



Pontifícia Universidade Católica do Rio Grande do Sul
Faculdade de Biociências
Programa de Pós-Graduação em Biologia Celular e Molecular

**Expressão e Purificação do Domínio Ets do Fator de Transcrição Spi-C
Humano Recombinante: Ensaio de Ligação a Promotores de Linfócitos B**

Autor
Ekaterini Gianniotis

Orientador:
Prof. Dr. Maurício Reis Bogo

Co-orientadores:
Prof. Dr. Diógenes Santiago Santos
Prof. Dr. Luiz Augusto Basso

Porto Alegre, RS
Outubro, 2008



Pontifícia Universidade Católica do Rio Grande do Sul
Faculdade de Biociências
Programa de Pós-Graduação em Biologia Celular e Molecular

**Expressão e Purificação do Domínio Ets do Fator de Transcrição Spi-C
Humano Recombinante: Ensaio de Ligação a Promotores de Linfócitos B**

Dissertação apresentada ao
Programa de Pós-Graduação
em Biologia Celular e Molecular
como requisito para obtenção
do grau de Mestre.

Porto Alegre, RS
Outubro, 2008

AGRADECIMENTOS

Agradeço ao Prof. Diógenes Santiago Santos pela oportunidade de ingressar no CPBMF o que proporcionou muito aprendizado e crescimento acadêmico.

Este trabalho foi financiado pelo CNPq e pela Hewlett Packard e a bolsa acadêmica foi financiada pela Capes.

Quero agradecer principalmente à Claudia P. Nunes, pelos ensinamentos, apoio e motivação constante. Especialmente pela ajuda na elaboração deste trabalho, tanto durante a parte experimental quanto na parte escrita. Pelo ombro amigo e sábio.

Agradeço a todos os colegas de laboratório pelo companheirismo. Especialmente àqueles grandes amigos pelo apoio, carinho, motivação e ajuda durante a realização dos experimentos e pelos bate-papos no dia-a-dia. Também agradeço à M^a Gleci A. Ferreira e Renilda Trapp de Mello pela amizade e “quebra-galhos”.

Aos meus pais pelo amor, carinho, compreensão, paciência e pela força para seguir em frente nos momentos difíceis.

Ao Augusto, pelo apoio e compreensão nos momentos de nervosismo e stress e alegria nos momentos de sucesso.

ÍNDICE

RESUMO	5
ABSTRACT	6
INTRODUÇÃO	7
1. A família de fatores de transcrição Ets	7
1.2 Os domínios existentes nas proteínas Ets	8
1.3 Interação entre proteínas Ets e o DNA	10
1.4 Spi-C	12
1.5 Células B e possíveis seqüências alvo para Spi-C	18
OBJETIVOS	22
ARTIGO CIENTÍFICO	23
CONSIDERAÇÕES FINAIS	38
REFERÊNCIAS	44
ANEXO I	48
<u>Testes de expressão em cepas de <i>Escherichia coli</i></u>	48
<u>Protocolo de produção de células que expressam o domínio Ets de Spi-C</u> <u>humana</u>	51
ANEXO II	52
<u>Testes de purificação do domínio Ets de Spi-C humana recombinante</u>	52
ANEXO III	54
<u>Testes de ligação ao DNA</u>	54
Ressonância plasmônica de superfície	54
Espectrofluorimetria	57

RESUMO

Os membros da família de fatores de transcrição Ets contêm um domínio de ligação ao DNA, chamado domínio Ets, que reconhece uma seqüência rica em purinas, cujo consenso central é formado pela seqüência 5'-GGAA/T-3'. O domínio Ets pode ser produzido como um fragmento de proteína que se enovela independentemente, numa estrutura estável, sendo suficiente para que ocorra ligação ao DNA. A proteína Spi-C pertence à subfamília Spi e se caracteriza por ser expressa durante diferentes estágios do desenvolvimento de células B e em macrófagos. Existem alguns trabalhos publicados que descrevem o reconhecimento por Spi-C de algumas seqüências de DNA, previamente estudadas para a proteína PU.1, membro da mesma subfamília. A expressão de Spi-C sobrepõem-se com a de PU.1 ao longo do desenvolvimento das células B. Existem promotores de genes envolvidos no desenvolvimento de células B que contêm sítios para a ligação de proteínas Ets. Estudos quanto à ligação da Spi-C nos mesmos podem contribuir para acrescentar dados sobre a função desse fator de transcrição e sobre a regulação da transcrição desses genes. Neste trabalho, apresentamos a construção do gene, a expressão heteróloga do domínio Ets da proteína Spi-C humana em *Escherichia coli* e a sua purificação por FPLC com rendimento de aproximadamente 0,7mg de proteína por grama de célula. Com o objetivo de investigar novas seqüências promotoras alvo para Spi-C, foram realizados estudos de ligação a seqüências de DNA localizadas nos promotores dos genes que codificam para a subunidade do receptor de IL-7 e para a proteína transmembrana integrante do complexo receptor de células B, Ig α . Estas proteínas participam do desenvolvimento de células B e são essenciais para a seleção das mesmas na medula óssea. Nossos resultados demonstram que o domínio Ets de Spi-C reconhece e se liga a essas seqüências de DNA (migração em gel) formando complexos de diferentes tamanhos. O domínio Ets foi capaz de se ligar a seqüências de DNA *in vitro* numa proporção maior do que 1:1 sendo que, quanto maior a concentração do domínio maior foi o retardo do DNA no gel. Surpreendentemente, o domínio Ets não foi capaz de discriminar as seqüências selvagens, daquelas com mutações no consenso central de ligação, mantendo o mesmo perfil de retardo das bandas, sugerindo que o restante da proteína é importante para esse reconhecimento.

Palavras-chave: Spi-C; desenvolvimento de células B; domínio Ets; ligação ao DNA.

ABSTRACT

The members of the Ets family of transcription factors contain a DNA binding domain called Ets, which recognizes a purine rich sequence that has a central 5'-GGAA/T-3' core. The Ets domain can be produced as a protein fragment that can independently fold into a stable structure, and it is sufficient to bind DNA by itself. The Spi-C protein is member of the subfamily Spi and is characterized by being expressed during different stages of B cell development and also in macrophages. There are some published works describing that Spi-C recognizes some of the DNA sequences which were previously studied for PU.1 protein, a member of the same subfamily. The expression of Spi-C overlaps the one of PU.1 during B cell development. There are genes involved in B cell development whose promoters have Ets binding sites. Binding studies for Spi-C can contribute to some issues about this transcription factor function and the regulation of these genes transcription. In this work, we present the heterologous expression of human Spi-C in *Escherichia coli* and its purification by FPLC yielding approximately 0.7 mg of protein per cell gram. Here, to investigate new target promoter sequences for Spi-C, DNA binding studies were done using sequences localized at promoters of genes that codify for the α subunit of the IL-7 receptor and for the B cell receptor complex transmembrane protein Ig α . These proteins participate in the development of B cells and are essential for their selection in the bone marrow. Our results show that Spi-C Ets domain is capable of recognizing and binding these DNA sequences by electrophoretic mobility shift assays forming different sized complexes. We suggest that Ets domain can bind to DNA sequences *in vitro* in a proportion higher than 1:1; the higher was the domain concentration, the more retarded were the protein-DNA complexes. Surprisingly, the Ets domain was not able to discriminate between the wild sequences, from the ones with mutations in the central core, maintaining the same shift pattern, suggesting that for DNA recognition the other parts of the protein are fundamental.

Keywords: Spi-C; B cell development; Ets domain; DNA binding.

INTRODUÇÃO

1. A família de fatores de transcrição Ets

As proteínas pertencentes à família Ets são fatores de transcrição que regulam positivamente ou negativamente a expressão de genes que estão envolvidos em vários processos biológicos na fase adulta e embrionária¹. As proteínas Ets controlam genes relacionados aos processos de proliferação, transformação, diferenciação celular, angiogênese, remodelamento tecidual, hematopoiese, apoptose, invasão e metástase^{2,3,4,5}.

A família Ets de proto-oncogenes foi originalmente identificada por sua homologia com a seqüência do oncogene *v-ets* do retrovirus E26⁶ que causa leucemia eritoblástica e mieloblástica em frangos⁷. Na bibliografia existe divergência quanto ao significado da sigla *Ets*. LePrince e colaboradores (1983) definem como **E**-**t**wenty-**s**ix **s**pecific⁶, Nunn e colaboradores (1983) se referem a **E**26 **t**ransformation-**s**pecific⁸ e a expressão **E**-**t**wenty-**s**ix foi citada por Sharrocks (2001)¹. Já na base de dados online de domínios conservados do “The National Center for Biotechnology Information” NCBI está definida como **E**rythroblast **T**ransformation **S**pecific⁹.

Esta família é composta por aproximadamente trinta genes^{4,5} que codificam fatores de transcrição eucarióticos que se ligam a seqüências específicas de DNA com uma região evolutivamente conservada de aproximadamente 85 aminoácidos, o domínio Ets^{10,11}. Os membros desta família estão presentes em organismos do reino Metazoa, desde esponjas até humanos¹², podendo desempenhar um papel tanto ativador, como repressor da transcrição de genes¹³.

A expressão inapropriada das proteínas Ets ou a expressão de proteínas Ets fusionadas com outras proteínas são importantes na formação de alguns tipos de câncer^{1,5}. Geralmente, os tumores exibem ausência de regulação das rotas de sinalização, resultando em ativação contínua e independente de ligação. Como

várias proteínas Ets são alvos finais destas rotas de sinalização, a função das mesmas é significativamente alterada no câncer.

Devido ao fato de alguns fatores de transcrição contribuírem para inibição de apoptose, é possível que altos níveis de fatores de transcrição Ets em células tumorais influenciem na resistência à terapia contra o câncer. Considerando estas afirmações, o nível de expressão de alguns fatores de transcrição Ets pode representar um marcador de prognóstico e/ou diagnóstico para pacientes oncológicos⁴.

