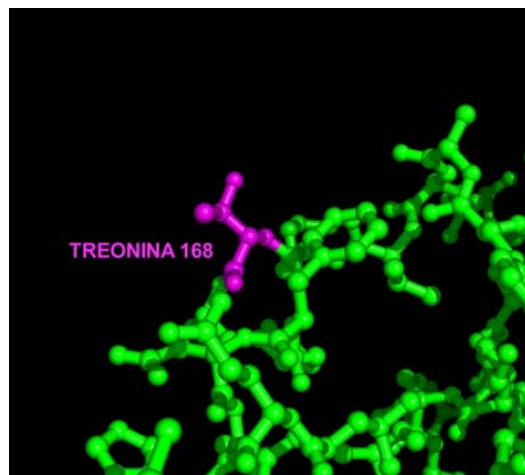


Figura 12. Exemplos de aminoácidos com diferentes probabilidades de serem fosforilados. A) Serina 252 com baixa probabilidade; B) Treonina 182 com probabilidade indefinida e C) Treonina 168 com alta probabilidade.

5.2. Expressão do mRNA do *rasgef1b* em células Raw 264.7 estimuladas com poli I:C ou LPS

Utilizando a técnica de RT-PCR, foi feita a análise da expressão do *rasgef1b* em células Raw 264.7 estimuladas com LPS e poly-rI:rC (agonistas de receptor do tipo Toll). Os resultados indicam que a expressão do mRNA do *rasgef1b* é estimulada em 2 e 4 horas após tratamento das células com cada um dos agonistas, quando comparadas com células não estimuladas (0h) (Figura 13).

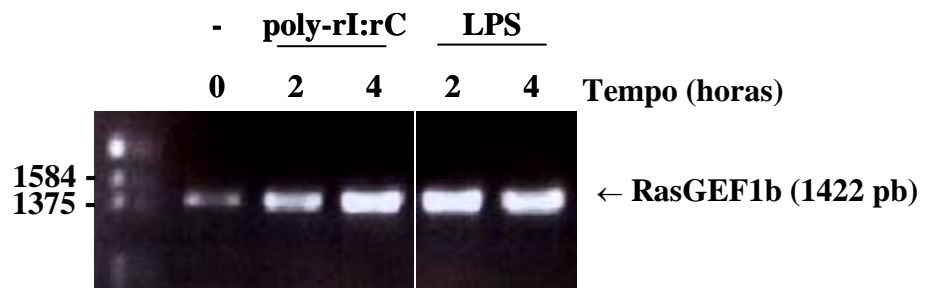


Figura 13. Análise, por RT-PCR, da expressão do *rasgef1b* em células RAW264.7 (macrófagos murinos de linhagem contínua) estimuladas com RNA de fita-dupla (poly-rI:rC, 100µg/ml) ou endotoxina bacteriana (LPS, 100ng/ml). O RNA total extraído das células, nos tempos indicados, foi transcrito reversamente e à primeira fita cDNA a qual foi utilizada como molde nas reações de PCR (28 ciclos) para a amplificação com iniciadores específicos para o cDNA correspondente ao mRNA do *rasgef1b*. Após fracionamento dos produtos de PCR, o gel de agarose foi corado em brometo de etídio e as bandas visualizadas em luz ultravioleta.

5.3. Amplificação do cDNA do *rasgef1b* de macrófagos murinos

Com o objetivo de expressar o *rasgef1b* recombinante em bactérias, utilizou-se PCR para amplificar a região codificadora do gene *rasgef1b*, gerando-se um produto de 1422 pb. As bandas mais intensas observadas no gel de agarose, corado com brometo de etídio, foram aquelas obtidas utilizando como molde o cDNA obtido de RNA extraído de macrófagos tratados com GPI-mucinas e o de macrófagos estimulados com o parasita *T. cruzi* (forma tripomastigota) (Figura 14). Após a clonagem desses fragmentos no vetor pCR[®] TOPO 2.1, foi verificado, por PCR do DNA extraído das colônias obtidas, que duas colônias selecionadas continham o cDNA do *rasgef1b*. As PCRs com o DNA extraído das colônias de bactérias transformadas foram realizadas com os mesmos iniciadores da primeira amplificação (dados não mostrados).

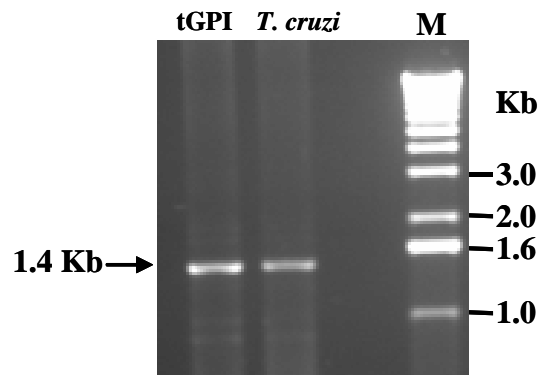


Figura 14. Amplificação da região codificadora do *rasgef1b* (seta) utilizando como molde cDNAs de macrófagos murinos estimulados com glicosilfosfatidilinositol derivados da membrana de *T. cruzi*, forma tripomastigota (tGPI) ou estimulados com o parasita *Trypanosoma cruzi*. M, marcador de tamanho molecular (kb).

5.4. Perfil de restrição do cDNA do *rasgef1b* clonado no vetor TOPO 2.1

Para verificar a presença de sítios de restrição no cDNA do *rasgef1b* clonado, foram feitas digestões enzimáticas dos plasmídeos extraídos de colônias positivas (5 e 6). Os fragmentos gerados pelas enzimas de restrição estão representados na Figura 15 (a e b). A digestão do plasmídeo obtido da colônia 6 com a enzima *Bam*H I forneceu dois fragmentos correspondentes aos dois sítios de restrição previstos para essa enzima presentes na posição 759 e 309 pb da sequência codificadora do gene, com base na comparação com o marcador de tamanho molecular utilizado. As digestões do DNA plasmidial da colônia 6 com as enzimas *Eco*R I, *Hind* III, *Sac* I, *Kpn* I e *Xho* I também forneceram fragmentos compatíveis com sítios presentes na sequência do *rasgef1b*, a saber: *Eco*R I, 1500 pb; *Hind* III, 1482 pb; *Sac* I, 281 pb; *Kpn* I, 205 pb e *Xho* I, 513 pb (comparação com o marcador de tamanho molecular utilizado). O controle positivo das reações de digestão foi o plasmídeo pcDNA3.1/V5-His (Invitrogen) contendo o cDNA do *rasgef1b* clonado, gentilmente cedido pelo Dr. Aristóbolo Mendes da Silva.

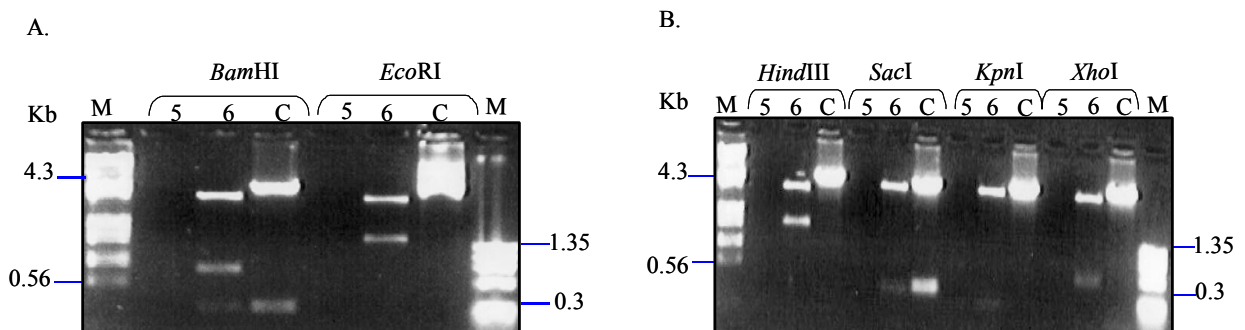
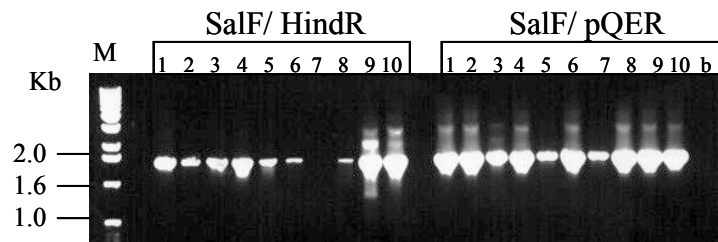


Figura 15. Perfil de restrição enzimática do cDNA do *rasgef1b* clonado no pCR[®] TOPO 2.1. As digestões feitas com o DNA plasmidial da colônia 6 forneceram fragmentos com tamanho molecular esperado de acordo a análise da presença dos sítios de restrição das enzimas utilizadas para as digestões na sequência do *rasgef1b*, depositada no Banco de dados do NCBI. M, marcador de tamanho molecular; C, cDNA do *rasgef1b* clonado no pcDNA3.1/V5-His.

5.5. Clonagem do gene *rasgef1b* no vetor pQE-30 e sequenciamento do produto gerado

O vetor pQE-30 foi escolhido para a clonagem do cDNA do *rasgef1b* por permitir a expressão em bactérias de uma proteína recombinante em fusão com resíduos de histidina de maneira a facilitar posteriormente sua purificação. Os iniciadores utilizados na amplificação da região codificadora do *rasgef1b*, para a clonagem no pQE-30, contém sítios para as enzimas de restrição *Sal* I (GCTGTCGACGGGAAAGTATGC) e *Hind* III (TCCAAAGCTTAAACTCTGCCC). O DNA plasmidial extraído da colônia 6 e do vetor pQE-30 foram digeridos com essas enzimas, possibilitando a formação de extremidades coesivas para a ligação entre o inserto e o vetor. Após a ligação, foi feita a transformação em bactérias competentes. Deste experimento, foram escolhidas, aleatoriamente, 10 colônias para testes por PCR. Foram utilizados dois pares de iniciadores para as reações. Um dos pares continha os dois iniciadores utilizados na amplificação inicial e o outro continha o iniciador que se anela no inserto clonado e outro no vetor pQE-30 (Figura 16a). Três colônias positivas nos testes de PCR foram escolhidas para a extração plasmidial em pequena escala. Outros testes, utilizando 4 combinações entre os pares de iniciadores que flanqueiam a região codificadora do gene e o par de iniciadores que flanqueiam o sítio múltiplo de clonagem do pQE-30, foram realizados e as três colônias mostraram-se novamente positivas (Figura 16b). O DNA plasmidial obtido da colônia 2 foi parcialmente seqüenciado utilizando os iniciadores SalF e HindR. Duas seqüências obtidas com cerca de 800 nucleotídeos, cada, foram agrupadas, gerando o consenso mostrado na Figura 6. A análise da seqüência consenso obtida utilizando o programa BLAST mostra uma similaridade de 96,7% do gene *rasgef1b* humano e murino (Figura 17).

A.



B.

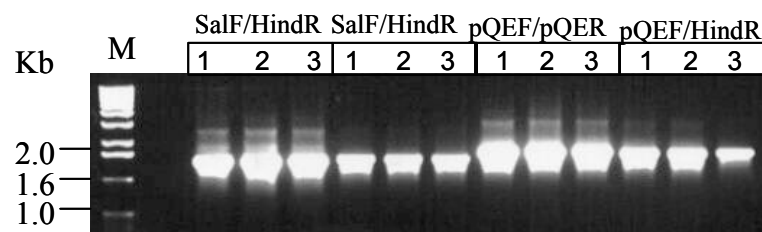


Figura 16. a) Teste das colônias derivadas da clonagem do cDNA do *rasgef1b* no pQE-30 com os iniciadores SalF e HindR e com o SalF e o pQE-30 reverso (pQER). b) Teste, por PCR, das três colônias recombinantes utilizando os iniciadores da clonagem inicial e os iniciadores que se anelam no vetor pQE-30 (pQEF e pQER).

```

QUERY      -----GGGACATCATGACCCGTCTGCAGCGACC--TACACATTGGCCCA 42
SBJCT      CCCACAGTCCATACAGCGGGACATCATGACC-GTCTGCAGCGACCCTTACACATTGGCCCA 779
          *****

QUERY      GCANCTGACCCACATCGACCTGGAGANACTCAATTACATTGGCCAGAGAGTTTGTTC 102
SBJCT      GCAGCTGACCCACATCGACCTGGAGAGACTCAATTACATTGGCCAGAGAGTTTGTTC 839
          *****

QUERY      GGCATTTGTGCAGAAGGATCCTTTGGACAACGATAAGAGTTGCTACAGCGAACGGAAGAA 162
SBJCT      GGCATTTGTGCAGAAGGATCCTTTGGACAACGATAAGAGTTGCTACAGCGAACGGAAGAA 899
          *****

QUERY      AACACGGAACCTGGAGGCGTACGTAGAAATGGTTTAAATCGTCTAAGCTACTTGGTTGTCTAC 222
SBJCT      AACACGGAACCTGGAGGCGTACGTAGAAATGGTTTAAATCGTCTAAGCTACTTGGTTGTCTAC 959
          *****

QUERY      AGAAATCTGCATGCCTGTGAAGAAGAAGCACCCGAGCGAGGATGATCGAGTATTCATCGA 282
SBJCT      AGAAATCTGCATGCCTGTGAAGAAGAAGCACCCGAGCGAGGATGATCGAGTATTCATCGA 1019
          *****

QUERY      TGTGGCTCGAGAGTGTTC AACATTGGAACTTCAATTCCTTGATGGCAATCATCTCTGG 342
SBJCT      TGTGGCTCGAGAGTGTTC AACATTGGAACTTCAATTCCTTGATGGCAATCATCTCTGG 1079
          *****

QUERY      TATGAATATGAGCCGGTCTCGAGACTTAAGAAAACCTGGGCCAAAGTGAAGACTGCAAA 402
SBJCT      TATGAATATGAGCCGGTCTCGAGACTTAAGAAAACCTGGGCCAAAGTGAAGACTGCAAA 1139
          *****

QUERY      GTTTGACATTCCTTGAGCATCAGATGGATCCATCCAGCAATTTCTACAACCTACAGAACGGC 462
SBJCT      GTTTGACATTCCTTGAGCATCAGATGGATCCATCCAGCAATTTCTACAACCTACAGAACGGC 1199
          *****

QUERY      TCTCCGAGGGGCGAGCGCAAAGGTCCCTGACTGCTCACAGTAGCAGAGAGAAGATTGTGAT 522
SBJCT      TCTCCGAGGGGCGAGCGCAAAGGTCCCTGACTGCTCACAGTAGCAGAGAGAAGATTGTGAT 1259
          *****

QUERY      ACCATTCTTCAGTCTCTTAATCAAAGACATTTATTTCCCTCAACGAGGGCTGTGTCAACCG 582
SBJCT      ACCATTCTTCAGTCTCTTAATCAAAGACATTTATTTCCCTCAACGAGGGCTGTGTCAACCG 1319
          *****

QUERY      TCTTCCAAATGGCCATGTCAACTTTGAGAAATTTTGGGAACCTGGCCAAACAAGTGAGTGA 642
SBJCT      TCTTCCAAATGGCCATGTCAACTTTGAGAAATTTTGGGAACCTGGCCAAACAAGTGAGTGA 1379
          *****

QUERY      ATTTATGACATGGAACAGGTGGAGTGTCCGTTCCGAGCGGGACCGGAAGGTCTTACAATA 702
SBJCT      ATTTATGACATGGAACAGGTGGAGTGTCCGTTCCGAGCGGGACCGGAAGGTCTTACAATA 1439
          *****

QUERY      CCTGCTCTCCGTGCTGTATTTCAGTGAAGATGCGCTCTACTTG--CTTCATACGAGAGTGA 761
SBJCT      CCTGCTCTCCGTGCTGTATTTCAGTGAAGATGCGCTCTACTTGGCTTCATACGAGAGTGA 1499
          *****

QUERY      AGGACCCGAAAATAATATAGAAA--GACAGATGGAAGTCTTTAAGGTCCAGCCTTTTGGG 819
SBJCT      AGGACCCGAAAATAATATAGAAAAGGACAGATGGAAGTCTTTAAGGTCCAGCCTTTTGGG 1559
          *****

gi|22749128|ref|NM_152545.1| Homo sapiens GPI-gamma 4 (GPIG... 815 0.0
gi|21928185|gb|AY129963.1| Mus musculus strain C3H GPI-gamm... 1550 0.0

```

Figura 17. Sequenciamento da região codificadora do *rasgef1b* clonado no vetor pQE-30. O alinhamento indica uma homologia entre a sequência do *rasgef1b* humana e de camundongo depositadas no Banco de Dados do NCBI.

5.6. Expressão da proteína RasGEF1b recombinante em sistema procarioto

A célula hospedeira escolhida para a expressão da proteína recombinante foi a bactéria *Escherichia coli*, de linhagem denominada XL1-Blue, baseado nas características convenientes de cultura dessa cepa, sistemas de expressão gênica bem desenvolvidos e a habilidade da produção da proteína necessária em larga escala. Entretanto, como em qualquer sistema de expressão procariótico, não há modificações pós-traducionais na proteína recombinante (87-90).

As colônias de número 1 e 2 foram escolhidas aleatoriamente para indução da expressão da proteína recombinante com IPTG. A região codificadora do gene da proteína dihidrofolato redutase clonado no vetor pQE-40 foi utilizado como controle positivo. Essa proteína recombinante, quando expressa, contém seis histidinas na porção aminoterminal. O vetor pQE-30 sem inserto foi utilizado como controle negativo. Na ausência de IPTG, não houve expressão de nenhuma das proteínas recombinantes. Quatro horas após indução com IPTG, tanto o controle positivo, quanto as duas colônias de RasGEF1b apresentaram as respectivas proteínas expressas. Dezoito horas após indução não houve alteração aparente no nível da RasGEF1b recombinante (Figura 18a). As massas das proteínas His-RasGEF1b e dihidrofolatoredutase recombinante comparadas com o padrão utilizado foram confirmadas com as massas teóricas dessas proteínas. No caso da dihidrofolatoredutase a massa teórica esperada é 26 kDa, e para a RasGEF1b 56 kDa, somadas a massa correspondente às seis histidinas (~1 kDa).

Para verificar se a banda correspondente à proteína induzida corresponde à RasGEF1b recombinante ligada às seis histidinas, foi feito *Western Blotting* com anticorpo monoclonal anti-histina. Também foi utilizada, nesse experimento, a colônia contendo o gene da proteína recombinante dihidrofolatoredutase como controle positivo. O anticorpo reconheceu as histidinas das proteínas analisadas (Figura 18b). O reconhecimento da His-RasGEF1b pelo anti-histidina confirma a expressão dos seis aminoácidos na mesma janela de leitura da proteína RasGEF1b, possibilitando, então, a posterior purificação da proteína recombinante com sistema de quelato de níquel, que apresenta alta afinidade pelas histidinas.

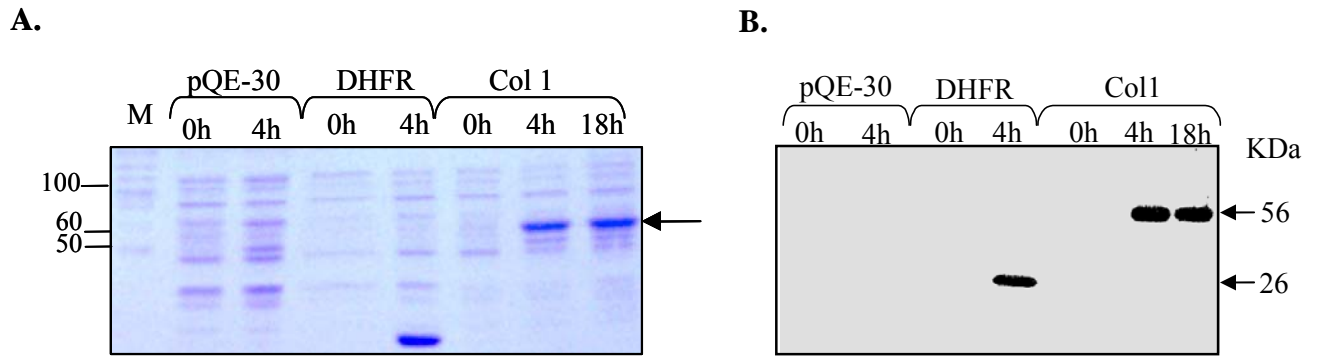


Figura 18. Indução da expressão da His-RasGEF1b. A) Os extratos totais bacterianos da colônia 1 (Col1), do pQE-30 (vetor vazio), da DHFR (dihidrofolato redutase clonada no pQE-40) e o marcador de massa molecular (M) foram fracionados em SDS-PAGE, posteriormente corado com Coomassie blue. Houve expressão da proteína recombinante His-RasGEF1b às 4 e 18 horas após indução com IPTG (seta). B) Western blot da colônia 1 utilizando anticorpo monoclonal anti-histidina. O anticorpo primário de camundongo IgG₁ anti-histidina foi utilizado na diluição de 1:4000 e o anticorpo secundário IgG de cabra anti-IgG camundongo conjugado com peroxidase foi diluído 3000 vezes. A revelação foi feita através de quimioluminescência.

5.7. Solubilização da proteína His-RasGEF1b recombinante produzida em *E. coli*

Após expressar a proteína His-RasGEF1b e confirmar sua expressão com um anticorpo monoclonal anti-histidina, testamos a sua solubilidade. A presença da His-RasGEF1b foi constatada na fração insolúvel dos extratos de bactérias XL1-Blue induzidas por IPTG. Não se pode prever, com exatidão, quais proteínas serão expressas de forma solúvel em *E. coli*. A formação de corpos de inclusão pode ser influenciada por vários fatores, incluindo a natureza da proteína, a célula hospedeira, as condições de crescimento e indução e o nível de expressão resultante da escolha do vetor.

O teste de solubilidade foi realizado para verificar se His-RasGEF1b estava presente na fração solúvel ou insolúvel do extrato das proteínas totais bacterianas. Com o tampão de lise utilizado, contendo ou não Tween[®] 20 na concentração de 0,25%, a proteína recombinante permaneceu no sedimento, no qual estavam as proteínas insolúveis (Figura 19). Foi feita a re-extração por sonicação e centrifugação, porém a proteína recombinante permaneceu insolúvel (2B₂, resultados não mostrados). Este experimento permitiu definir que a purificação da RasGEF1b deveria ser sob condições desnaturantes.

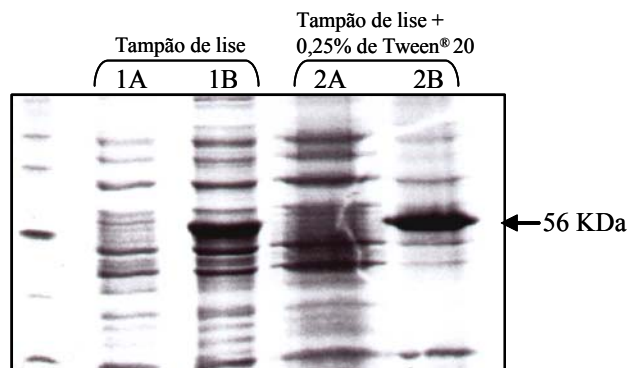
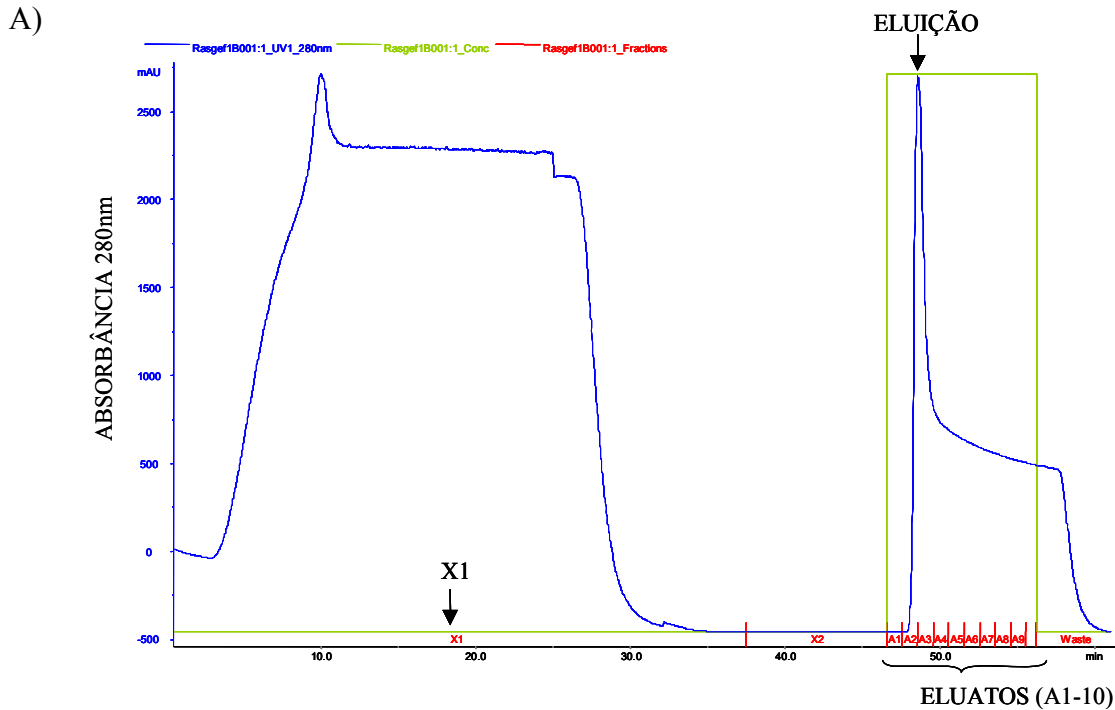


Figura 19. Solubilização da His-RasGEF1b. As amostras do teste de solubilidade 1A, 1B, 2A e 2B foram fracionadas em SDS-PAGE 12%, corado com Coomassie blue. O sedimento do tubo 2 (2B) foi solubilizado em tampão de lise contendo 0,25% de Tween[®] 20. A proteína His-RasGEF1b permaneceu na fração insolúvel do extrato protéico total bacteriano (seta).