1.2 Os domínios existentes nas proteínas Ets

O domínio de ligação ao DNA das proteínas da família Ets, chamado domínio Ets, reconhece uma seqüência rica em purinas, cujo consenso central é formado pela seqüência 5'-GGAA/T-3'^{11,14}. Estruturalmente, o domínio Ets se apresenta como “winged” hélice-volta-hélice (“winged helix-turn-helix”-wHTH) e é composto de três α -hélices e quatro folhas β -pregueadas, dispostos da seguinte maneira: α 1- β 1- β 2- α 2- α 3- β 3- β 4¹⁵. O domínio Ets pode ser produzido como um fragmento de proteína estável e é suficiente para que ocorra ligação ao DNA, sendo, portanto, não só uma unidade estrutural, mas também funcional^{13,16}.

As proteínas Ets estão classificadas em subfamílias, tendo como base a homologia entre os domínios Ets e a composição estrutural, além da presença de outros domínios conservados na seqüência da proteína^{1,5,13}. Algumas das subfamílias contêm os domínios Ets na porção C-terminal e outras na porção N-terminal¹⁷.

Os fatores de transcrição podem regular a expressão gênica em duas etapas. A primeira etapa é a ligação do fator de transcrição a um elemento com seqüência específica de “enhancer” e/ou promotor. A segunda etapa consiste na interação do

domínio de ativação do fator com a maquinaria de transcrição basal, resultando na regulação da taxa de transcrição para um determinado gene. Para a maioria dos fatores de transcrição, essas funções se encontram em partes diferentes da proteína, sendo importante definir os domínios que participam da ligação ao DNA e os domínios de transativação¹⁸. O domínio de transativação (TAD – domínio ácido de transativação) está presente em algumas proteínas Ets na sua porção N-terminal e é rico em glutamina e vários resíduos ácidos que, como o nome sugere, é necessário para a transativação¹⁹ (Figura 1).

Algumas subfamílias contêm outro domínio além dos Ets e TAD, o qual também é evolutivamente conservado, denominado “pointed” (PNT), que também está localizado na região N-terminal (Figura 1). O domínio PNT tem estrutura de hélice-alça-hélice (“helix-loop-helix”-HLH). O domínio PNT contém 65-85 aminoácidos e está presente em mais ou menos onze membros da família Ets. Sua função está relacionada com interações proteína–proteína e oligomerização. Análises estruturais deste domínio indicam que este forma uma estrutura independente de um feixe monomérico de cinco-hélices^{3,5,20}.

Subfamily (member)

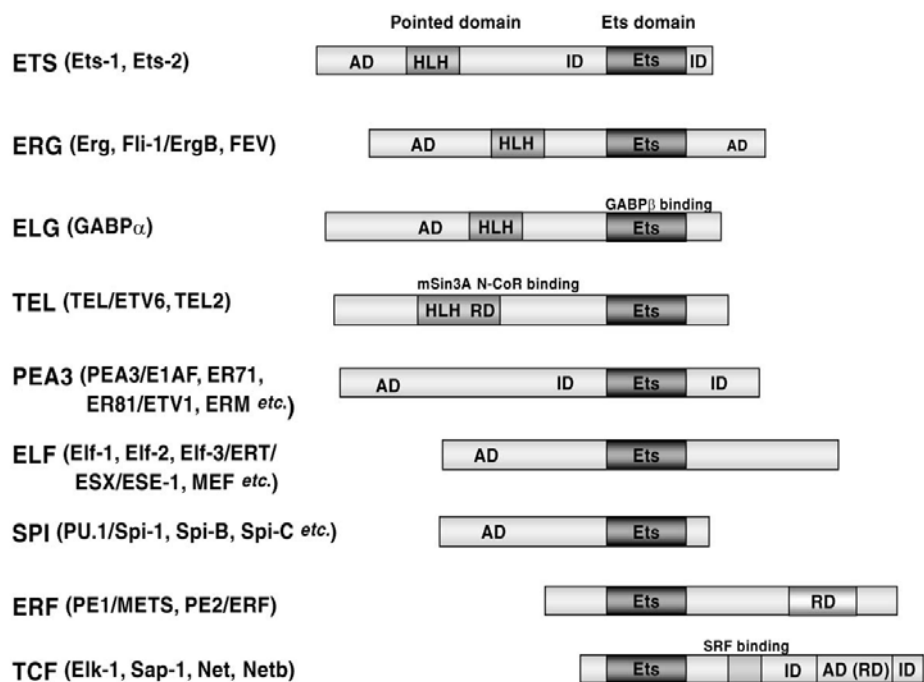


Figura 1. Esquema dos domínios existentes nos membros da família de proteínas Ets. Ets: Domínio Ets de ligação ao DNA; HLH: domínio “helix–loop–helix” (domínio pointed); AD: domínio de ativação (domínio TAD); ID: domínio de auto-inibição; RD: domínio de repressão. Fonte: OIKAWA e YAMADA (2003)

1.3 Interação entre proteínas Ets e o DNA

Tendo como modelo a determinação do domínio Ets da proteína PU.1 ligada ao DNA por cristalografia, verificou-se que a terceira α -hélice é responsável pelo contato com o sulco maior do DNA, onde se encontra o consenso GGAA/T. Dois resíduos de arginina altamente conservados entre os membros da família Ets, estão localizados nesta α -hélice e fazem contato direto com os nucleotídeos GGA do consenso (Figura 2). Outros dois resíduos fazem contato direto com o DNA: uma lisina localizada na “wing”, entre β 3 e β 4, se liga “upstream” do consenso GGAA e outra lisina da “volta” da estrutura hélice-volta-hélice se liga “downstream” e na fita oposta.

Domínio Ets de PU.1

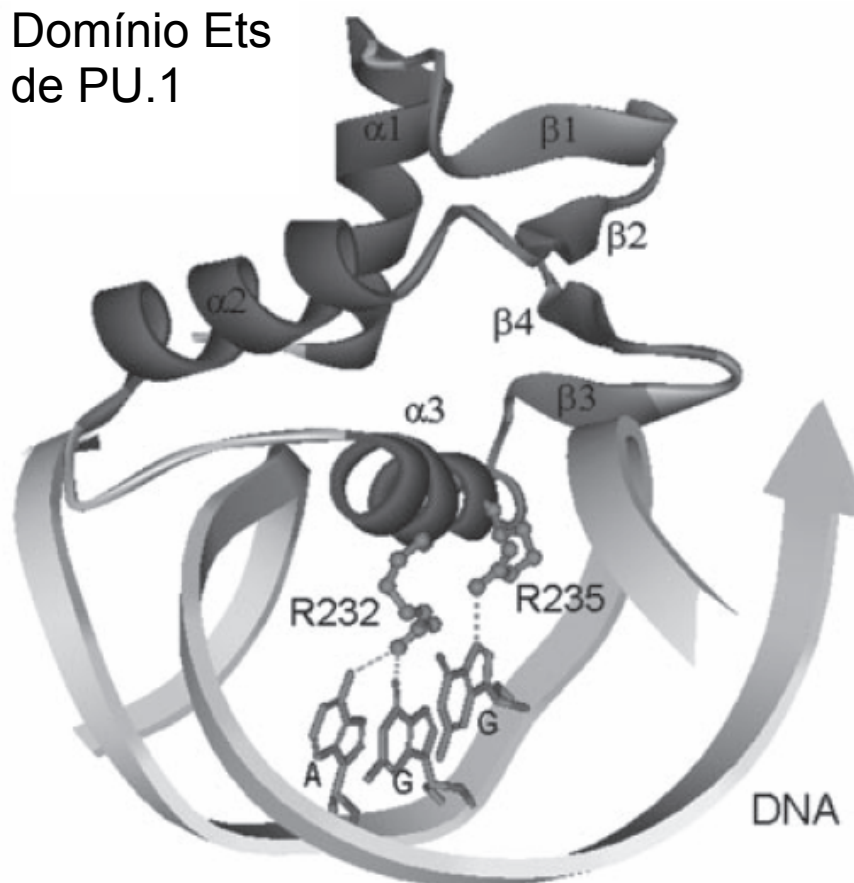


Figura 2. Estrutura terciária do domínio Ets de PU.1 (Protein Data Bank ID 1PUE). A estrutura completa é composta por quatro folhas β -pregueadas antiparalelas deitadas numa região de hélice-volta-hélice. A hélice3 (α 3) se acomoda no sulco maior do DNA e reconhece o consenso GGAA. O reconhecimento ocorre principalmente por meio de pontes de hidrogênio entre o consenso e as argininas. As voltas entre β 3 e β 4 e entre α 2 e α 3 interagem com os grupos fosfatos do sulco menor e estabilizam o complexo. Fonte: LISTMAN et al. (2005)

Diferentes membros da família de proteínas Ets demonstram especificidades diferentes de ligação ao DNA. Os aminoácidos adjacentes ao domínio Ets exercem influência na afinidade e especificidade de ligação, assim como as seqüências adjacentes ao consenso central 5'-GGAA/T-3' do DNA¹³. A partir de substituições nas argininas conservadas do domínio Ets de PU.1, que se ligam diretamente ao DNA, demonstrou-se que elas são necessárias, não somente para a sua interação com o DNA, mas também para interação com outras proteínas transativadoras que, como ela, se ligam ao DNA formando um complexo funcional. Além disso, as argininas também são responsáveis pelo sinal de localização nuclear. Isso pôde ser observado em proteínas cujas argininas foram mutadas, as quais permaneceram no

citoplasma celular²². Além dessas interações entre o domínio Ets e o DNA alvo, os contatos de fosfato que se estabelecem flanqueando o consenso GGA são importantes para fornecer maior estabilidade ao complexo domínio Ets-DNA¹³.

Muitos dos fatores de transcrição Ets sofrem mecanismos de auto-regulação que inibem a sua ligação ao DNA, o que teria a função de prevenir a ligação desnecessária ao DNA. O desbloqueio da ligação proteína-DNA pode ocorrer por meio de mecanismos de fosforilação, acetilação e interações com fatores de transcrição corregulatórios para ligação a seqüências específicas e reconhecimento de determinados promotores^{1,5}.

1.4 Spi-C

Spi-C pertence à família de fatores de transcrição com domínio Ets e é um membro da subfamília Spi, já que a seqüência de seu domínio Ets tem maior identidade com os domínios dos outros membros da mesma, PU.1, Spi-B, Spi-D. Além dessas proteínas, a Spi de lampreia, considerada evolutivamente a mais primitiva, também está incluída nessa subfamília pela alta identidade. A partir dos estudos realizados até o presente, acredita-se que o grupo Spi exerça um papel importante no desenvolvimento de linfócitos e se presume que tenha evoluído de um ancestral comum durante a transição entre vertebrados que possuem e não possuem mandíbula. A presença dos diferentes membros do grupo Spi sobrepõe-se à ocorrência dos linfócitos durante a evolução. Portanto sugeriu-se que as proteínas do grupo Spi e os linfócitos co-evoluíram^{23,24}.