5.8. Purificação da proteína His-RasGEF1b recombinante por cromatografia de afinidade

Após ser verificada a presença da His-RasGEF1b na fração insolúvel do extrato protéico das bactérias, partimos para a sua purificação, sob condições desnaturantes, em coluna HiTrap™ Chelating contendo níquel. As proteínas com um ou mais resíduos de histidinas se ligam na coluna de níquel com uma afinidade maior do que as interações antígeno–anticorpo ou enzima-substrato. O método cromatográfico de afinidade utilizado para a purificação da His-RasGEF1b forneceu bons resultados, uma vez que a proteína recombinante foi detectada em grandes quantidades em algumas frações eluídas da coluna. Entretanto, a desanilização feita posteriormente acarretou uma perda significativa de proteína.

Foram utilizados 500 mL de cultura bacteriana contendo a construção pQE-30*rasgef1b* para a expressão da proteína His-RasGEF1b com a finalidade de purificá-la. Como não houve alteração nos níveis de expressão da proteína recombinante após 4h, a indução com IPTG foi feita durante este período. Posteriormente, as bactérias foram lisadas com tampão contendo uréia, para solubilizar a proteína de interesse. A purificação da proteína recombinante foi feita através de uma coluna de afinidade contendo níquel e para a posterior troca de solvente (desalting) utilizou-se uma coluna de gel filtração. O primeiro cromatograma obtido representa a passagem da amostra pela coluna de afinidade, onde é possível visualizar o pico referente ao início da eluição e sua absorbância correspondente (Figura 20a). A proteína foi eluída com o tampão fosfato contendo 500mM de imidazol, concentração que faz com que as proteínas que têm histidinas ligadas ao níquel da coluna sejam deslocadas e coletadas. Os eluatos obtidos (A1 a A10) continham a proteína recombinante, o que foi observado em SDS-PAGE (Figura 20b). A partir de meio litro de cultura bacteriana foi possível obter cerca de 1mg de proteína recombinante. Foi feita a desalinização para a troca de tampões desses eluatos. O tampão de eluição contendo 8M de uréia foi trocado por PBS 1X (Figura 21a). As frações obtidas dessa desanilização foram fracionadas em SDS-PAGE para ser verificado em qual delas está presente a proteína His-RasGEF1b (21b). Após a desanilização foi verificada uma perda de até 80% comparada ao rendimento inicial.



B)

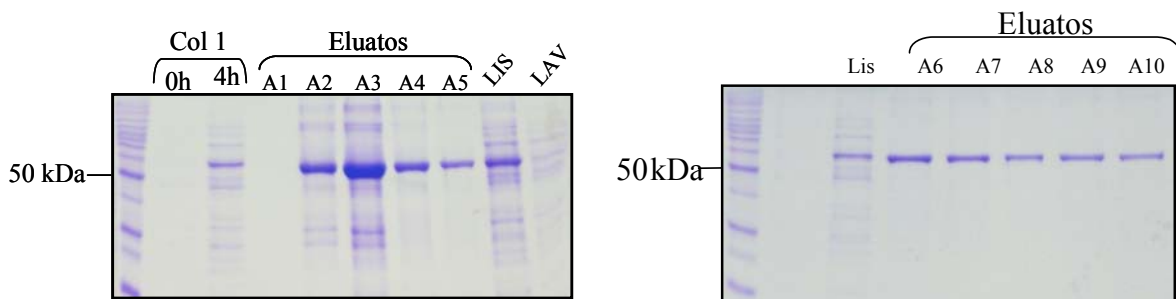
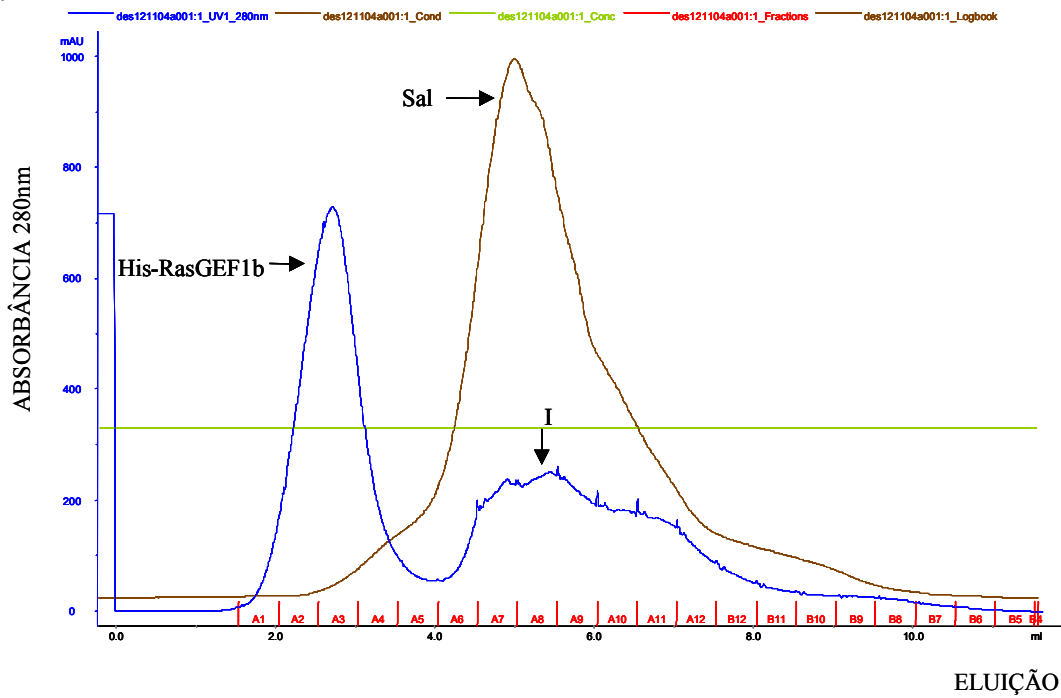


Figura 20. Purificação da proteína His-RasGEF1b recombinante por cromatografia de afinidade. a) O cromatograma obtido da purificação da proteína recombinante com a coluna HiTrapTM Chelating mostra a entrada do lisado total na coluna e a eluição das proteínas não ligadas (X1), e o pico correspondente à absorbância das frações purificadas. b) Fracionamento dos eluatos A1 a A10 em SDS-PAGE corado com Coomassie blue. A proteína His-RasGEF1b está presente em todos os eluatos conforme mostram as figuras (painel da esquerda eluatos 2 a 5, painel da direita, eluatos 6 a 10). Foram fracionados, juntamente com os eluatos, alíquotas das proteínas totais bacterianas da colônia 1 nos tempos 0 e 4h, o lisado total que foi aplicado na coluna de afinidade (LIS), e a fração obtida da etapa de lavagem da coluna durante a purificação (LAV).

A.



B.

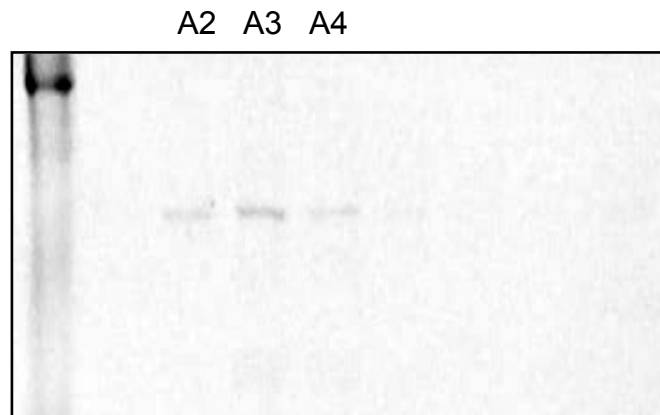
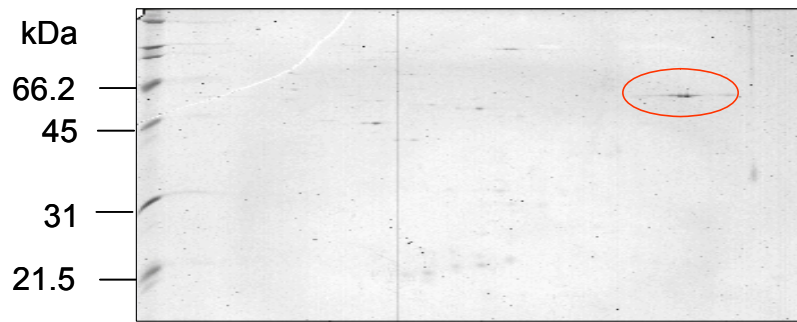


Figura 21. Desanilização do eluato A4 originado da purificação por cromatografia de afinidade. A) O cromatograma mostra que durante a passagem do eluato A4 pela coluna de gel filtração a proteína His-RasGEF1b é liberada nas primeiras frações de eluição, livre do tampão contendo uréia (sal) e imidazol (I), que são liberados posteriormente. B) As amostras de 1 a 9 obtidas do “desalting” do eluato A4 foram fracionadas em SDS-PAGE. A proteína His-RasGEF1b foi eluída primeiramente, sendo visível nas frações em A2, A3 e A4.

5.9. Eletroforese bidimensional e análise por espectrometria de massa da proteína His-RasGEF1b purificada

Na eletroforese bidimensional (2-D PAGE) (91), as proteínas são separadas de acordo com a carga pela focalização isoelétrica (IEF) na primeira dimensão e com a massa molecular na segunda dimensão. O ponto isoelétrico teórico da RasGEF1b é 8,19; o qual sofre variação para 8,25 com a adição dos seis aminoácidos histidina. A proteína RasGEF1b recombinante foi submetida à separação em gel 2D (Figura 22a) e o “spot” referente à esta proteína foi confirmado por Western blotting (Figura 22b). Esse spot específico foi retirado do gel, reduzido, carboxiamidado e digerido com tripsina. Depois da digestão, a amostra foi analisada por espectrometria de massa associada à desorção a laser assistida por matriz. O resultado dessa análise gerou sequências de peptídeos correspondentes (fingerprint) aos fragmentos da proteína RasGEF1b (Figura 23). Através da fragmentação dos peptídeos gerados pela digestão com tripsina, as massas dos peptídeos na mistura foram precisamente medidas utilizando o MALDI Tof-Tof para produzir uma impressão digital das mesmas. Isso permitiu encontrar o tamanho total da proteína RasGEF1b recombinante e a sua identificação em bancos de dados (Figura 24).