Entre todas as subfamílias Ets, a subfamília Spi é a que apresenta menor identidade entre os genes da família Ets¹⁷. Embora a identidade entre os domínios Ets dos membros da subfamília Spi seja significativa, a Spi-C é o membro que mais difere dentro dessa subfamília, sendo os aminoácidos que flanqueiam o seu domínio

Ets únicos, podendo até ser considerada membro de uma nova subfamília^{25,26}. Portanto, até o presente, Spi-C é a proteína que apresenta maior divergência entre todas as proteínas da família Ets. Ao analisar a identidade entre as α -hélices dos domínios dos membros das subfamília Spi constatou-se que a terceira delas é a que apresenta maior identidade entre os membros PU.1, Spi-B e Spi-C²⁵ sugerindo que esses fatores de transcrição podem compartilhar seqüências alvo de DNA.

A proteína Spi-C está presente em camundongos^{25,27}, humanos²⁶, ciclídeo (peixe de estrutura óssea)²³ e raia (peixe cartilaginosa)²⁴. O gene *SPIC* humano está localizado no cromossomo 12q22-24 e contém seis exons, sendo cinco deles codificantes^{25,26}.

Em 1999 foi publicada, simultaneamente, a descoberta da proteína Spi-C de camundongo por dois grupos. Com isso, a proteína recebeu dois nomes: Spi-C (por ser posterior a Spi-B)²⁵ e Prf (por “PU.1-related factor”)²⁷. A primeira nomenclatura, proposta por Bemak e colaboradores, é a mais utilizada na comunidade científica.

Quantidades significativas de RNA mensageiro (RNAm) de Spi-C foram detectadas por meio de RT-PCR em tecidos de órgãos linfóides (baço, linfonodos e medula óssea) de camundongo, onde há presença de linfócitos B maduros, enquanto baixos níveis foram detectados em macrófagos recém isolados e em uma linhagem celular de monócito-macrófago (J774). Diferentes tipos celulares também foram examinados e o mesmo não ocorre em células B em diferenciação ou em células B maduras presentes no plasma²⁵. Outro estudo similar mediu o RNAm de Spi-C em vários estágios de diferenciação de células B na medula óssea e detectou quantidades significativas deste RNA em células B precursoras (pré-B) e B maduras. O mesmo RNAm, entretanto, não foi observado em células B progenitoras (pró-B) e B imaturas²⁷. A Figura 3 esquematiza as etapas do desenvolvimento de células B para maior esclarecimento sobre o assunto.

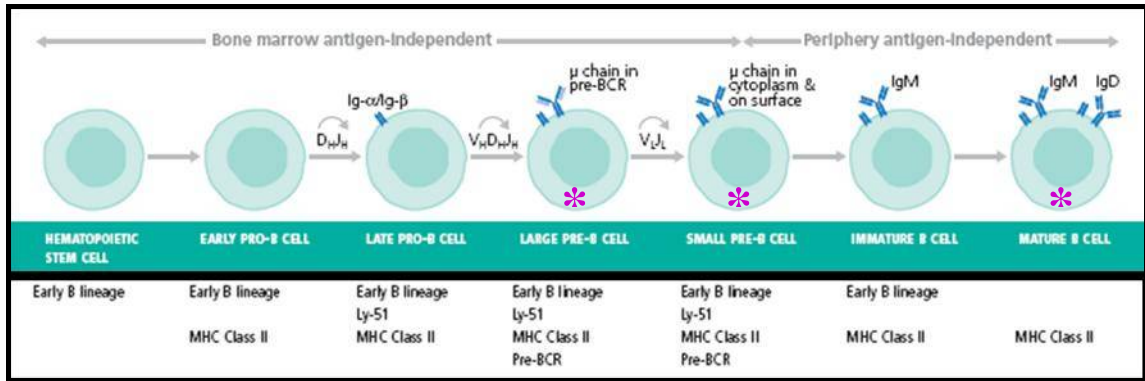


Figura 3. Etapas do desenvolvimento de linfócitos B desde a célula tronco hematopoiética até a célula B madura. RNAm de Spi-C foi detectado nos estágios de célula B marcados com ✿. Fonte: <http://www.bdbiosciences.com>

Carlsson e colaboradores verificaram que a proteína Spi-C humana segue o mesmo perfil de expressão que a de camundongo, indicando que há quantidades significativas de RNAm de Spi-C em tecidos linfóides. Foram detectados, menor quantidade em células de linfoma de células B (Raji), nenhum nível em células linfóides T (MOLT-4) e em células mielóides (K562 e HL-60) (Figura 4).

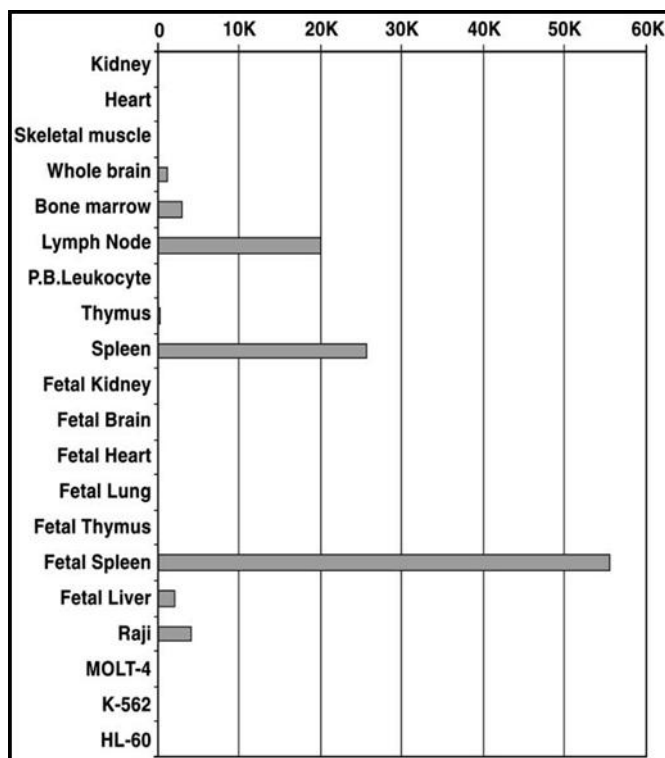


Figura 4. Perfil da quantificação relativa de RNAm de Spi-C humana por RT-PCR. Maior quantidade foi detectada em tecidos linfóides: linfonodos, baço e medula óssea. Fonte: CARLSSON et al. (2002)

Estes fatos indicam que Spi-C deve ter uma função específica na regulação da diferenciação e maturação dos linfócitos B. Além disso, os resultados sugerem que Spi-C tem uma função diferente dos outros membros da mesma subfamília que são significativamente expressos em todas as etapas ou em fases diferentes da diferenciação dos linfócitos B, como também em outros tipos celulares (linfócitos T e células mielóides)²⁶.

A atividade de transativação de Spi-C de camundongo se deve à região N-terminal que contém os primeiros 21 aminoácidos. Esta fração da proteína tem aminoácidos ácidos conservados em todos os representantes já estudados. Assim, pode-se afirmar que este domínio de transativação é importante para a função dos ortólogos de Spi-C¹⁹.

Tanto Spi-B quanto PU.1 possuem uma região rica em prolina, glutamato, serina e treonina chamada PEST, que interage com PIP (PU.1 Interaction Partner). Geralmente, essa região está envolvida na sinalização para degradação de

proteínas intracelulares, mas esta função é desconhecida em PU.1. PIP é essencial para diferenciação terminal de linfócitos B e T. Desta forma, camundongos com silenciamento do gene que codifica para PIP ($PIP^{-/-}$) não geram anticorpos nem têm resposta citotóxica de células T. Para verificar se a proteína Spi-C interage com PIP, foram realizados estudos por detecção de formação de um complexo tendo como DNA alvo a seqüência “enhancer” da cadeia leve λB na região λ_{2-4} de imunoglobulina. Esse complexo é formado pela interação das proteínas PIP-Spi-B ou PIP-PU.1 com o DNA alvo. Entretanto, foi demonstrado que a Spi-C se liga ao DNA, mas não interage com PIP *in vitro*, ao contrário das outras proteínas pertencentes a sua subfamília, como também não afeta a formação do mesmo complexo com PU.1. Esses resultados levam à hipótese de que como Spi-C é expressa de forma diferente e específica nas células B, essa proteína poderia atuar como um mediador de controle negativo, competindo com PU.1 ou Spi-B pelo sítio de ligação. Essa mediação poderia ocorrer devido ao seu potencial de transativação ser consideravelmente menor do que o do complexo PU.1-PIP ligado ao DNA alvo. Desse modo, sua expressão na célula pode ser diminuída nos casos em que for necessário o acesso de outras proteínas Ets (PU.1) e seus cofatores (PIP) a elementos de controle da imunoglobulina¹⁹.

Um estudo de análise do papel funcional de Spi-C, utilizando um sistema de dois híbridos e tendo a Spi-C humana como “isca”, demonstrou que Spi-C interage, especificamente, com a porção C-terminal da proteína STAT6 (Tradutor e Ativador de Sinal de Transcrição 6) em levedura e em células COS7 (células epiteliais de fígado de macaco). Além disso, essa interação proteína-proteína é funcional, já que foi demonstrado que a transcrição do promotor de IgE ($I\epsilon$) estimulada por IL-4 tem transativação sinérgica realizada pelas proteínas SPI-C e STAT6, quando ambas estão ligadas ao DNA. Essa interação representa um novo mecanismo para ativação

de genes alvo dependentes de STAT6 e IL-4 no processo de diferenciação de linfócitos B. Vale ressaltar que a PU.1 também forma o mesmo complexo, sugerindo uma possível competição entre Spi-C e PU.1, tanto pela ligação ao promotor, quanto pela ligação direta com STAT6²⁸. A citocina IL-4 induz a mudança de isotipo para IgE em células B. A transcrição de IgE ocorre em células B em desenvolvimento em resposta a IL-4 nos centros germinais nos órgãos linfóides secundários²⁹.

Schweitzer e colaboradores realizaram estudos comparando a atividade de Spi-C humana com a proteína PU.1 humana em relação à ligação a DNAs alvos, utilizando seqüências que foram previamente descritas para PU.1 e que controlam a transcrição de genes específicos. Esse estudo relacionou a transcrição dos genes e de marcadores celulares de desenvolvimento de célula B na presença ou ausência de ambos os fatores de transcrição a partir de células pro-B derivadas de fígado fetal. Um dos DNAs alvo estudados para ligação, por meio de ensaio de ligação ao DNA por eletroforese em gel, foi a seqüência *IgH* intronic enhancer. Quando sua transcrição é ativada ocorre processo de rearranjo e expressão de IgH. Outro DNA alvo foi o promotor de receptor *FcγRIIb*, cuja expressão na superfície das células B se dá em todo o processo de desenvolvimento das mesmas e na célula B madura, tendo como função bloquear a proliferação através da inativação de MAPKs (proteíno-quinases ativadas por mitógenos). Os transcritos destes genes também foram analisados por PCR em tempo real e os marcadores celulares por citometria de fluxo. Também foram utilizados ensaios de transfecção de vetor contendo gene repórter da luciferase para observar a atividade relativa dos fatores de transcrição nos genes alvo mencionados acima.

Os resultados destes experimentos indicam que a proteína Spi-C promove a diferenciação de células B e que a sua expressão em estágios específicos pode ser importante para o desenvolvimento das mesmas. Foi demonstrado que Spi-C regula

negativamente o gene *FcγRIIb*, fato interessante para estágios que requerem proliferação, como em pré-B e células B maduras, sugerindo que a mesma atua por bloquear os efeitos antiproliferativos de PU.1 e do receptor *FcγRIIb*. Outro comportamento contrário à proteína PU.1 é o aumento da transcrição de IgH e de sua presença na superfície celular, significando a diferenciação para células B maduras. Em relação à ligação ao DNA, tanto PU.1 quanto Spi-C podem interagir com sítios definidos dos genes já mencionados. Spi-C pode inibir direta ou indiretamente a ligação de PU.1 ao DNA, dado que confirma a idéia de que Spi-C e PU.1 têm atividades antagônicas. Portanto, a Spi-C pode antagonizar o efeito de PU.1 para promover a diferenciação das células B³⁰.

1.5 Células B e possíveis seqüências alvo para Spi-C

A citocina IL-7 é um fator de crescimento produzido por células estromais da medula óssea, pelo fígado fetal, timo e por células epiteliais para promover a sobrevivência, proliferação e diferenciação dos linfócitos B e T^{31,32}. O receptor de IL-7 (IL-7R) é expresso na superfície dessas células desde as células progenitoras na medula óssea até as células pré-B grandes³². Foi demonstrado, que o desenvolvimento de células B é interrompido no estágio de célula pré-pró B na ausência da sinalização proveniente do IL-7R³³. Foi descoberto que os fatores de transcrição Ets PU.1 e GABP interagem com o sítio Ets do promotor da subunidade alfa do IL-7R (IL-7R α) para ativar a transcrição, resultando em promoção do desenvolvimento de célula B³⁴.

A inserção do complexo do receptor de células pré-B (pBCR) na membrana é necessário para a transição para o estágio de célula pré-B e é um ponto de controle no desenvolvimento de células B. Na ausência de complexos pBCR ou BCR (em célula B madura) intactos, a maturação de células B pode ser alterada ou

interrompida. As diferentes partes integrantes do pBCR são transcritas e expressas em células pró-B, incluindo as cadeias sinalizadoras Ig α e Ig β ³⁵. Ig α é uma glicoproteína transmembrana codificada pelo gene *mb1* e é necessária para a ancoragem das cadeias de imunoglobulina “pesada” na membrana plasmática de célula B e para a sinalização transmembrana após estimulação do BCR pelo antígeno³⁶. Existem seqüências “upstream” adjacentes ao sítio de iniciação da transcrição que são reconhecidas por proteínas Ets como Ets1, Ets2 e PU.1. Estas proteínas se ligam a estas seqüências *in vitro*³⁷.

Células pró-B são capazes de proliferar em resposta à IL-7, na ausência de outras células ou fatores. A expressão de pBCR em células pró-B (necessária para a transição de estágio) possibilita a resposta em menores concentrações de IL-7. A proliferação de células que expressam o pBCR (células pré-BCR+) em sua membrana plasmática em resposta a baixas concentrações de IL-7 e sinais mediados pelo IL-7R elevam a densidade de colônias de células pré-B e promovem a expressão do BCR maduro. Para seleção de células pBCR+, o microambiente da medula óssea estabelece a condição de que células pró-B necessitam proliferar em níveis que estão abaixo do limiar de concentração de IL-7. A Figura 5 ilustra a seleção para proliferação de células B em desenvolvimento no microambiente da medula óssea.

Tanto o IL-7R quanto o pBCR geram sinais para a ativação da via ERK/MAPK (ERK: quinases reguladas por sinais extracelulares) que é requerida para a proliferação de células B em resposta à baixos níveis de IL-7. Isso proporciona um mecanismo onde somente a célula contendo um pBCR funcional seja capaz de sobreviver, expandir-se e proliferar em colaboração com os sinais do ambiente³⁵.

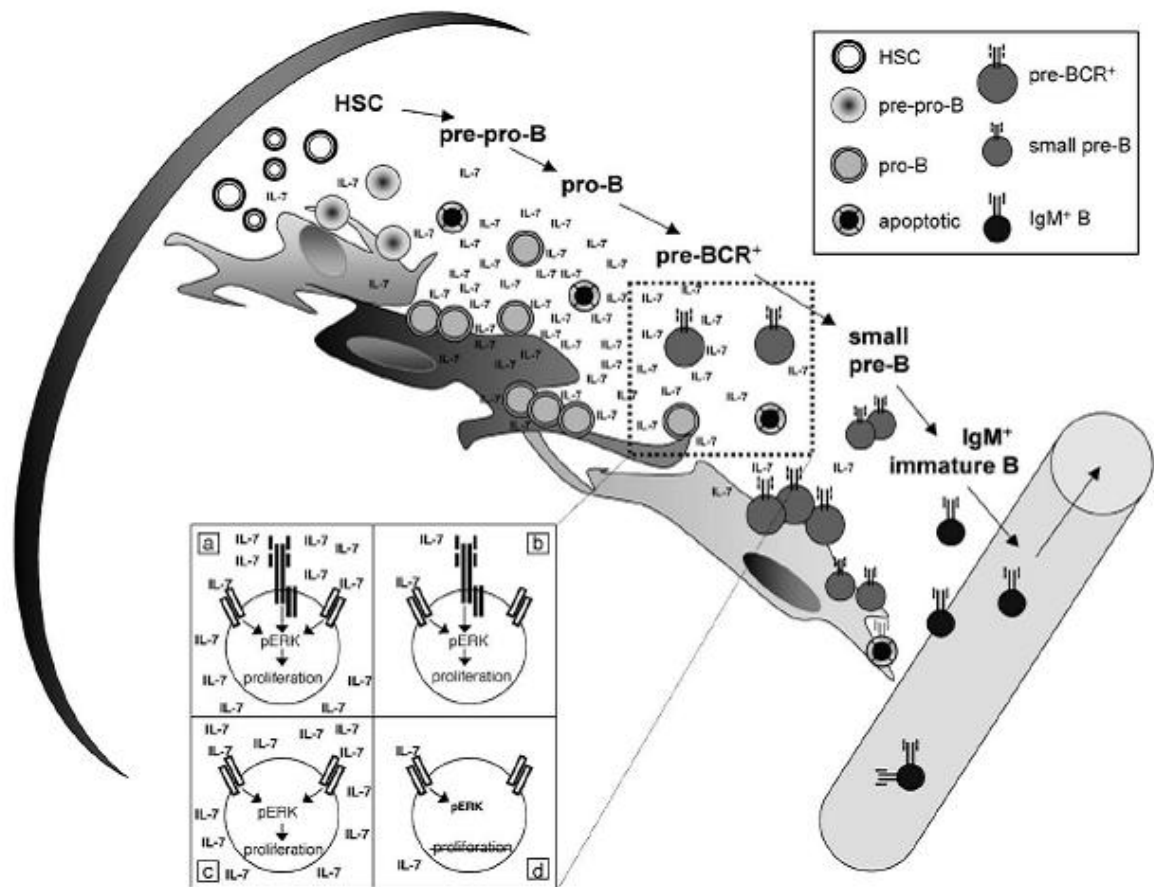


Figura 5. Modelo da compartimentalização de IL-7 na medula óssea. Células-tronco hematopoiéticas (HSC) localizam-se na periferia do osso e células B precursoras em desenvolvimento migram em direção à região central através de modificações nas áreas físicas em relação à produção, liberação e/ou acesso de IL-7 de células estromais. As áreas de alta concentração de IL-7 suportam a expansão de células pró-B antes e durante a recombinação da cadeia pesada. A área adjacente que contém IL-7 em quantidades limitadas seleciona as células que expressam o pBCR e restringe a população de pró-B ao compartimento com alta concentração de IL-7. A proliferação de células pBCR+ em resposta a baixas concentrações de IL-7 eleva a densidade de colônias de células pré-B e promove a expressão do BCR. Células que expressam o BCR continuam a seleção e migram para a periferia via o sinus central. Quadro ampliado: compartimentos que contém altas concentrações de IL-7 (a,c), quantidade suficiente de ERK1,2 é fosforilada para alcançar o limiar para proliferação com ou sem expressão de pBCR. Em regiões onde IL-7 é limitante (b,d) menor é a quantidade de sinais provenientes dos complexos IL-7/IL-7R para ativar ERK1,2. Se o pBCR é expresso (b), os sinais “downstream” do mesmo se somam com a quantidade de ERK1,2 que são fosforiladas, permitindo que as células proliferem. Entretanto, células pró-B pBCR- não têm fonte adicional de ERK (d) e, com isso, não se dividem neste ambiente. Fonte: FLEMING e PAIGE (2002)

Levando em consideração a importância dos genes que codificam tanto para IL-7R α como para Ig α pode-se afirmar que a regulação da transcrição dos mesmos deve ser importante para formação de respostas imunes apropriadas.