A)

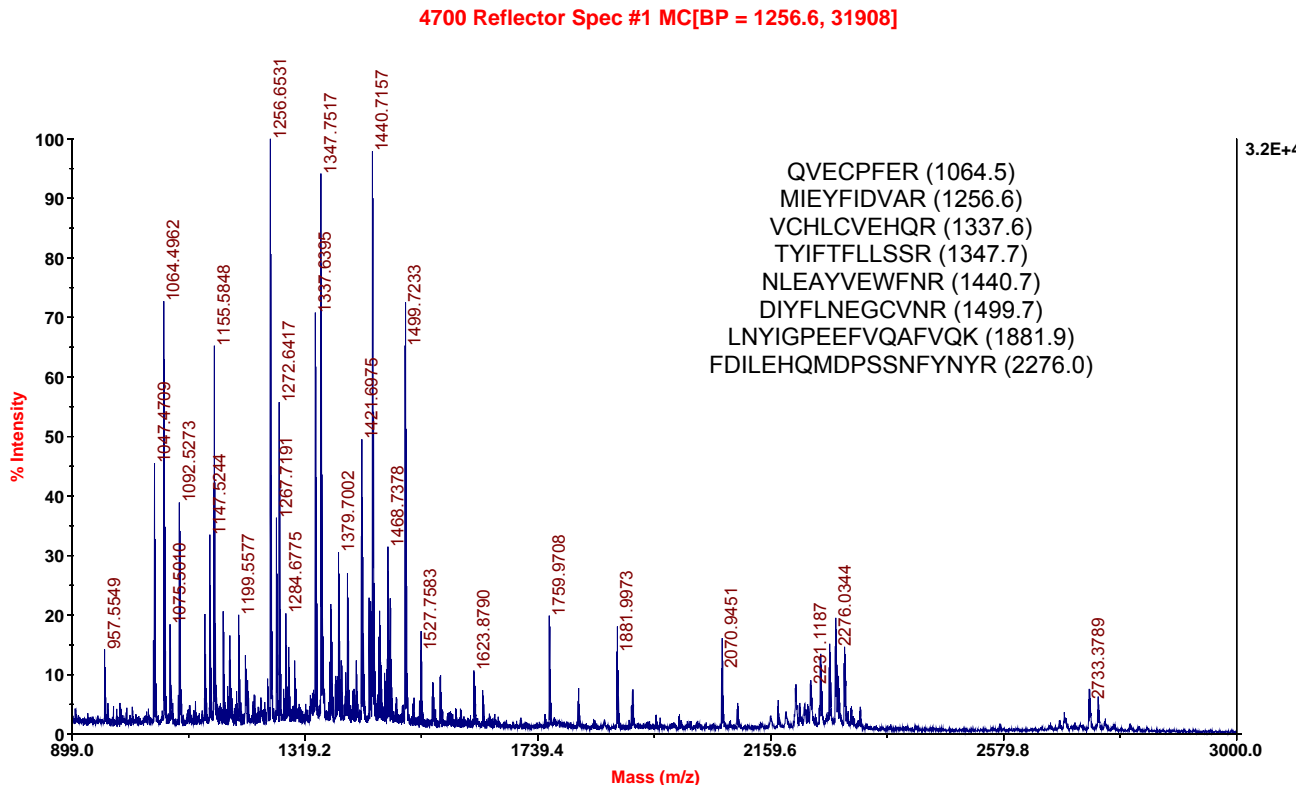


B)



Figura 22. Eletroforese bidimensional da proteína RasGEF1b recombinante obtida a partir da fração A3. A amostra foi separada na primeira dimensão numa faixa de pI de 3 a 10 e na segunda dimensão por SDS-PAGE 12%. As massas moleculares relativas estão indicadas na parte esquerda da figura (kDa). O gradiente de pH é não-linear e a orientação é 3-10 (esquerda para direita). (A) Gel SDS-PAGE corado com Coomassie blue coloidal com a proteína RasGEF1b dentro do círculo vermelho. (B) Western Blotting utilizando um anticorpo monoclonal anti-histidina.

A)

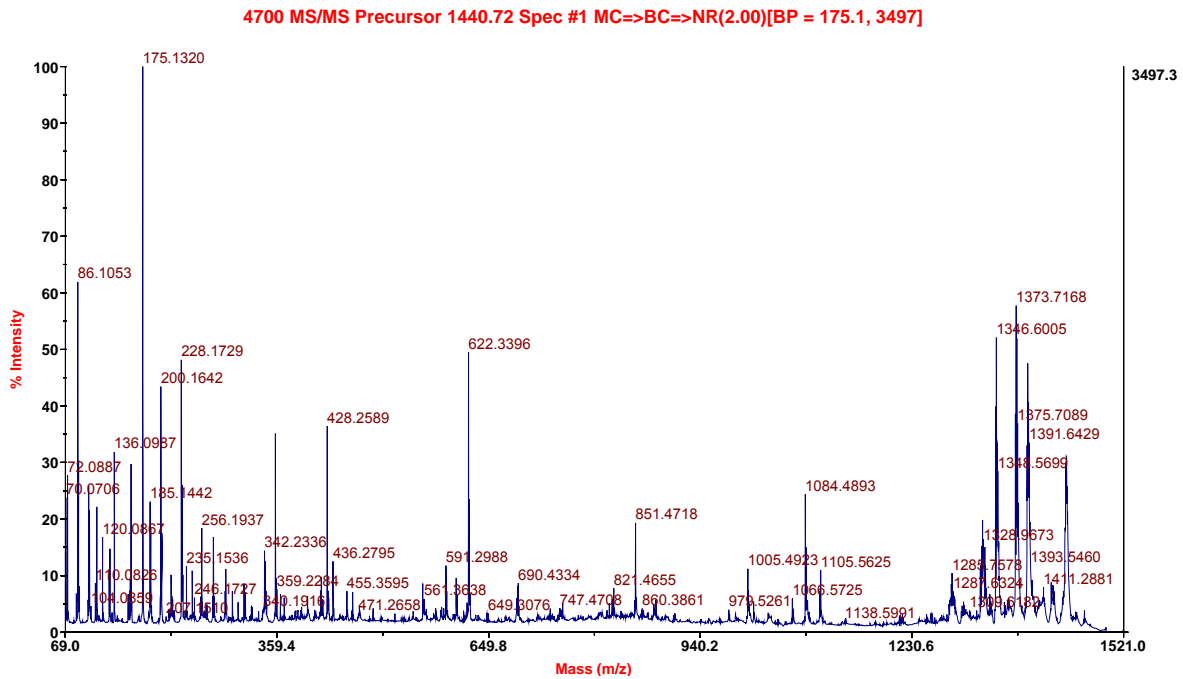


B)

mpqtppfsamfdssgynrnlyqsaedscgglyyhdnllsgslealiqhlvpnvdyypdrtyiffllssrlf
 mhpyelmakvchlcvehqrlsdpsdsknqmrkiapkilqlltewtetfpydfrdermmrnkdlahrias
 geeqtyrkvnvqqmmqclirklaalsqyeevlakisststdrlvtlktkpqsiqrdiitvcndpytlaqqllthiel
 erlnyigpeefvqafvqkdpldndkscyserkktrnleayvewfnrlsylvateicmpvkkkhrarmieyfi
 dvarecfnignfnslmaiisgmnm spvsrlkktwakvktakfdilehqmdpssnfynrtalrgaaqrslt
 ahsrekivipffsllikdiyflnegcvnrlpnghvnfekfwelakqvsefntwkqvecpferdrkilqylltpv
 vfsedalylysesegpenhiekrwkslrssllgrv

Figura 23. Análise por espectrometria de massa do “spot” correspondente à proteína RasGEF1b (Figura 21b). (A) Espectro dos peptídeos gerados pela tripsinização do “spot” correspondente à proteína recombinante RasGEF1b. (B) Os peptídeos seqüenciados formam o “fingerprinting” da RasGEF1b recombinante e estão sublinhados na seqüência de aminoácidos da proteína.

A)



B)

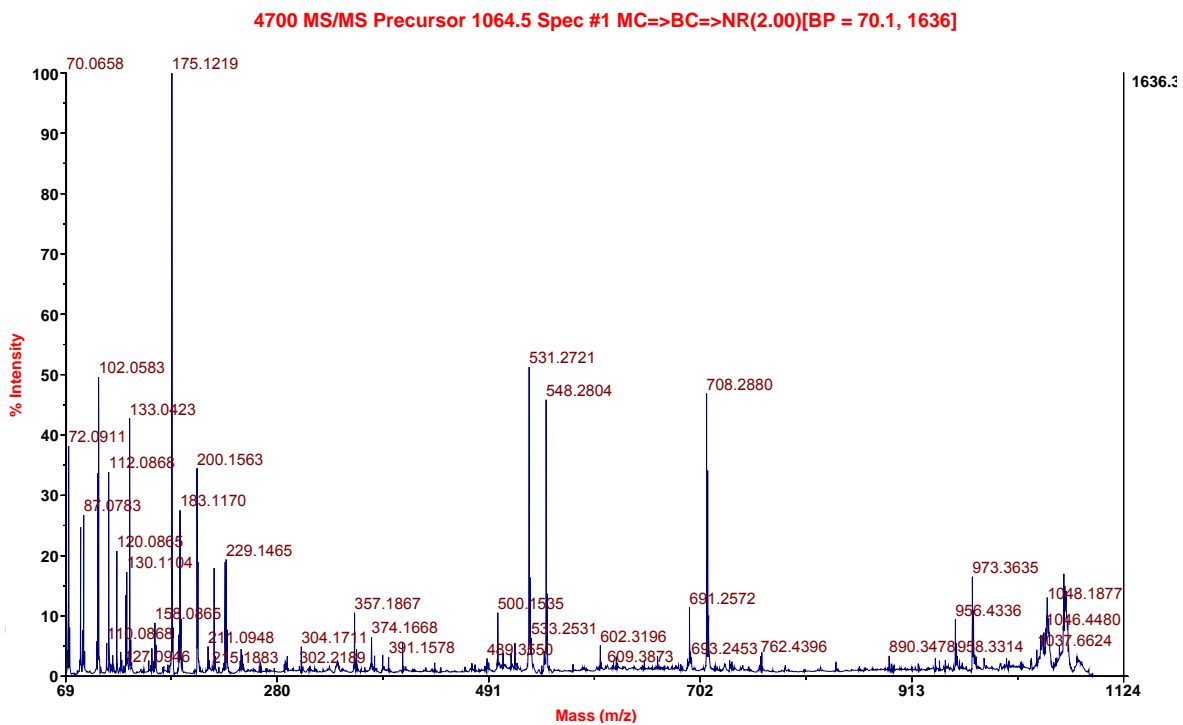


Figura 24. Os peptídeos da primeira análise por espectrometria de massa (Figura 23a) são seqüenciados em uma segunda análise MS/MS. Em A, MS/MS do fragmento precursor de massa 1440,72 Dáltons; em B, MS/MS do fragmento precursor de massa 1064,5 Dáltons.

5.10. Produção de anticorpos anti-His-RasGEF1b em coelho e galinha.

O sangue de coelhos e os ovos das galinhas foram coletados imediatamente antes da primeira e 15 dias após a terceira imunização. Os anticorpos foram purificados de acordo com os protocolos descritos em Materiais e Métodos e um ensaio de ELISA foi realizado para a análise da produção de anticorpos anti-His-RasGEF1b. Não houve produção de anticorpos contra a proteína recombinante, tanto em coelhos quanto em galinhas como pode ser visto a partir das leituras das absorbâncias antes e após as imunizações (Tabelas 6 e 7 respectivamente).

Tabela 6 - ELISA com soro dos coelhos

Diluição	Pré imune			Imune		
	Adjuvante	50µg	150µg	Adjuvante	50µg	150µg
1:20	0.247	0.492	0.092	0.939	0.266	0.115
1:40	0.126	0.188	0.042	0.825	0.123	0.066
1:80	0.067	0.115	0.027	0.643	0.061	0.038
1:160	0.037	0.076	0.002	0.400	0.018	0.026
1:320	0.021	0.021	0.000	0.196	0.003	0.009
1:640	0.008	0.012	0.003	0.097	0.007	0.008
1:1280	0.002	0.003	0.003	0.032	0.001	0.002
Branco	0.003	0.000	0.001	0.001	0.001	0.002

Tabela 7 - ELISA com anticorpos das galinhas

Diluição	Pré imune						Imune					
	Adjuvante		50µg		150µg		Adjuvante		50µg		150µg	
1:20	1.121	0.840	1.373	1.404	1.060	1.061	0.964	0,94	1.468	1.840	1.105	1.023
1:40	0.736	0.492	0.756	0.834	0.599	0.629	0.671	0.637	0.936	0.826	0.696	0.825
1:80	0.388	0.250	0.492	0.528	0.390	0.387	0.312	0.302	0.495	0.418	0.456	0.514
1:160	0.220	0.181	0.312	0.312	0.272	0.254	0.203	0.227	0.324	0.221	0.306	0.313
1:320	0.145	0.141	0.189	0.189	0.158	0.166	0.148	0.148	0.196	0.115	0.203	0.274
1:640	0.130	0.113	0.123	0.114	0.122	0.119	0.115	0.117	0.128	0.125	0.138	0.199
1:1280	0.101	0.168	0.097	0.095	0.133	0.098	0.085	0.128	0.147	0.129	0.100	0.146
Branco	0.144	0.066	0.177	0.072	0.086	0.078	0.075	0.085	0.087	0.097	0.084	0.113

Diluição: representa a titulação dos anticorpos dos animais

Branco: representa os poços controles sem anticorpo dos animais

Adjuvante: representa soros de animais que foram imunizados somente com adjuvante

Imune: representa soros de animais imunizados com a proteína His-RasGEF1b

50µg e 150µg: representam as doses de proteínas utilizadas nas imunizações dos animais das colunas “Imune”.

5.11. Localização subcelular da proteína FLAG-RasGEF1b em sistema eucarioto através de centrifugação diferencial e eletroforese bidimensional

O desafio da análise proteômica está na tarefa de examinar grande número de proteínas mantendo alta sensibilidade para a detecção de proteínas com baixo número de cópias. Não há passo de amplificação para proteínas, análogo ao método da reação em cadeia da polimerase, para amplificar DNA ou RNA. Isso significa que proteínas muito abundantes, sem fracionamento e enriquecimento das amostras biológicas, mascaram as proteínas pouco abundantes. A maioria das proteínas regulatórias como as cinases, fosfatases ou GTPases existem em baixo número de cópias, mas com localização subcelular bem específica (92).