Os trabalhos anteriormente descritos, utilizando diferentes seqüências para ligação de Spi-C foram baseados em seqüências previamente estudadas para PU.1,

já que os mesmos pertencem à mesma subfamília e sua expressão ocorre simultaneamente em alguns estágios de células do sistema imune.

Contudo, o estudo de sítios para ligação de proteínas Ets nos promotores de genes envolvidos no desenvolvimento de células B quanto à ligação da Spi-C pode contribuir para acrescentar dados sobre a função desse fator de transcrição e a regulação da transcrição dos genes em questão.

OBJETIVOS

Objetivo geral

Purificar o domínio Ets da Spi-C humana e caracterizar a ligação entre o domínio Ets do fator de transcrição Spi-C e o DNA alvo por espectrometria de ressonância plasmônica de superfície e mobilidade de DNA em gel por eletroforese. Utilizar seqüências promotoras importantes no desenvolvimento de linfócitos B para obter mais informações sobre a função de Spi-C.

Objetivos específicos

- Construir o gene que codifica o domínio Ets de Spi-C humana;
- Clonar o gene no vetor pCR[®]-Blunt (Invitrogen);
- Subclonar no vetor de expressão pET23a(+) (Novagen);
- Expressar domínio Ets de Spi-C em *Escherichia coli* na forma solúvel;
- Purificar a proteína recombinante por meio de Cromatografia Líquida de Rápida Eficiência (FPLC);
- Realizar análises por espectrometria de massas e seqüenciamento N-terminal;
- Pesquisar seqüências de DNA interessantes para testes de ligação ao DNA;
- Realizar ensaios de ligação utilizando ressonância plasmônica de superfície (SPR) e espectrofluorimetria para obter constantes de ligação;
- Realizar ensaios de retardo da migração do DNA em gel (EMSA) a promotores de linfócitos B e seus mutantes.

ARTIGO CIENTÍFICO PARA SUBMISSÃO

Título: Expression and purification of Ets domain of hSpi-C transcription factor: binding assays to B lymphocyte promoters

Periódico: Molecular and Cellular Biochemistry: An International Journal for Chemical Biology in Health and Disease. Springer Netherlands.

Expression and purification of Ets domain of hSpi-C transcription factor: binding assays to B lymphocyte promoters

Ekaterini Gianniotis^{1,2}, Mauricio R. Bogo^{1,3}, Andrea B. Scholles², Claudia P. Nunes², Gaby Renard², Jocelei Chies², Luiz A. Basso^{1,2}, Diógenes Santiago Santos^{1,2}

(1) Programa de Pós-Graduação em Biologia Celular e Molecular, PUCRS. Av. Ipiranga, 6690, Prédio 12, Partenon, Porto Alegre, 90610000, Brazil.

(2) Centro de Pesquisas em Biologia Molecular e Funcional, Instituto de Pesquisas Biomédicas, PUCRS. Av. Ipiranga, 6681, Tecnopuc, Prédio 92A, Partenon, Porto Alegre, 90619900, Brazil

(3) Laboratório de Biologia Genômica e Molecular, PUCRS. Av. Ipiranga, 6690, Prédio 12C, sala 172, Partenon, Porto Alegre, 90610000, Brazil

Abstract

The members of the Ets family of transcription factors contain a DNA binding domain called Ets which recognizes a purine rich sequence that has a central 5'-GGAA/T-3' core. The Ets domain can be produced as a protein fragment which can independently fold into a stable structure and it is sufficient to bind DNA by itself. The Spi-C protein is member of the subfamily Spi and is characterized by being expressed during different stages of B cell development and in macrophages. There are some published works describing that Spi-C recognizes some of the DNA sequences which were previously studied for PU.1 protein, member of the same subfamily. The expression of Spi-C overlaps the one of PU.1 during B cell development. There are genes involved in B cell development whose promoters have Ets binding sites. For these, binding studies with Spi-C can contribute information on this transcription factor function and these genes transcriptional regulation. In this work we present the heterologous expression of human Spi-C in *Escherichia coli* and its purification by FPLC yielding to approximately 0.7 mg of protein per cell gram. Here, to investigate new target promoter sequences for Spi-C, DNA binding studies were done using sequences localized at promoters of genes that codify for the α subunit from the IL-7 receptor and for the B cell receptor complex transmembrane protein Ig α . These proteins participate in the development of B cells and are essential for their selection in the bone marrow. Our results show that the Spi-C Ets domain is capable of recognizing and binding these DNA sequences by electrophoretic mobility shift assays forming different sized complexes. We suggest that the Ets domain can bind to DNA sequences *in vitro* in a proportion higher than 1:1, where after all the DNA fragments were already occupied proteins continued binding instead of staying free in the medium and formed higher and more retarded complexes. Surprisingly, the Ets domain was incapable of discriminating between the wild sequences from the ones with mutations in the central core maintaining the same shift pattern, suggesting that the other parts of the protein are fundamental for DNA recognition.

Keywords: Spi-C; B cell development; Ets domain; DNA binding

Introduction

The recognition of DNA is one of the most important themes of study implicated in a large number of cellular processes, being mainly involved in regulation of fundamental steps in gene expression, when regulatory proteins specifically recognize promoter or enhancer DNA sequences. One of the most studied and known proteins are the Ets transcription factors, due to importance in various biological processes [1]. This family of transcription factors shares a conserved winged helix-turn-helix DNA binding domain [2]. The Ets domain contains approximately 85 amino acids, which suggests conserved functional properties [3,4]. Despite considerable variation among family members, their major characteristic is the recognition of sequences containing the core GGAA/T, the Ets binding site [4,5]. In general, both the full Ets proteins and the Ets domains bind as a monomer to the core sequence, the Ets binding site [6].

The Spi-C protein, a member of the Ets family of transcription factors belongs to the Spi subfamily. It is present in different species of Metazoa including humans. The expression pattern and homology among orthologues suggest that these Spi proteins have important regulatory functions during the evolution of the adaptive immune system. The *SPIC* gene is localized to chromosome 12q22-24 consisting in six exons, one non-coding and five coding exons [7]. This protein is expressed in particular stages during B cell development and in macrophages [8,9]. DNA regulatory sequences of B cell differentiation that are targets for PU.1, a subfamily co-member, have been shown to be targets as well for Spi-C, showing that this transcription factor regulates the expression of development related genes in B cell. Furthermore Spi-C activity seems to promote B cell differentiation [10,11,12].

The *mb1* gene encodes Ig α , an early B-cell-specific transmembrane protein required for docking of immunoglobulin heavy chains of pre-BCR (pre-B cell receptor) and BCR (mature B cell receptor) in the B-cell plasma membrane and for transmembrane signaling after stimulation of the BCR by antigen [16]. The different pre-BCR elements are transcribed and expressed in pro-B cells, including the signaling chains Ig α and Ig β . Membrane insertion of the pre-BCR complex is required for the transition to the pre-B cell stage, and is a checkpoint in the B cell development. In the absence of intact pBCR or BCR complexes, maturation can be altered or arrested. Pro-B cells are capable of proliferating in response to IL-7 in the absence of other cells or factors. pBCR expression in pro-B cells turns more sensitive the response to IL-7, resulting in a reduced threshold of IL-7 responsiveness in pro-B cells.

Proliferation of pBCR⁺ cells in response to low concentrations of IL-7 and the IL-7R-mediated signals enhances the density of pre-B cells colonies and promotes BCR expression. For selection, the bone marrow microenvironments establish the condition that pro-B cells need to proliferate in IL-7 levels that are below threshold. At this condition, only the B lineage cells expressing a pBCR will proliferate and continue developing [17].

Synthetic oligonucleotides containing presumptive Ets binding sites have been used world wide in electrophoretic mobility shift assays (EMSAs), DNase I footprinting assays, and most recently the surface plasmon resonance (SPR) as an alternative to conventional biochemistry methods for studying DNA-protein affinity interactions [13,14]. In order to elucidate new Spi-C binding sequences we chose two DNA sequences which were previously shown to bind PU.1 protein and are involved in

B cell development. One of the target sequences is the IL-7R α promoter which was discovered to interact with the Ets factors PU.1 and GABP within the Ets site to activate transcription for promotion of B cell development [15]. Other interesting sequence which has two Ets sites is the adjacent upstream sequences of transcription initiation sites of *mb1* gene.

In this study we present the construction, cloning, expression, purification and EMSA studies of the human Spi-C (hSpi-C) Ets DNA binding domain to promoter and enhancer DNA sequences of genes involved in B cell development. We demonstrate that human Spi-C can recognize and bind to these promoter sequences *in vitro*.

Materials and methods

DNA synthesis and cloning of coding sequence for hSpi-C Ets domain

DNA sequence synthesis was performed by designing ten oligonucleotides (Invitrogen™, São Paulo, Brazil) based on the hSpi-C coding sequence (GenBank accession number NM_152323) which were overlapped and extended by a *Pfu Turbo*® DNA Polymerase (Stratagene, La Jolla, US) in a multi-step technique according to the patent application PCT/Br 2006/000267 [18] to obtain the coding sequence. Primers 1 and 10 were designed to contain, respectively, *NdeI* and *HindIII* restriction sites (underlined).

The oligonucleotides sequences were as follows:

1. 5'-AGCATATGAAAGGCAGGAAGAAGCTC-3'
2. 5'-AAAGGCAGGAAGAAGCTCCGACTGTTTGAATACCTTCACGAATCCC-3'
3. 5'-CCCCTGAATACAAGATGCCATCTCCGATTATACAGGGATTCGTG-3'
4. 5'-ATTCAGTGGGTAGATAAAACCAAAGGCATCTTTCAGTTTGTATCAAAAAA-3'
5. 5'-GCCTTTTCTTTTCCCCCAAAGCTCGGCAAGTTTTTCTTTGTTTTTGTATA-3'
6. 5'-AAGAAAAGGCAACAGGAAGACCATGACTTACCAGAAAATGGCCAGGGCAC-3'
7. 5'-CGGATTTTGGTAATTTCCCCTTCTTCCGTAATTTCTGAGTGCCCTGGC-3'
8. 5'-CCAAAATCCGGAGGAAGCTGACTTACCAGTTCAGTGAGGCCATTCTCCAA-3'
9. 5'-CTCTTTCCCCAGGAAATAGGATGGAGAGAGTCTTTGGAGAATG-3'
10. 5'-TAAAGCTTCACTCTTTCCCCAGGAA-3'

The complete sequence was purified from 2% agarose gel (stained with 0.5 $\mu\text{g mL}^{-1}$ of Ethidium Bromide) using PureLink™ Quick Gel Extraction kit (Invitrogen™, Carlsbad, CA, USA), cloned in pCR®-Blunt vector (Invitrogen™, Carlsbad, CA, USA). *Escherichia coli* (*E.coli*) DH10B (Invitrogen™, Carlsbad, CA, USA) was transformed and selected in Luria-Bertani (LB) medium with Kanamycin 50 $\mu\text{g/mL}$. The QIAprep® Miniprep kit (Qiagen, Santa Clarita, CA) was used to extract recombinant plasmids from the cells. After digestion with *NdeI* and *HindIII* (Invitrogen™, Carlsbad, CA, USA), the insert was subcloned into the pET23a(+) expression vector (Novagen, Gibbstown, USA) pré-digested with the same enzymes and propagated in the same *E. coli* strain. The identity and lack of mutations of the insert of the resulting plasmid (pET23a(+):Spi-CEts) was confirmed by DNA automatic sequencing (MegaBACE 1000 – GE Healthcare, Uppsala, Sweden) with the ET Terminators cycle sequencing kit.

Expression of the recombinant protein

E.coli BL21 (DE3) (Novagen, Gibbstown, USA) strain was transformed with the recombinant pET23a(+):hSpi-CEts plasmid and propagated in agar plates containing Ampicilin 50µg/mL. 11 litres of LB medium containing Ampicilin 50µg/mL were inoculated with previous culture, originated from a single colony, until optical density (OD₆₀₀) reached 0,8. The expression of the target protein was then induced with 0.5mM IPTG (isopropyl β-D-thiogalactopyranoside) and grown for 3 additional hours, harvested and stored at -20°C.

Purification

Five grams of the frozen cells were resuspended in 25mL of buffer A (50mM potassium phosphate buffer pH 7.0) with addition of protease inhibitor cocktail tablet Complete[®] (Roche, Mannheim, Germany) and lysed by sonication 20 times (60% amplitude, 10 seconds each). The lysate was clarified by centrifugation for 30minutes at 48,000xg and treated with Deoxyribonuclease I from bovine pancreas (Sigma-Aldrich[®], St. Louis, EUA) and MgCl₂ 5mM for additional 30 minutes and then centrifuged again. The supernatant was dialyzed three times against 2L of buffer A and the soluble fraction after centrifuging was loaded onto a pre-rinsed strong cation exchange SP Sepharose FF 26/11 column (GE Healthcare, Uppsala, Sweden). The unbound material was washed with 6 column volumes of buffer A and an isocratic salt gradient with NaCl containing buffer A (buffer B), from 0,5M to 1M NaCl in 20 column volumes, was developed to separate the target protein from most contaminants. The protein was eluted between 0.75 – 0.80 M NaCl. The same procedure was repeated twice and all the eluted protein was gathered and concentrated to 5mL using a cut-off 3,000Da membrane (Millipore[™], Billerica, USA). The concentrated sample was loaded onto a size exclusion HiPrep 26/60 Sephacryl S-100 HR column (GE Healthcare, Uppsala, Sweden), which was previously equilibrated with buffer A. The chromatography was carried out with buffer A for elution and at a 0.3mL/min flow rate. The protein was eluted in less than one column volume and Spi-C Ets domain containing fractions were pooled. All procedures were carried out at 4°C and the FPLC procedure was monitored by UV detection at 215nm. The protein content of samples was analyzed by SDS-PAGE 15% and stained with coomasie brilliant blue.

The homogeneous protein was concentrated to 2.2mL and dialyzed against HBS buffer (10mM Hepes, 150mM NaCl, 3mM EDTA and 0.005% surfactant p20) pH 7.4 (three times of 0.5L) and stored at -80°C. Total protein concentration was determined by Bradford's method [19] using bovine serum albumin as standard (Bio-Rad Laboratories, Hercules, USA).

N-terminal amino acid sequencing

The N-terminal amino acid residues of homogeneous recombinant hSpi-C Ets domain were identified by automated Edman degradation sequencing using a PPSQ 21A gas-phase sequencer (Shimadzu, Kyoto, Japan).

Mass spectrometry analysis

Recombinant hSpi-C Ets domain was analyzed using electron-spray ionization mass spectrometry (ESI-MS) in an adaptation of Chassigne and Lobinski system [20]. Samples were analyzed on a triple quadrupole mass spectrometer (model QUATTRO II) equipped with a standard

ESI probe (Micromass, Altrincham, United Kingdom) and adjusted to a flow rate of ca. 250 $\mu\text{l min}^{-1}$. The source temperature (80°C) and needle voltage (3.6 kV) were maintained constant throughout experimental data collection, applying a drying gas (nitrogen) flow of 200 liters h^{-1} and a nebulizer gas flow of 20 liters h^{-1} . The mass spectrometer was calibrated with intact horse heart myoglobin and its typical cone voltage-induced fragments. The ESI-MS molecular mass of the hSpi-C Ets domain was determined by adjusting the mass spectrometer to give a peak with a half-height of 1 mass unit, and the cone sample to skimmer lens voltage controlling the transfer of ions to the mass analyzer was set to 38 V. 50 pmol sample was injected into electrospray transport solvent. The ESI spectrum was obtained in the multichannel acquisition mode, with scanning from 500 to 1,800 m/z at a scan time of 7 s. The mass spectrometer is equipped with MassLynx and Transform software for data acquisition and spectrum handling.

DNA sequences, annealing and preparation

We designed complementary oligonucleotides containing 36 nucleotides each and annealed them to generate double stranded DNA. The E74 was designed with and without a biotin molecule at 5' terminus on one of the strands. The human target DNA sequences (except from E74 from *Drosophila sp.*) and their respective mutants are given below. The core sequences are in italics:

Drosophila E74 promoter [21]: 5'- AGCTTCTCTAGCTGAATAACCGGAAGTAACTCATCG -3'
IL-7 receptor α promoter [15]: 5'- AGAAGGCAAGTTCAGAAACAGGAAGTCTGAGTAAA -3'
IL-7 receptor α promoter mut: 5'- AGAAGGCAAGTTCAGAAACACGACGTCTGAGTAAA -3'
Iga *MB1* promoter region 1 [22]: 5'- GGGGGGCTGGGCAGGAAGTAGGGTGGGGTTTGGGGT -3'
Iga *MB1* promoter region 1 mut: 5'- GGGGGGCTGGGCATCAAGTAGGGTGGGGTTTGGGGT -3'
Iga *MB1* promoter region 2: 5'- GAGGGCCACTGGAGCCCATCTCCGGGGGGCTGGGCA -3'
Iga *MB1* promoter region 2 mut: 5'- GAGGGCCACTGGAGCCCATCTTTGGGGGGCTGGGCA -3'
IgE germline promoter [11]: 5'- TCCCAAGAACAGAGAGAAAAGGGAAGTTCAGGGCG -3'
IgE germline promoter mut: 5'- TCCCAAGAACAGAGAGAAAAGTTAACTTCAGGGCG -3'

The oligonucleotides were resuspended in MilliQ water and the annealing was performed by adding equal amounts of each complementary oligonucleotide and submitted to 80°C with decreasing of 1 degree per minute until reached 15°C. The 5'biotin sequence was annealed at the same manner as the others except for the fact that the addition of the complementary strand was ten times the quantity of the 5'biotin-oligonucleotide. The concentration of the different annealed double-stranded DNA (36bp) was estimated by a fluorimetric assay (Quant-it™ dsDNA BR Assay Kit, QuBit Platform, Invitrogen, Carlsbad, USA). Briefly, 1 μl of double stranded synthesized DNA reproducing the promoter region of interest was diluted (200 X) in a reaction buffer containing the fluorescent dye (SybrGreen). The mixture was then incubated at room temperature for 15 minutes. The sample concentration was further measured in a fluorometer and compared with a calibrated curve generated by the serial dilution of a stock protein of known concentration. The assay has a broad range, enabling detection of DNA ranging from 100 $\text{pg} \cdot \mu\text{l}^{-1}$ to 1,000 $\text{ng} \cdot \mu\text{l}^{-1}$.

Surface plasmon resonance analysis (SPR)

The preparation of the sensor surfaces and the interpretation of the sensorgrams were performed using the BiacoreX system (Pharmacia biosensor, Uppsala, Sweden). For the

immobilization procedure, 5'-biotin-E74 DNA sequence was diluted to 30ng/mL in HBS buffer pH 7.4 (10mM Hepes, 150mM NaCl, 3mM, 0.005%, surfactant P20 – Biacore AB, Uppsala, Sweden) and 50µL were applied to a streptavidin pre-activated surface (Sensor Chip SA, Biacore AB, Uppsala, Sweden) at a 50µL/min flow rate. The 5'-biotin-E74 DNA was immobilized in one flow cell (FC2) on the sensorchip, while another remained as reference cell. The binding activity of the Spi-C Ets domain was assayed by injecting the purified protein over both flow cells. The protein was injected (50 µL) at different concentrations. The running buffer for immobilization and binding assays was the HBS-EP (Biacore AB, Uppsala, Sweden) and all the experiments were conducted at 25°C. The binding activity was measured as Resonance Units (RU).s⁻¹. The SPR signal obtained in each individual reaction cycle was recorded as a sensorgram, which is a real-time pattern plotted in RU versus time (s). After each experiment, injection of 1M, 2M NaCl and 0.1M potassium phosphate pH 2.0 were used for regeneration of the surface chip. Data are expressed in relative RU. Analysis of the results is based on relative binding of the Spi-C Ets domain to the E74 immobilized sequence. The efficiency of the regeneration procedure was examined as the baseline value returned to the initial value before the experiment.