As células HEK 293T transfectadas com pFLAG*Grasgef1b* foram lisadas com tampão de lise (10mM KCl) e então centrifugadas a $700 \times g$ por 10 minutos para sedimentar o núcleo e debris celulares. O sobrenadante foi centrifugado novamente a $10,000 \times g$ por 25 minutos para obter o sedimento de membranas pesadas. O sobrenadante correspondente foi centrifugado novamente a $100,000 \times g$ para gerar o sedimento de membranas leves e o sobrenadante citosólico. Anticorpos monoclonais foram utilizados para proteínas específicas das diferentes frações: α -histona-4 para fração nuclear, α -calreticulina para membranas pesadas, α -rab4 para membranas leves e α -actina para o citosol. Utilizando anticorpo monoclonal anti-FLAG, foi demonstrado que a proteína FLAG-Rasgef1b está presente, principalmente, nas frações de núcleo e de membranas pesadas, como indicado pela forte reatividade com anti-histona e anti-calreticulina, respectivamente (Figura 25).

Dois principais problemas têm impedido o desenvolvimento de procedimentos rápidos e padronizados para fracionamento subcelular. Primeiro, diferentes compartimentos celulares compartilham propriedades físicas similares e co-fracionam pelo menos em alguma extensão em gradientes convencionais. Segundo, as culturas de tecidos são mais comumente utilizadas para o fracionamento, porque as células podem ser mais facilmente manipuladas, diferente dos tecidos animais. Entretanto, depois da homogeneização, as culturas de células são mais difíceis de fracionar do que a maioria dos tecidos, presumidamente devido às diferenças na organização do citoesqueleto. A completa purificação é, com poucas exceções, dificilmente possível (93).

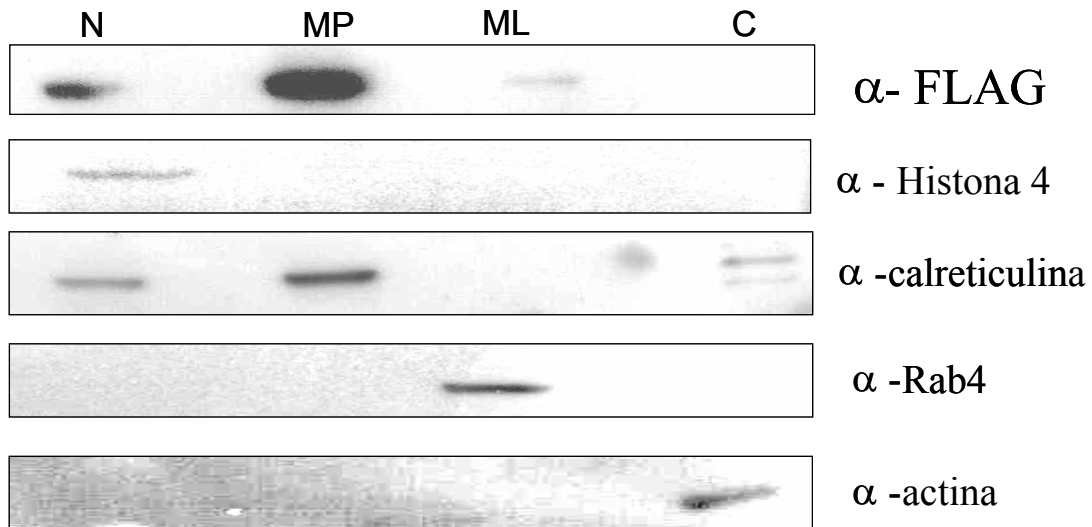


Figura 25. Fracionamento subcelular de células HEK293T expressando a proteína FLAG-RasGEF1b utilizando o método de centrifugação diferencial. N (núcleo); MP (membranas pesadas); ML (membranas leves); (C) citósol.

Após ser determinado que a proteína FLAG-RasGEF1b está localizada principalmente na fração de membranas pesadas, esta fração foi submetida à eletroforese bidimensional e Western blotting com anticorpo monoclonal anti-FLAG (Figura 26). O ponto isoelétrico (PI) da proteína RasGEF1b em fusão com o peptídeo FLAG foi determinado como sendo 7.14. Esta alteração de PI se deve à inclusão do peptídeo FLAG (metionina- ácido aspártico- tirosina- lisina- ácido aspártico- ácido aspártico- ácido aspártico- lisina), que possui caráter ácido.

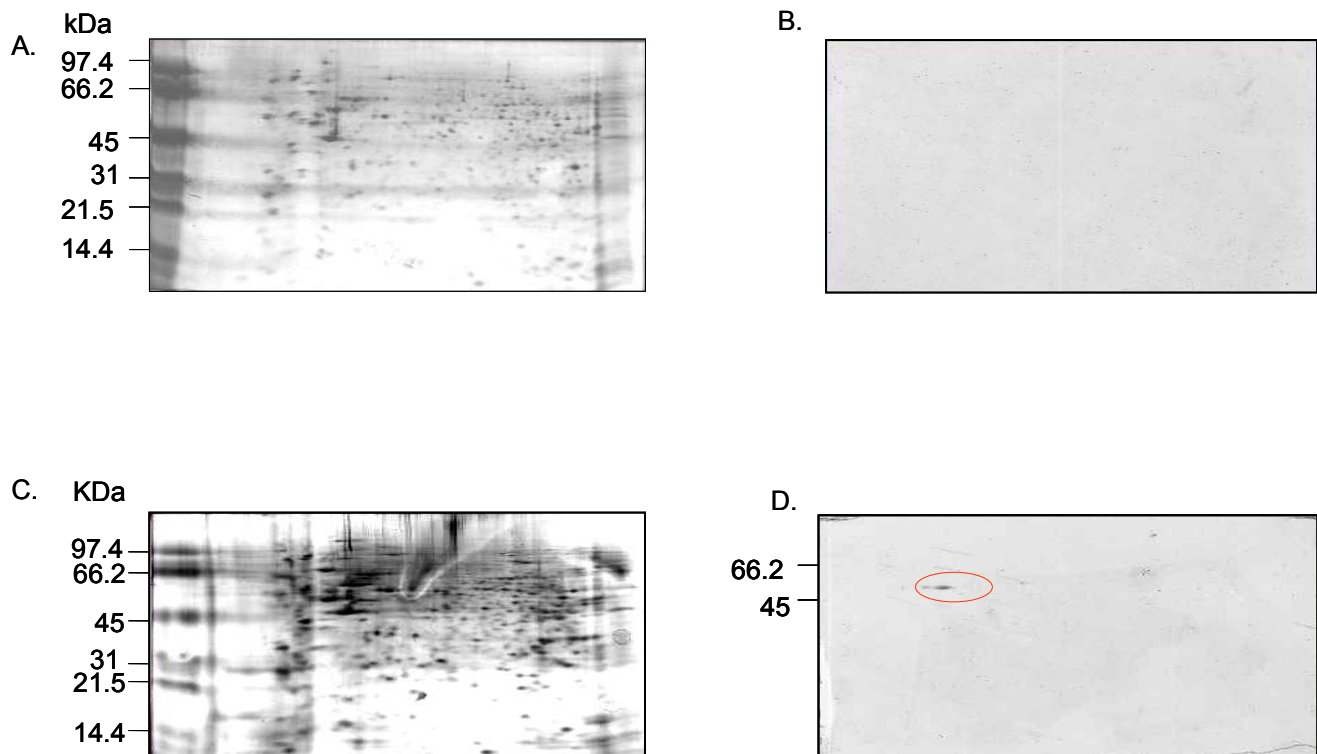


Figura 26. Gel bidimensional da fração de membranas pesadas das células HEK 293T transfectadas com o vetor pcDNA3.1 (controle negativo; A, B) ou transfectadas com pFLAGrasgef1b (C, D). O gradiente de pH dos géis é não linear e a orientação é 3 a 10 (esquerda para a direita). A amostra foi separada na primeira dimensão numa faixa de pI de 3 a 10 e na segunda dimensão por SDS-PAGE 12%. As massas moleculares relativas estão indicadas na esquerda (kDa). Coloração por coomassie coloidal (A,C); Western blotting com anti-FLAG (B,D); o “spot” correspondente à proteína FLAG-RasGEF1b no filme de raio X de Western Blotting está indicado dentro do círculo vermelho.

6. Discussão

A habilidade do *Trypanosoma cruzi* em ativar macrófagos, é, pelo menos em parte, atribuída às âncoras de glicosilfosfatidilinositol de glicoproteínas (GPI-mucinas) expressas na superfície do estágio tripomastigota do parasita. As técnicas de DDRT-PCR (differential display reverse transcriptase-polymerase chain reaction) e de Northern blot reverso (reverse Northern blot) foram utilizadas para estudar a modulação da expressão gênica em macrófagos murinos expostos a GPI-mucinas e em tecidos cardíacos de camundongos infectados com *T. cruzi*. Entre vários cDNAs identificados como derivados dos mais abundantes mRNAs presentes em macrófagos estimulados com GPI-mucinas, quando comparados a células não estimuladas, foi confirmado, em nosso laboratório, a expressão diferencial dos genes A1, interleucina 18, e *gpi*¹⁴ (o qual foi re-denominado *rasgef1b*). Alguns desses genes mostraram-se também com expressão acentuada em tecido cardíaco (A1, *rasgef1b* e DAP-12) de animais infectados. Também foi verificada a expressão do *rasGEF1b* em clones de linfócitos Th1 e Th2 murinos pré ativados por mitógenos, mas não em linfócitos sem pré ativação. Esses resultados mostram uma correlação entre a expressão do *rasGEF1b*, a ativação de macrófagos e de linfócitos. A expressão do *rasgef1b* mostrou-se também acentuada em monócitos humanos estimulados com GPI-mucinas ou LPS. A seqüência completa do transcrito do gene *rasgef1b* incluindo a região 5' flanqueadora foi definida. Constatou-se que o gene *rasgef1b* possui uma única cópia, presente em camundongos, e em humanos, e apresenta homologia a diferentes membros da família dos fatores de troca de nucleotídeos guanina (36).

O objetivo principal desta dissertação de mestrado foi expressar a proteína RasGEF1b, codificada pelo gene *rasgef1b*, em sistema procarioto e eucarioto. Para a expressão em bactérias, a região codificadora do *rasgef1b* foi amplificada e clonada no vetor de expressão pQE-30, em fusão com seis aminoácidos histidina. Isso permitiu a purificação da proteína recombinante em cromatografia de afinidade, utilizando uma coluna carregada com níquel. A proteína His-RasGEF1b foi obtida em grandes quantidades e sua seqüência foi confirmada por espectrometria de massa. Quanto ao sistema eucarioto, a seqüência codificadora do gene *rasgef1b* foi clonada no vetor pFLAG, em fusão com o peptídeo FLAG, de nove aminoácidos. A expressão da proteína FLAG-RasGEF1b ocorreu em células de mamífero e sua localização foi determinada através de centrifugação diferencial e eletroforese bidimensional.

O gene *rasgef1b* foi inicialmente denominado *gpiγ4*, porém, devido à sua homologia com outros membros da família RasGEF, sua sequência no banco de dados do NCBI, por anotação gênica, teve seu nome alterado para *rasgef1b*. Este gene mostra-se muito conservado entre várias espécies filogeneticamente distantes. O domínio RasGEF está relacionado com a função catalítica da troca de GDP por GTP pelos RasGEFs. Esta homologia corrobora com a nossa hipótese que o *rasgef1b* tem como principal função a troca de nucleotídeos guanina de proteínas Ras. Após purificarmos a proteína recombinante produzida em bactérias, tentamos produzir anticorpos policlonais para esta proteína, porém, não obtivemos nenhum resultado positivo. Isso pode ser explicado pela alta conservação da proteína RasGEF1b entre várias espécies (Figura suplementar 2), com produção e função estabelecidas e conservadas.

Proteínas com grande similaridade na sua sequência primária, e/ou com similaridade estrutural e funcional, pertencem a uma mesma família protéica. Uma forte relação evolucionária é geralmente evidente dentro de uma família protéica. Duas ou mais famílias com pouca similaridade na sequência primária fazem uso de um mesmo motivo estrutural principal e têm similaridades funcionais. Estas famílias são agrupadas em superfamílias. Motivos estruturais tornam-se especialmente importantes em definir família de proteínas e superfamílias. Sistemas de comparações e classificações aperfeiçoados para proteínas levam inevitavelmente à elucidação de novas relações funcionais. Dada uma função central de proteínas em sistemas vivos, estas comparações estruturais podem ajudar a iluminar vários aspectos da bioquímica, de evolução de proteínas individuais para a história evolucionária de vias metabólicas completas (86).