Electrophoretic Mobility Shift Assay (EMSA)

In the reaction mixtures, the amount and concentration of DNA was kept constant while the protein concentration was varied. The binding reactions contained hSpi-C Ets domain at the indicated concentrations, 500ng of double-stranded target DNA, 6% glycerol, 2 µL of 5x sample solution (1x sample solution contains 25mg.mL⁻¹ polysucrose and 0.5mg.mL⁻¹ xylene cyanol) and 1x TBE (89mM Tris-borate pH 8.3, 2mM EDTA) enough to complete 10µL of final volume. Reaction mixtures were kept on room temperature (approximately 20°C) for 20 minutes and then loaded into a native 6% polyacrylamide gel pre-electrophoresed for 1h. Electrophoresis was continued for an additional 2hours and 10minutes. Gels were run in 1x TBE at 30V at room temperature and stained with staining kit Owl Silver Stain (Thermo Scientific Owl Separation Systems, Inc, Portsmouth, UK)

Results

hSpi-C Ets domain was obtained in a heterologous system

The sequence coding for Ets-binding domain from hSpi-C was constructed with the addition of restriction sites at the both terminus. The entire sequence contains 330pb. It was cloned in pCR[®]-Blunt vector and subcloned in pET23a(+) expression vector. To confirm the insertion of the constructed sequence the recombinant vector was *Nde*I and *Hind*III digested, shown in Figure 1.

BL21(DE3) *E. coli* strain, transformed with the recombinant expression vector, was grown in 11L of LB until 3 hours after IPTG induction to obtain 15g of harvested cells.

Purification and analysis of the recombinant hSpi-C Ets domain

In order to obtain a homogeneous protein we developed a two column strategy, which yielded essentially pure Spi-C Ets domain (Figure 2). The purification protocol reaches approximately 0.7mg of recombinant protein per gram of cells. The total solution volume containing the homogeneous

recombinant protein after dialysis against HBS buffer pH 7.4 was approximately 2.2mL which was frozen in liquid nitrogen for storage.

Analysis of purified hSpi-C Ets domain

The ESI-MS of homogeneous Ets-binding domain from hSpi-C revealed a subunit molecular mass of 12,500.28 Da, consistent with the theoretical molecular mass (12,500.28 Da). The first 8 N-terminal amino acid residues were identified as KGRKKLRL, which confirmed the identity of the protein.

Surface plasmon resonance analysis

The SPR assay was performed by immobilizing biotinilated DNA into a streptavidin surface on the FC2 flow-cell in the Sensor Chip SA followed by injections with different concentrations of hSpi-C Ets domain on both FC1 (reference cell) and FC2 (cell with immobilized DNA). The immobilization response for 5'-biotin-E74 promoter sequence was approximately 105 RU. We observed that at every protein injection the sensorgrams registered an association and dissociation phase on both FC1 and FC2 flow-cells. Then some experiments injecting high salt concentrations or low pH buffer (pH lower than carboxyl group isoelectric point) were performed to confirm the protein affinity not only for the DNA but also for the carboximethylated dextran on chip surface. The chip surface did not regenerate with high concentrations of NaCl injections. We were able to observe some surface regeneration only when pH2.0 phosphate buffer was injected on FC1 flow-cell (pH lower than carboxyl group isoelectric point).

These interactions made prevented the acquisition of binding constants from the SPR system. Even though, the response on FC2 was higher than on FC1 confirming that there was more protein binding on the cell that contains immobilized E74 DNA than on the reference one. Thus, the protein recognizes and binds to E74 sequence.

Electrophoretic mobility shift assay

Since SPR assay was unfeasible with Spi-C Ets-binding domain, to characterize DNA-protein complexes, gel retardation experiments were done to confirm the protein binding ability to DNA target. Some specific PU.1 sequences were tested to the Ets-binding domain of Spi-C transcription factor.

The wild type and the Ets core mutated sequences were designed and used to perform EMSA assay to determine whether the Spi-C Ets domain has affinity and its binding specificity for them. The purpose of the mutations at the core sequence was to investigate how these modifications affect the binding of Spi-C Ets-binding domain.

Three double stranded 36 bp sequences from human IL-7 receptor α promoter, I γ *MB1* promoter region 1 and I γ *MB1* promoter region 2 and their respective mutants were constructed to function as DNA probe for the Spi-C Ets-binding domain. For positive control we constructed the canonical *Drosophila* E74 promoter sequence (see materials and methods). hSpi-C Ets domain resulted to recognized and bound all these sequences, even the mutated ones. The band shift profile was practically identical with all the probe sequences resulting in a significant delay in electrophoretic mobility of hSpi-C-DNA complexes. In addition, the higher the concentration of Spi-C, the more

retarded is the complex mobility when compared to free DNA migration. Thus the migration of DNA-protein complexes was retarded gradually with increasing concentrations of Spi-C.

To investigate if this anomalous recognition was due to the missing N and C-terminal amino acids we decided to test the human IgE germline promoter sequence which was previously demonstrated to be bonded by full Spi-C protein, and its mutant which was not recognized by the protein. We observed that the hSpi-C Ets-domain binds to this sequence with the same band pattern than the previous ones.

The band shifts generated with E74 DNA are shown in Figure 3. The Spi-C Ets-binding domain gives clearly four major band shifts: I, II, III and IV (lanes 2-8). The modification at the sequences (core) have no large effects on the band pattern, as the retarded complexes generated with these mutants were similar to those observed in the wild type sequences. The little differences in migration between the protein-DNA complexes formed with E74 sequence, IgE germline promoter, IL-7R α promoter, Ig α *MB1* promoter region 1 and Ig α *MB1* promoter region 2 can be due to different number of Spi-C Ets-binding domain proteins bound per sequence or even to different conformations of the DNA target caused by the protein.

Discussion

One of the most important cell groups – the lymphocytes – is responsible for the immunity defense. There are two major types of lymphocytes: B lymphocytes or B cells, which when activated differentiate into plasma cells that secrete antibodies; and T lymphocytes or T cells.

During the humoral immune response, antibodies produced by B cells cause the destruction of extracellular microorganisms and prevent the spread of intracellular infections. During B cell development in the bone marrow membrane immunoglobulins (mIg) are expressed on the surface of B lymphocytes from a very early developmental stage (pre B cell) when it is restricted to the IgM isotype. As soon as the B cell leaves the bone marrow it starts to express a second isotype on the surface, IgD. The mIg associates with other transmembrane proteins to form the B-cell antigen receptor complex (BCR). Signaling through the B cell antigen receptor (BCR) is required for peripheral B lymphocyte maturation, maintenance, activation and silencing [23].

The most important proteins concerned in cell signaling are the transcription factors which are involved in gene regulation binding to DNA sequences as enhancers and/or promoters of target genes, or interacting with other regulatory proteins. There is a family of transcription factors that is widely studied, the Ets family, which presents the Ets DNA binding domain. This family has some members that are known to act as regulators of specific genes for immune cells, the Spi subfamily integrants.

High levels of messenger mRNA of Spi-C Ets transcription factor has been detected in lymphoid organs, spleen and bone marrow, more specifically in B cell lineage. Spi-C seems to promote B-cell differentiation in pre-B and mature B lymphocytes whereas low level of expression is present in pro-B and secreting plasma B cells [8,9,11,12]. There is low level of Spi-C expression in macrophage cell lines as well [8], but its function in this cell type remains to be defined.

Spi-C was found to be implicated in the regulation of B-cell specific genes in mouse, such as the κ Y element from the SP6 κ promoter [8], Ig λ light chain λ_{2-4} enhancer element [10], IgH intronic enhancer and Fc γ R11b gene promoter [12]. Whereas the human Spi-C was found to interact with the signal transducer and activator of transcription 6 (STAT6), and both can synergistically transactivate IgE germline promoter [11].

There are some studies showing that Spi-C binds some of the sequences that its subfamily co-member PU.1 does [8-12] suggesting a modulatory relationship Spi-C/PU.1 activity to promote B cell differentiation [12].

In the present study, we choose DNA targets related with B-cell differentiation since a little is known about hSpi-C transcription factor. The sequences tested to be target of Spi-C Ets-binding domain play different functions in B cell development as well as are recognized by human PU.1 Ets factor. Accordingly to Shore et al. [24], the isolated Ets DNA-binding domains of Elk-1 and SAP-1 exhibit similar specificities of binding to their respective full proteins, because the Ets domains contains all the information required to direct sequence-specific DNA binding. However, little is known about the contribution made by this so-called Ets-binding domain. We addressed this question by making use of the Spi-C Ets domain to perform the binding experiments.

Attempting to obtain binding constants from Spi-C Ets domain-DNA affinity we performed SPR assays using the protein as the analyte and a double stranded oligonucleotide as the ligand, the E74 promoter – universal Ets binding sequence. For competition assays we designed oligonucleotides based in promoters and enhancers related to B cell development genes and were annealed to obtain double stranded sequences. As positive control, we chose the E74 promoter [11,21]. The other sequences that were selected had already been tested for other Ets proteins like PU.1 from the same subfamily of Spi-C and as negative control, mutation on the core sequence was done as described elsewhere [15,22]. Unfortunately, the DNA binding domain of Spi-C showed to have high affinity for the chip surface beyond than for the immobilized DNA. In this manner, it was not possible to acquire binding constants using the SPR technology.

As an alternative binding assay, we applied a conventional biochemical assay: the electrophoretic mobility shift assay (EMSA). Native polyacrylamide gels were used to perform these assays with the aim to obtain high definition bands of binding complexes. As a result, four specific retarded DNA-protein complexes were detected at the Ets-binding domain of Spi-C for each sequence. To determine which nucleotides are required for protein binding at the Ets DNA-binding site, the sequences carrying two mutations at the core sequence 5' GGAA 3' of the DNA tested were employed, however, no modification was presented at band shift pattern. We demonstrated that the Spi-C Ets-binding domain recognized and bound these sequences. Surprisingly, the EMSA failed to detect significant differences in binding to wild type and mutated Ets-binding DNA sequences.

A possible explanation for the Ets-binding domain to be not sufficient to discriminate between them is that many ETS-domain transcription factors are subjected to autoregulation, whereby their DNA-binding activity is usually masked until an appropriate trigger is in place. Another Ets transcription factor, ETS-1, which is not expressed by B cells, presents regions that are located both amino and carboxy-terminally to the ETS domain forming an inhibitory module, which blocks DNA binding by Ets-1 [25,26]. This autoinhibition can be relieved by triggering a conformational change and the unfolding of one of the α -helices [25,27]. Thus, a key difference between our own and the previous

study is, the Spi-C Ets binding domain by itself that could not have the possible autoinhibition function from the N- and C-terminal amino acids of the protein. Due to relatively low stringency of binding, many Ets-domain transcription factors possess autoinhibitory domains, including Elk [28], PU.1 [29] and members of PEA3 subfamily [30,31,32].