A sequência de *rasgef1b* contém dois domínios relacionados à RasGEFN e RasGEF. O domínio RasGEFN é um domínio presente em subtipos de GEFs para pequenas GTPases do tipo Ras. A cristalografia de Sos-1 revelou que o domínio RasGEFN está em uma α -hélice com papel estrutural (94). O RasGEFN também apresenta padrão de zíper de leucina que está normalmente envolvido na associação com outras proteínas. O domínio RasGEF é também encontrado em GEFs específicos para proteínas Ras. Domínios similares têm sido encontrados em leveduras como CD25 (95) e SCD25 (96), e estão envolvidos no controle da divisão celular. No domínio RasGEF do *rasgef1b* há três sítios de endereçamento nuclear. De fato, outros GEFs tais como RCC1 (97) e CDC42 (98) foram encontrados, respectivamente, localizado e direcionado para o núcleo sob certas condições fisiológicas. Então, os motivos de localização nuclear presentes no

domínio RasGEF do *rasgef1b* podem indicar que este GEF pode agir em associação com uma GTPase tipo Ras encontrada no núcleo da célula (36).

Após as modificações pós-traducionais, as proteínas podem ser direcionadas para diversos compartimentos celulares, como por exemplo, o núcleo. Uma sequência sinal direciona a proteína para o núcleo (a sequência de localização celular, NLS) e não é removida. Muitas vezes a NLS está localizada na porção amino-terminal, mas esta sequência pode estar localizada em qualquer parte da sequência de aminoácidos de uma proteína nuclear. A importação nuclear é mediada por um número de proteínas que circulam entre o citosol e o núcleo, incluindo as importinas α e β e a pequena GTPase Ran (86). Dois programas utilizados forneceram predições de localização citoplasmática (pTarget e NucPred) e outros dois, forneceram predição de localização nuclear (DBsubloc e Psort). É possível que a RasGEF1b esteja localizada em ambos os compartimentos, citosol e núcleo. Entretanto, a localização no citosol não é condizente com os resultados obtidos no laboratório, os quais indicam a localização nuclear e membranas pesadas da proteína FLAG-RasGEF1b. O GEF Sos é trazido do citoplasma para o receptor ativado numa maneira fosfotirosina-dependente através de proteínas adaptadoras tais como a Grb2. A Grb2 contém domínios SH3 que são ligados constitutivamente à região carboxi-terminal rica em prolina do Sos, e o complexo Grb2-Sos é recrutado para os receptores ativados por interações entre o domínio SH2 de Grb2 e os resíduos de fosfotirosina no receptor. Porque a Ras está localizada na membrana, a ativação do receptor resulta no aumento da efetiva concentração de Sos próximo à Ras, facilitando a troca de nucleotídeo guanina (99). Outra hipótese seria a localização do RasGEF1b na membrana próximo à Ras para que exerça a sua função e talvez no citoplasma quando inativo, assim como o GEF Sos.

O fracionamento subcelular permite a separação de organelas baseado nas suas propriedades físicas e consiste em dois passos principais: rompimento da organização celular e fracionamento do homogenato para separar as diferentes populações de organelas. O homogenato pode então ser fracionado por centrifugação diferencial em várias frações. Foi detectada a presença da proteína FLAG-RasGEF1b nas frações de membrana pesada e de núcleo, sendo nesta última, em menor quantidade. A fração de membranas pesadas compreende a membrana plasmática, o complexo de Golgi e o retículo endoplasmático (68). Entretanto, o fracionamento muitas vezes não é completo, e uma pequena contaminação entre as frações pode ocorrer (93). Experimentos de imunofluorescência estão em andamento para a confirmação por microscopia confocal da localização subcelular da proteína FLAG-RasGEF1b recombinante.

A modelagem de uma proteína pelo método de homologia baseia-se no conceito de evolução molecular. Em outras palavras, parte do princípio de que a semelhança entre as estruturas primárias desta proteína e de proteínas homólogas com estruturas tridimensionais conhecidas implique possivelmente em similaridade estrutural entre elas (100).

A sobreposição da modelagem tridimensional da RasGEF1b com o Sos mostra a existência de domínios bastante conservados, particularmente o domínio que interage com a Ras. É possível que a RasGEF1b interaja com a Ras da mesma forma que a Sos. Estudos de mutações sítio-dirigidas baseadas em modelagem tridimensional podem fornecer maiores informações sobre o tipo de interação RasGEF-Ras e quais os aminoácidos envolvidos. Neste sentido, deve ser salientado que, em experimentos de co-imunoprecipitação realizados pelo Dr. Aristóboło Mendes da Silva, a proteína FLAG-RasGEF1b co-precipitou com outra proteína reconhecida por um anticorpo anti-Ras.

No momento, a função da RasGEF1b é uma incógnita. Portanto, serão necessários estudos que estabeleçam sua atividade funcional. De posse de futuros resultados, poderemos então, determinar o envolvimento do *rasGEF1b* na ativação de receptores do tipo Toll, dando continuidade a uma via de sinalização celular estimulada pelo *T. cruzi* e outros agentes infecciosos ainda por ser elucidada.

7. Perspectivas

As descobertas deste e de outros trabalhos do nosso laboratório relacionados com o gene *rasgef1b* sugerem as próximas etapas de investigação. Acreditamos que a obtenção de maiores quantidades da proteína recombinante, na forma solúvel, permitirá a realização de outros estudos funcionais, como por exemplo, a análise da sua interação com a proteína Ras, através de experimentos de precipitação ou de co-imunoprecipitação. Nestas análises, poderemos estudar o processo de troca de GDP por GTP resultando na ativação da GTPase na presença His-RasGEF1b recombinante.

Estamos iniciando a análise de macrófagos murinos estimulados com agonistas do tipo Toll através da técnica de “shotgun” associada à espectrometria de massa. Com esta técnica, pretendemos caracterizar a proteína RasGEF1b nativa. Para a produção de anticorpos, pretendemos sintetizar um epítopo antigênico para a imunização de animais. Além disso, um refinamento na modelagem será necessário para guiar o trabalho experimental, sendo de grande interesse o estudo da interação Ras-RasGEF1b por mutagênese sítio-dirigida. A localização da RasGEF1b, como mencionado, será confirmada por experimentos de microscopia confocal.

8. Conclusão

A análise da seqüência de nucleotídeos da região codificadora do gene *gpiγ4* demonstrou a existência de domínios RasGEF e RasGEFN, e sua classificação como membro hipotético da família RasGEF denominado *rasgef1b*. A seqüência de aminoácidos da proteína codificada por esse gene, RasGEF1b, é altamente conservada em diferentes espécies filogeneticamente distintas. A homologia entre as seqüências analisadas variou de 57% a 97%.

A proteína recombinante His-RasGEF1b foi expressa em sistema procarioto e sua seqüência foi confirmada por espectrometria de massa. Não foi possível produzir anticorpos policlonais contra a RasGEF1b recombinante em coelhos e galinhas. Acreditamos que isso se deve à alta similaridade entre a proteína murina e a de galinha e, possivelmente, a de coelho.

A proteína FLAG-RasGEF1b foi expressa em sistema eucarioto e está localizada principalmente nas frações de membrana pesada e núcleo.

9. Dados suplementares

Figura suplementar 1

Database: PDB protein database, 25,879 sequences; 5,763,090 total letters

> [gi|4389274|pdb|1BKD|S](https://www.ncbi.nlm.nih.gov/blast/blast.cgi?gi|4389274|pdb|1BKD|S) S Chain S, Complex Of Human H-Ras With Human Sos-1
Length=439

Score = 80.1 bits (196), Expect = 1e-15, Method: Composition-based stats.
Identities = 88/382 (23%), Positives = 172/382 (45%), Gaps = 46/382 (12%)

```

Query  40  SGSLEALIQHLPNVDPDRTYIFTFLLSSRLFMHPYELMAKVCHLCVEHQRLSEGDD  99
      +G++ LI+ L  ++ Y D  ++ TFL + R F  P EL++   L +E  + E
Sbjct  30  AGTVIKLIERLTYHM--YADPNFVRTFLTTRYSFCKPQELLS----LIIERFEIPEPRFR  83

Query  100 KNQMRKIAPKILQLLTEWTETFPYDF-RDERMMRNLKDLAHRMASGEEQTYRKNVQMMQ  158
      K  ++ +  ++L +  W E  YDF RD  +++ +++  + +  + +K V+ + +
Sbjct  84  KEYIQPVQLRVLNVCRHWVEHHFYDFERDAYLLQRMEEF---IGTVRGKAMKKWVESITK  140

Query  159  CLIRKLAALSQYEEVLAKLSSTATDRLTVLKTTPQSIQR-DIMTVCSDPYTLAQQLTIE  217
      + RK                               SS T  +  ++P I+ D++T+  P  +A+QLT +E
Sbjct  141  IIQRKKITFQ-----SSPPTVEWHI--SRPGHIETFDLLTL--HPIEIARQLTLE  187

Query  218  LERLNYIGPEEFVQAFVQKDPDNDKSCYSEKTRNLEAYVEWFNRLSYLVATEICMPV  277
      +  + P E V +  K+ D + + + K R+   WF +  +V TE
Sbjct  188  SDLYRAVQPSELVGSVWTK--DKEINSPNLLKMIRHTTNLTLWFEKC--IVETE-----  238

Query  278  KKKHRARMIEYFIDVARECFNIGNFNLSMAIISGMNMSPVSRLLKKTWAKVKTAKFDILEH  337
      + R  ++  I++ +  + NFN ++ ++S MN SPV RL  T+ ++ + +  ILE
Sbjct  239  NLEERVAVVSRIIEILQVFQELNNFNGVLEVVSAMNSSPVYRLDHTFEQIPSRQKKILEE  298

Query  338  QMDPSSNFYNYRTALRGAAQRSLTAHSSREKIVIPFFSLLIKDIYFLNEG---CVNRLPN  394
      + S + Y                ++ L  S  +PFF + + +I  EG  + R
Sbjct  299  AHELSEDHY-----KKYLAKLRSINPPCVPFPGIYLTNLIKTEEGNPEVLKRHGK  348


Query  395  GHVNFEEKFWELAKQVSEFMTWK  416
      +NF K  ++A+  E  ++
Sbjct  349  ELINFSKRRKVAEITGEIQYQ  370

```

Figura Suplementar 1. Blastp da sequência da proteína RasGEF1b com a sequência 1bkd

Figura suplementar 2

Database: All non-redundant GenBank CDS translations+PDB+SwissProt+PIR+PRF excluding environmental samples,794,285 sequences; 1,308,253,377 total letters

> [gi|50746537|ref|XP_420541.1](#)  PREDICTED: similar to RasGEF domain family, member 1B; GPI-gamma 4; likely ortholog of mouse GPI-gamma 4 [Gallus gallus]
Length=573

Score = 890 bits (2301), Expect = 0.0, Method: Composition-based stats.
Identities = 429/473 (90%), Positives = 457/473 (96%), Gaps = 1/473 (0%)

Query	1	MPQTPPFSA	MPQTPPFSA	60
Sbjct	102	MPQTPPFSA	MPQTPPFSA	161
Query	61	TYIFTFLSS	TYIFTFLSS	120
Sbjct	162	TYIFTFLSS	TYIFTFLSS	221
Query	121	FPYDFRDER	FPYDFRDER	180
Sbjct	222	FPYDFRDER	FPYDFRDER	280
Query	181	ATDRLTVLK	ATDRLTVLK	240
Sbjct	281	ATDRLTVLK	ATDRLTVLK	340
Query	241	NDKSCYSER	NDKSCYSER	300
Sbjct	341	NDKSCYSER	NDKSCYSER	400
Query	301	NFNSLMAII	NFNSLMAII	360
Sbjct	401	NFNSLMAII	NFNSLMAII	460
Query	361	TAHSSREKI	TAHSSREKI	420
Sbjct	461	TAHSSREKI	TAHSSREKI	520
Query	421	PFERDRKVL	PFERDRKVL	473
Sbjct	521	PFERDRKVL	PFERDRKVL	573

Figura Suplementar 2: Blastp entre as seqüências das proteínas RasGEF1b murina com uma proteína hipotética de galinha.

Tabela suplementar 1 - Representação dos aminoácidos por códigos de letras

Aminoácidos		
Nome completo	Código de letra única	Código de três letras
Alanina	A	Ala
Arginina	R	Arg
Asparagina	N	Asn
Ácido Aspártico	D	Asp
Cisteína	C	Cys
Glutamina	Q	Gln
Ácido Glutâmico	E	Glu
Glicina	G	Gly
Histidina	H	His
Isoleucina	I	Ile
Leucina	L	Leu
Lisina	K	Lys
Metionina	M	Met
Fenilalanina	F	Phe
Prolina	P	Pro
Serina	S	Ser
Treonina	T	Thr
Triptofano	W	Trp
Tirosina	Y	Tyr
Valina	V	Val

10. Endereços Eletrônicos

Protparam: <http://br.expasy.org/cgi-bin/protparam>

Clustal W: <http://www.ebi.ac.uk/clustalw/>

Protein-protein BLAST, blastp: <http://www.ncbi.nlm.nih.gov/blast/>

SignalP: www.cbs.dtu.dk/services/SignalP/

NetAcet: www.cbs.dtu.dk/services/NetAcet/

NetNglyc: www.cbs.dtu.dk/services/NetNGlyc/

NetOGlyc: www.cbs.dtu.dk/services/NetOGlyc/

NetPhos: www.cbs.dtu.dk/services/NetPhos/

YinOYang 1.2. Prediction Server: www.cbs.dtu.dk/services/YinOYang/

Prosite: <http://www.expasy.ch/prosite>

PTarget: bioinformatics.albany.edu/~ptarget/

NucPred: www.sbc.su.se/~maccallr/nucpred/

DbSubLoc: www.bioinfo.tsinghua.edu.cn/dbsubloc.html

Psort: <http://www.psort.org/>

HNN Secondary Structure Prediction Method: http://npsa-pbil.ibcp.fr/cgi-bin/secpred_hnn.pl

ESyPred3DwebServer1.0: <http://www.fundp.ac.be/sciences/biologie/urbm/bioinfo/esypred/>.

Pymol: <http://pymol.sourceforge.net/>

PROCHECK v.3.5.4: <http://www.biochem.ucl.ac.uk/~roman/procheck/procheck.html>

PHYLIP: <http://evolution.genetics.washington.edu/phylip.html>

11. Referências

1. Fonte: World Health Report. (2002)
<http://www.who.int/tdr/diseases/chagas/diseaseinfo.htm>, consultado em abril de 2005.
2. Sherlock, I.A., Vetores, in: Brener, Z., Andrade, Z. and Barral-Neto, M. (2nd Eds), *Trypanosoma cruzi* and Chagas' disease. Editora Guanabara Koogan SA, Rio de Janeiro, Brazil (2000) 21-40
3. Dias, J.C.P. (1992). Epidemiology of Chagas disease. In *Chagas Disease (American Trypanosomiasis): its impact on transfusion and clinical medicine* (Wendel, S., Brener, Z., Camargo, M.E. and Rassi, A) eds. ISBT. São Paulo.49-80.
4. Koberle, F. (1968) Chagas' disease and Chagas' syndrome: the pathology of American Trypanosomiasis. *Adv. Parasitol.*6, 63-116.
5. Teixeira, A.R., Teixeira, G., Macedo, V., Prata, A. (1978) Acquired cell-mediated immunodepression in acute Chagas' disease. *J. Clin. Invest.* 62,1132-1141.
6. Zhang, L., Tarleton, R.L. (1999). Parasite persistence correlates with disease severity and localization in chronic Chagas' disease. *J. Infect. Dis.* 180: 480-486.
7. Andrade, Z.A. (1999). Immunopathology of Chagas Disease. *Mem Inst Oswaldo Cruz, Rio de Janeiro, Vol. 94, I:* 71-80.
8. Andrade, Z.A. & Andrade, S.G. (1979). Patologia. Em Andrade & Brener, *Trypanosoma cruzi* e doença de Chagas. Ed. Guanabara Koogan. Rio de Janeiro, Brasil. p 199.
9. Biron, C.A., Gazzinelli, R.T (1995). Effects of IL-12 on immune response to microbial infections: a key mediator in regulating disease outcome. *Curr. Opin. Immunol*7:485-497.
10. Gazzinelli, R.T., Amychay, D., Sharton-Kersten, T., Grunwald, E., Farber, J.M., Sher, A. (1996) Role of macrophages derived cytokines in the induction and regulation of cell mediated immunity to *Toxoplasma gondii*. *Curr. Top. Microb. Immunol* 219: 127-140.
11. Sharton-Kersten, T., Sher, A. (1997). Role of natural killer cells in innate resistance to protozoan infection. *Curr. Opin. Immunol* 9:44-51.
12. Trinchieri, G. (1995) Interleukin 12: a proinflammatory cytokine with regulatory function that bridge innate resistance and adaptative immunity. *Ann. Rev. Immunol.* 13:251-276.
13. Fearon, D.T. and Locksley, R.M. (1996). The instructive role of innate immunity in the acquired immune response. *Science* 272:50-54.
14. Ulevitch, R.J. and Tobias, P.S. (1995). Receptor-dependent mechanisms of cell stimulation by bacterial endotoxin. *Ann Rev Immunol* 13: 437-457.

15. Roach, T.I., Barton, C.H., Chatterjee, D. and Blackwell, J.M (1993). Macrophage activation: lipoarabinomannan from avirulent and virulent strains of *Mycobacterium tuberculosis* differentially induces the early genes c-fos, KC, JE and tumor necrosis factor- α . *J Immunol* 150:1886–1896.
16. Gazzinelli, R.T., Hieny, S., Wynn, T.A., Wolf, S. and Sher, A. (1993) Interleukin-12 is required for the T-lymphocyte-independent induction of interferon- γ by an intracellular parasite and induces resistance in T-cell deficient hosts. *Proc Natl Acad Sci USA* 90:6115–6119.
17. Klinman, A.K., Yi, S.L. Beaucage, J. Conover and Krieg, A.M. (1996) CpG motifs present in bacteria D.M DNA rapidly induce lymphocytes to secrete interleukin 6, interleukin 12 and interferon γ . *Proc Natl Acad Sci USA* 93:2879–2883.
18. Brightbill, H.D., Brightbill, D.H., Libraty, S.R., Krutzik, R.B., Yang, J.T., Belisle, J.R., Bleharski, M., Maitland, M.V., Norgard, S.E., Plevy, S.T., Smale et al. (1999) Host defense mechanisms triggered by microbial lipoproteins through Toll-like receptors. *Science* 285:732–736.
19. Gazzinelli, R.T. and Ropert, C. (2000) Signaling of immune system cells by glycosylphosphatidylinositol (GPI) anchor and related structures derived from parasitic protozoa *Curr Opin Microb* 3: 395-403.
20. Schofield, L. and Hackett, F. (1993) Signal transduction in host cells by glycosylphosphatidylinositol toxin of malaria parasites. *J Exp Med* 177:145–153.
21. Tachado, S.D., and Schofield, L. (1994) Glycosylphosphatidylinositol toxin of *Trypanosoma Brucei* regulates IL-1 alpha and TNF- α expression in macrophages by protein tyrosine kinase mediated signal transduction. *Biochem Biophys Res Com* 205:984–991.
22. Camargo, M.M., Almeida, C.I., Pereira, S.E.M. , Ferguson, J.A.M., Travassos, R.L., Gazzinelli, R.T. (1997). Glycosylphosphatidylinositol-anchored mucin-like glycoproteins isolated from *Trypanosoma cruzi* trypomastigotes initiate the synthesis of proinflammatory cytokines by macrophages. *J. Immunol.* 158:5890-5901.
23. Coelho, S.C., Klein, A., Talvani, A., Coutinho, S. F., Takeuchi, O., Akira, S., Canizzaro, H., Gazzinelli, R.T., Teixeira, M.M. (2002) Glycosylphosphatidylinositol-anchored mucin-like glycoproteins isolated from *Trypanosoma cruzi* trypomastigotes induce in vivo leukocyte recruitment dependent on MCP-1 production by IFN-gamma-primed-macrophages *J. Leukoc. Biol.* 71:837-844.
24. Ropert, C., Almeida, I.C., Closel, M., Travassos, L. R., Ferguson, M.A.J., Cohen, P., Gazzinelli, R.T. (2001) Requirement of mitogen-activated protein kinases and I kappa B phosphorylation for induction of proinflammatory cytokines synthesis by macrophages indicates functional similarity of receptors triggered by glycosylphosphatidylinositol anchors from parasitic protozoa and bacterial lipopolysaccharide *J. Immunol.* 166:3423-3431.

25. Campos, M.A.S., Almeida, I.C., Takeuchi, O., Akira, S., Procópio, D.O., Valente, E., Travassos, L.R., Smith, J.A., Golenbock, D.T., Gazzinelli, R.T. (2001) Activation of Toll-like receptor-2 by glycosylphosphatidylinositol anchors from a protozoan parasite. *J. Immunol.* 167:416-423.
26. Talvani, A., Ribeiro, C.S., Aliberti, J.C., Michailowsky, V., Santos, P.V., Murta, S.M. Romanha A.J., Almeida, I. C., Farber, J., Lannes-Vieira, J, Silva, J.S., Gazzinelli, R.T. (2000) Kinetics of cytokine gene expression in experimental chagasic cardiomyopathy: tissue parasitism and endogenous IFN-gamma as important determinants of chemokine mRNA expression during infection with *Trypanosoma cruzi*. *Microbes Infect.* 2:851-866.
27. dos Santos, P.V., Roffe, E., Santiago, H.C., Torres, R.A., Marino, A.P.M.P., Paiva, C.N., Silva, A.A., Gazzinelli, R.T., Lannes-Vieira, J. (2001) Prevalence of CD8(+)alpha beta T cells in *Trypanosoma cruzi*-elicited myocarditis is associated with acquisition of CD62L(Low)LFA-1(High)VLA-4(High) activation phenotype and expression of IFN-gamma-inducible adhesion and chemoattractant molecules. *Microbes Infect.* 3:971-984.
28. Aliberti, J.C.S., Souto, J. T., Marino, A.P.M.P., Lannes-Vieira, J., Teixeira, M.M., Farber, J., Gazzinelli, R.T., Silva, J.S. (2001) Modulation of chemokine production and inflammatory responses in interferon-gamma-and tumor necrosis factor-R1-deficient mice during *Trypanosoma cruzi* infection. *Am. J. Pathol.* 158:1433-1440.
29. Michailowsky, V., Silva, N.M., Rocha, C.D., Vieira, L.Q., Lannes-Vieira, J., Gazzinelli, R. T. (2001) Pivotal role of interleukin-12 and interferon-gamma axis in controlling tissue parasitism and inflammation in the heart and central nervous system during *Trypanosoma cruzi* infection. *Am. J. Pathol.* 159:1723-1733.
30. Liang, P., Pardee, A.B. (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* 257,967-971.
31. Liang, P., Averboukh, L., Pardee, A.B. (1993) Distribution and cloning of eukaryotic mRNAs by means of differential display: refinements and optimization. *Nucleic Acids Res.* 21:3269-3275.
32. Zhang, H., Zhang, R., Liang, P. (1996) Differential screening of gene expression difference enriched by differential display. *Nucleic Acids Res.* 24:2454-2455.
33. Olivier, J.P., Raabe, T., Henkemeyer, M., Dickson, B., Mbamalu, G., Margolis, B., Schlessinger, J., Hafen, E., Pawson, T. (1993) A *Drosophila* SH2-SH3 adaptor protein implicated in coupling the sevenless tyrosine kinase to an activator of Ras guanine nucleotide exchange, Sos. *Cell* 73:179-19.
34. Cen, H., Papageorge, A.G., Zippel, R., Lowy, D.R., Zhang, K. (1992) Isolation of multiple mouse cDNAs with coding homology to *Saccharomyces cerevisiae* CDC25: identification of a region related to Bcr, Vav, Dbl and CDC24. *EMBO J.* 11:4007-4015.
35. Schweighoffer, F., Faure, M., Fath, I., Chevallier-Multon, M. C., Apiou, F., Dutrillaux, B., Sturani, E., Jacquet, M., Tocque, B. (1993) Identification of a human guanine nucleotide-releasing factor (H-GRF55) specific for Ras proteins. *Oncogene* 8:1477-1485.

36. Ferreira, L.R.P., Abrantes E.F., Rodrigues, C.V., Caetano, B., Cerqueira C.G., Salim A.C., Reis L.F.L., Gazzinelli, R.T. (2002). Identification and characterization of a novel mouse gene encoding a Ras-associated guanine nucleotide exchange factor: expression in macrophages and myocarditis elicited by *Trypanosoma cruzi* parasites. *J. Leukoc. Biol.* 72: 1215-1227.
37. Lowy, D.R. and Willumsen, B.M (1993). Function and regulation of Ras. *Ann Rev. Biochem.* 62:851-891.
38. Coleman L.M., Marshall J.C., and Olson F.M. (2004). RAS and RHO GTPases in G1-phase cell-cycle regulation. *Nat. Rev. Mol. Cel.Biol.* 5:355-366
39. Boguski MS, McCormick F. (1993) Proteins regulating Ras and its relatives. *Nature*, 366(6456):643-654.
40. Campbell, S.L., Khosravi-Far, R., Rossman, K.L., Clark, G.J. & Der, C.J. (1998) Increasing complexity of Ras signaling. *Oncogene* 17:1395-413.
41. Wilkins, A., Insall, R.H. (2001) Small GTPases in *Dictyostelium*: lessons from a social amoeba. *Trends Genet* , 17:41-48.
42. Fukui, K., Sasaki, T., Imazumi, K., Matsuura, Y., Nakanishi, H., and Takai Y.(1997) Isolation and characterization of a GTPase activating protein specific for the Rab3 subfamily of small G proteins. *J Biol Chem* 272: 4655–4658.
43. Trahei, M. and McCormick, F. (1987). A cytoplasmic protein stimulates normal N-ras p21 GTPase, but does not affect oncogenic mutants. *Science* 238: 542–545.
44. Settleman, J., Albright, C.F., Foster, L.C, AND Weinberg, R.A. (1992) Association between GTPase activators for Rho and Ras families. *Nature* 359: 153–154.
45. Buday, L. and Downward, J. (1993) Epidermal growth factor regulates p21ras through the formation of a complex of receptor, Grb2 adapter protein, and Sos nucleotide exchange factor. *Cell* 73: 611–620.
46. Wada, M., Nakanishi, H., Satoh, A., Hirano, H., Obaish, H., Matsuura, Y., and Takay Y. (1997) Isolation and characterization of a GDP/GTP exchange protein specific for the Rab3 subfamily small G proteins. *J Biol Chem* 272: 3875–3878.
47. Yaku, H., Sasaki, T., and Takai Y. (1994) The Dbl oncogene product as a GDP/GTP exchange protein for the Rho family: its properties compared with those of Smg GDS. *Biochem Biophys Res Commun* 198:811–817.
48. Horiuchi, H., Kawata, M., Katayama, M., Yoshida, Y., Musha, T., Ando, S., and Takai Y. (1991) A novel prenyltransferase for a small GTP-binding protein having a C-terminal Cys-Ala-Cys structure. *J Biol Chem* 266: 16981–16984.

49. Bowtell, D., Fu, P., Simon, M. & Senior, P. (1992) Identification of murine homologues of the *Drosophila* son of sevenless gene: potential activators of ras. *Proc Natl Acad Sci USA* 89:6511-5.
50. Chardin, P., Camonis, J.H., Gale, N. W., van Aelst, L., Schlessinger, J., Wigler, M. H. & Bar-Sagi, D. (1993) Human Sos1: a guanine nucleotide exchange factor for Ras that binds to GRB2. *Science* 260:1338-43.
51. Ebinu, J.O., Bottorff, D.A., Chan, E.Y., Stang, S.L., Dunn, R. J. & Stone, J.C. (1998) A Rap guanine nucleotide exchange factor enriched highly in the basal ganglia. *Science* 280:1082-6.
52. Tanaka, S., Morishita, T., Hashimoto, Y., Hattori, S., Nakamura, S., Shibuya, M., Matuoka, K., Takenawa, T., Kurata, T., Nagashima, K. & et al. (1994) C3G, a guanine nucleotide-releasing protein expressed ubiquitously, binds to the Src homology 3 domains of CRK and GRB2/ASH proteins. *Proc Natl Acad Sci U S A* 91, 3443-3447.
53. Gulbins, E., Coggeshall, K.M., Baier, G., Katzav, S., Burn, P. & Altman, A. (1994) Activation of Ras in vitro and in intact fibroblasts by the Vav guanine nucleotide exchange protein. *Mol Cell Biol.* 14(2): 906-13.
54. Wei, W., Mosteller, R. D., Sanyal, P., Gonzales, E., McKinney, D., Dasgupta, C., Li, P., Liu, B. X. & Broek, D. (1992) Identification of a mammalian gene structurally and functionally related to the CDC25 gene of *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA* 89:7100-4.
55. Cen, H., Papageorge, A.G., Zippel, R., Lowy, D R. & Zhang, K. (1992) Isolation of multiple mouse cDNAs with coding homology to *Saccharomyces cerevisiae* CDC25: identification of a region related to Bcr, Vav, Dbl and CDC24. *Embo J* 11: 4007-15.
56. Martegani, E., Vanoni, M., Zippel, R., Coccetti, P., Brambilla, R., Ferrari, C., Sturani, E. & Alberghina, L. (1992) Cloning by functional complementation of a mouse cDNA encoding a homologue of CDC25, a *Saccharomyces cerevisiae* RAS activator. *Embo J* 11:2151-7.
57. Quilliam, L.A., Rebhun, J.F. & Castro, A.F. (2002) A growing family of guanine nucleotide exchange factors is responsible for activation of Ras-family GTPases. *Prog Nucleic Acid Res Mol Biol* 71:391-444.
58. Felsenstein, J. 2005. PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.
59. Lemmon, M.A., Ferguson, K. M. & Schlessinger, J. (1996) PH domains: diverse sequences with a common fold recruit signaling molecules to the cell surface. *Cell* 85:621-4.
60. Quilliam, L.A., Khosravi-Far, R., Huff, S.Y. & Der, C.J. (1995) Guanine nucleotide exchange factors: activators of the Ras superfamily of proteins. *Bioessays* 17:395-404.

61. Daniel J., Spiegelman G.B. and Weeks G. (1993) Characterization of a third ras gene, rasB, that is expressed throughout the growth and development of *Dictyostelium discoideum*. *Oncogene* 8:1041–1047
62. Daniel J., Bush J., Cardelli J, Spiegelman G.B. and Weeks G. (1994) Isolation of two novel ras genes in *Dictyostelium discoideum*; evidence for a complex, developmentally regulated ras gene subfamily. *Oncogene* 9: 501–508.
63. Chubb J.R., Thomas G.M. and Insall R.H. (2000) The *Dictyostelium* RasS protein is required for macropinocytosis, phagocytosis and the control of cell movement. *J Cell Sci* 113: 709–719.
64. Kyriakis, J.M., Avruch, J. (1996) Sounding the alarm: protein kinase cascades activated by stress and inflammation. *Sounding the alarm: protein kinase cascades activated by stress and inflammation.. J. Biol. Chem.* 271:24313-24316.
65. Cohen, P. (1997) The search for physiological substrates of mitogen- and stress-activated protein kinases in mammalian cells. *Trends Cell Biol.* 7:353-361.
66. Campos, M.A.S., Almeida, I.C., Takeuchi, O., Akira, S., Procópio, D.O., Valente, E., Travassos, L.R., Smith, J.A., Golenbock, D.T., Gazzinelli, R.T. (2001) Activation of Toll-like receptor-2 by glycosylphosphatidylinositol anchors from a protozoan parasite. *J. Immunol.* 167:416-423.
67. Medzhitov, R. (2001) Toll-like receptors and innate immunity. *Nat. Rev. Immunol.* 1:135-145.
68. Berger, T. and Kretzler, M. (2002) Interaction of DAP3 and FADD only after cellular disruption. *Nat. Immunol.* 3: 3–5.
69. Nucleic Acids Res. Bairoch, A., Bucher, P., Hofmann, K. (1997) The PROSITE database, its status in 1997. Department of Medical Biochemistry, University of Geneva, 1 rue Michel Servet 1211 Geneva 4, Switzerland. bairoch@cmu.unige.ch. Jan 1;25(1):217-221.
70. Emanuelsson O., Nielsen H., Brunak S. and von Heijne G.J.(2002) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *Mol. Biol.*, 300: 1005-1016.
71. Sujun Hua & Zhirong Sun: Support Vector Machine Approach for Protein Subcellular Localization Prediction, *Bioinformatics*, in press.
72. Nakai, K. and Horton, P. (1999). PSORT: a program for detecting the sorting signals of proteins and predicting their subcellular localization, *Trends Biochem. Sci*, 24(1) 34-35.
73. Heddad A., Brameier M. and MacCallum R.M.(2004) Evolving Regular Expression-Based Sequence Classifiers for Protein Nuclear Localisation. *EvoBIO 2004 Workshop, Lecture Notes in Computer Science Vol. 3005.* pp 31-40.

74. 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. *The Proteomics Protocols Handbook*, Humana Press pp. 571-607.
75. Bendtsen, J.D., Nielsen, H., von Heijne, G. and Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.*, 340:783-795.
76. Gupta, R., Jung, E. and Brunak, S. (2004) Prediction of N-glycosylation sites in human proteins.. In preparation, 2004.
77. Julenius, K., Mølgaard, A., Gupta, R. and Brunak, S. (2005) Prediction, conservation analysis and structural characterization of mammalian mucin-type O-glycosylation sites. *Glycobiology*, 15:153-164.
78. Gupta, R., Hansen, J. and Brunak, S. Identifying intracellular O-(beta)-GlcNAc 'yin-yang' switches in the available human proteome. Manuscript in preparation.
79. Kiemer, L., Dyrlov, J. Bendtsen and Blom, N. (2005) Prediction of N-terminal acetylation sites. *Bioinformatics*, 21(7):1269-70
80. Blom, N., Gammeltoft, S., and Brunak, S. (1999) Sequence- and structure-based prediction of eukaryotic protein phosphorylation sites. *Journal of Molecular Biology*: 294(5): 1351-1362.
81. Higgins, D., Thompson, J., Gibson, T. Thompson, J.D., Higgins D.G., Gibson T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22:4673-4680.
82. Altschul, Stephen, F., Madden, T.L., Schäffer, A.A., Zhang, J.Z.Z., Miller, W., and Lipman, D.J. (1997) "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs". *Nucleic Acids Res.* 25:3389-3402
83. Combinaison de classifieurs statistiques, Application a la prediction de structure secondaire des proteines PhD Thesis Guermeur, Y.
84. Lambert, C., Leonard, N., De Bolle X, Depiereux E. (2002) ESyPred3D: Prediction of proteins 3D structures. *Bioinformatics*. Sep;18 (9):1250-1256.
85. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.*, 26, 283-291.
86. Nelson, D.L; Cox, M.M. (2000) *Lehninger: principles of biochemistry*. 3. ed. New York: Worth Publishers, ISBN: 1-57259-931-6.
87. Smith, D.B., and Johnson, K.S. (1988) Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene (Amst.)* 67:31-40.

88. di Guan, C., Li, P., Riggs, P.D., and Inouye, H. (1988) Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein. *Gene (Amst.)* 67: 21–30.
89. LaVallie, E.R., DiBlasio, E.A., Kovacic, S., Grant, K.L., Schendel, P.F., and McCoy, J.M. (1993) *Bio/Technology* 11:187–193.
90. Zhijian, L., Elizabeth, A., DiBlasio-Smith, GrantK.L., Nicholas W. Warne, LaVallie E.R., Collins-Racie, L.A., Follettie, M.T., Williamson,J.M. and McCoy, J.M. (1996). Histidine Patch Thioredoxins. *J. Biol. Chem.* 271: 5059-5065.
91. O' Farrel, PH. (1975) High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* 250:4007-4021.
92. Harrison, P.M., Kumar, A. Lang, N., Snyder, M., Gerstein, M. (2002). A question of size: the eukaryotic proteome and the problems in defining it. *Nucleic Acids Res.* 30:1083-1090.
93. Huber, L.A. (2003) Opinion: is proteomics heading in the wrong direction? *Nat Rev Mol. Cell Biol.* 4:74-80.
94. Boriack-Sjodin, P.^a, Margarit, S.M., Bar-Saji,D., Kuryan, J.(1998). The structural basis of the activation of Ras By Sos. *Nature*, 394:337-343.
95. Boguski, M.S., McCornick, F.(1993). Proteins regulating Ras and its relatives. *Nature*, 366: 643-654.
96. Miosga, T., Zimmermann, F.K. (1996). Sequence analysis of the CEN12 region of *Saccharomyces cerevisiae* on a 43.7 Kb fragment of chromosome XII including an open reading frame homologous to the human cystic fibrosis transmembrane conductance regulator protein CFTR. *Yeast*, 12: 693-708.
97. Talcott, B., Moore, M.S. (2000). The nuclear import of RCC1 requires a specific nuclear localization sequence receptor, karyopherin alpha3/Qip. *J. Biol. Chem.*275:10099-10104.
98. Toenjes, K.A., Sawyer, M.M, Johnson, D.I.(1999) The guanine nucleotide exchange factor CDC24p is targeted to the nucleus and polarized growth sites. *Curr. Biol.*, 9:1183-1186.
99. Boriack-Sjodin, P.A., Margarit, S.M., Bar-Saji, D., Kuriyan, J. (1998) The structural basis of the activation of Ras by Sos. *Nature* 394: 337-343.
100. Teichmann, S.A., Murzin A.G., Chothia C. (2001) Determination of protein function, evolution and interactions by structural genomics. *Curr Opin in Structural Biology*, v.11, 354-363.