Apparently, as a DNA binding unit that serves to bring other functional components to their DNA target, the Ets domains, as well as other DNA binding domains, are expected to contain additional structural features that allow them to interact with other regulatory protein modules. Furthermore, regulatory proteins may directly alter the conformation of certain residues that interact with DNA, thereby further modifying the DNA binding specificity [33]. In addition, interactions with the paired DNA-binding domain of the transcription factor Pax-5 are known to promote DNA binding by Ets-1 and a subset of other ETS-domain proteins to a 'weak' Ets site in the *mb-1* promoter [34].

It has been demonstrated that full hSpi-C interacts with STAT6 for binding to the IgE germline promoter *in vitro* and do not bind with the mutated sequence [11]. IgE transcription occurs in developing B cells in response to IL-4 in the germinal centers in a secondary lymphoid organ [35]. We decided to include the IgE germline promoter (I ϵ) sequence and its mutated one in our experiments to test if DNA recognition by hSpi-C Ets domain had a different profile from the others we tested. Taking this into account, they would function as a reference for the difference in recognition between wild type and mutated sequence. Given that the band shifts profile by EMSA of these reference sequences was similar we can suggest that certainly other parts of Spi-C or the interaction with other transcription factors are important on the specificity and differentiation of target DNA. In addition, the latter DNA sequences for EMSA assay corroborate the hypothesis about the importance of the interaction of Ets domain transcription factors with other proteins for the precise recognition of the target DNA sequence for binding. The mechanisms by which individual members of transcription factor families achieve sequence-specific DNA binding appear to be increasingly diverse.

Acknowledgements

Financial support for this work was provided by Hewlett-Packard (Palo Alto, USA). We would like to thank Prof. Mario Sérgio Palma for the N-terminal sequencing and mass spectrometry analysis at the Laboratório de Zooquímica, Instituto de Insetos Sociais da Universidade do Estado de São Paulo, Rio Claro, SP, Brazil.

References

- 1.Sementchenko VI and Watson DK (2000) Ets target genes: past, present and future. *Oncogene* 19:6533-6548
- 2.Donaldson LW, Petersen JM, Graves BJ et al (1994) Secondary structure of the ETS domain places murine Ets-1 in the superfamily of winged helix-turn-helix DNA-binding proteins. *Biochemistry* 33:13509-13516

3. Gunther CV, Nye JA, Bryner RS et al (1990) Sequence-specific DNA binding of the proto-oncoprotein *ets-1* defines a transcriptional activator sequence within the long terminal repeat of the Moloney murine sarcoma virus. *Genes Dev* 4:667-679
4. Karim FD, Urness LD, Thummel CS et al (1990) The ETS-domain: a new DNA-binding motif that recognizes a purine-rich core DNA sequence. *Genes Dev* 4:1451-1453
5. Bosselut R, Duvall JF, Gégonne A et al (1990) The product of the *c-ets-1* proto-oncogene and the related *Ets2* protein act as transcriptional activators of the long terminal repeat of human T cell leukemia virus HTLV-1. *EMBO J* 9:3137-3144
6. Pio F, Kodandapani R, Ni C-Z et al (1996) New insights on DNA recognition by *ets* proteins from the crystal structure of the PU.1 ETS domain-DNA complex. *J Biol Chem* 271:23329-23337
7. Carlsson R, Hjalmarsson A, Liberg D et al (2002) Genomic structure of mouse SPI-C and genomic structure and expression pattern of human SPI-C. *Gene* 299:271-278
8. Bemark M, Mårtensson A, Liberg D et al (1999) Spi-C, a novel Ets protein that is temporally regulated during B lymphocyte development. *J Biol Chem* 274:10259-10267
9. Hashimoto S, Nishizumi H, Hayashi R et al (1999) Prf, a novel Ets family protein that binds to the PU.1 binding motif, is specifically expressed in restricted stages of B cell development. *Int Immunol* 11:1423-1429
10. Carlsson R, Persson C, Leanderson T (2003) SPI-C, a PU-box binding ETS protein expressed temporarily during B-cell development and in macrophages, contains an acidic transactivation domain located to the N-terminus. *Mol Immunol* 39:1035-1043
11. Carlsson R, Thorell K, Liberg D et al (2006) SPI-C and STAT6 can cooperate to stimulate IgE germline transcription. *Biochem Biophys Res Commun* 344:1155-1160
12. Schweitzer BL, Huang KJ, Kamath MB et al (2006) Spi-C has opposing effects to PU.1 on gene expression in progenitor B cells. *J Immunol* 177:2195-2207
13. Pio F, Assa-Munt N, Yguerabide J et al (1999) Mutants of ETS domain PU.1 and GGAA/T recognition: Free energies and kinetics. *Protein Sci* 8:2098-2109
14. Teh HF, Peh WY, Su X et al (2007) Characterization of protein-DNA interactions using surface plasmon resonance spectroscopy with various assay schemes. *Biochemistry* 46:2127-2135

15. DeKoter RP, Schweitzer BL, Kamath MB et al (2007) Regulation of the Interleukin-7 Receptor α promoter by the Ets transcription factors PU.1 and GA-binding Protein in developing B cells. *J Biol Chem* 282:14194-14204
16. Matsuuchi L and Gold MR (2001) New views of BCR structure and organization. *Curr Opin Immunol* 13:270-277
17. Fleming HE and Paige CJ (2002) Cooperation between IL-7 and the pre-B cell receptor: a key to B cell selection. *Semin Immunol* 14:423-430
18. Santos DS, Chies JM, Basso LA, Renard G, Fonseca IO (2006) Method for the obtention of chimeric nucleotide sequences and chimeric nucleotide sequences. Patent application WO2007/068068 A2.
19. Bradford MM, Mcorrie RA, Williams WL (1976) A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248-254
20. Chassaigne H, Lobinski R (1998). Characterization of horse kidney metallothionein isoforms by electrospray MS and reversedphase HPLC-electrospray MS. *Analyst* 123:2125-2130
21. Kas K, Finger E, Grall F et al (2000) ESE-3, a novel member of an epithelium-specific Ets transcription factor subfamily, demonstrates different target gene specificity from ESE-1. *J Biol Chem* 275:2986-2998
22. Sigvardsson M, Clark DR, Fitzsimmons D et al (2002) Early B-Cell Factor, E2A, and Pax-5 cooperate to activate the early B Cell-Specific *mb-1* promoter. *Mol Cell Biol* 22:8539-8551
23. Geisberger R, Lamers M, Achatz G (2006) The riddle of the dual expression of IgM and IgD. *Immunology* 118:429-437
24. Shore P, Sharrocks AD (1995) The ETS-domain transcription factors Elk-1 and SAP-1 exhibit differential DNA binding specificities. *Nucleic Acids Res* 23:4698-4706
25. Jonsen MD, Petersen JM, Xu Q-P et al (1996) Characterization of the cooperative functions of inhibitory sequences in Ets-1. *Mol Cell Biol* 16:2065-2073
26. Skalicky JJ, Donaldson LW, Petersen JM, et al (1996) Structural coupling of the inhibitory regions flanking the ETS domain of murine Ets-1. *Protein Sci* 5:296-309
27. Petersen JM, Skalicky JJ, Donaldson LW et al (1995) Modulation of transcription factor Ets-1 DNA binding: DNA-induced unfolding of an alpha helix. *Science* 269:1866-1869

28. Janknecht R, Zinck R, Ernst WH et al (1994) Functional dissection of the transcription factor Elk-1. *Oncogene* 9:1273-1278
29. Brass AL, Kehrl E, Eisenbeis CF et al (1996) Pip, a lymphoid-restricted IRF, contains a regulatory domain that is important for autoinhibition and ternary complex formation with the Ets factor PU.1. *Genes Dev* 10:2335-2347
30. Xin JH, Cowie A, Lachance P et al (1992) Molecular cloning and characterization of PEA3, a new member of the Ets oncogene family that is differentially expressed in mouse embryonic cells. *Genes Dev* 6:481-496.
31. Brown LA, Amores A, Schilling TF et al (1998) Molecular characterization of the zebra fish PEA3 ETS-domain transcription factor. *Oncogene* 17:93-104.
32. Laget MP, Defosse PA, Albagli O et al (1996) Two functionally distinct domains responsible for transactivation by the Ets family member ERM. *Oncogene* 12:1325-1336
33. Garvie, CW, Hagman, J, Wolberger, C (2001) Structural studies of Ets-1/Pax5 complex formation on DNA. *Mol Cell* 8:1267-1276.
34. Fitzsimmons D, Hodsdon W, Wheat W et al (1996) Pax-5 (BSAP) recruits Ets proto-oncogene family proteins to form functional ternary complexes on a B-cell-specific promoter. *Genes Dev* 10:2198-2211
35. Oettgen HC (2000) Regulation of the IgE isotype switch: new insights on cytokine signals and the functions of epsilon germline transcripts. *Curr Opin Immunol* 12:618-623

Figures

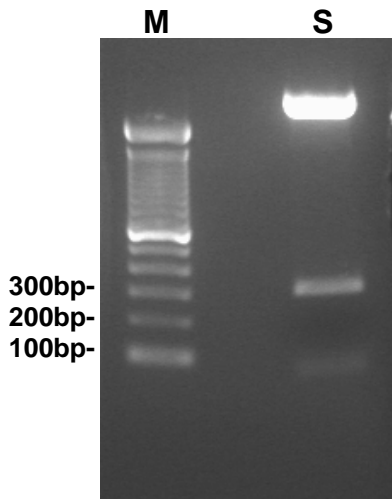


Fig 1 NdeI and HindIII cleavage of pET23a(+):hSpi-CEts. Lane M: 100bp DNA Ladder (Invitrogen); lane S: pET-23a+::Spi-CEts clones. The DNA fragment (330 bp) was confirmed into pET23a(+) vector (3666 bp) at the clones.

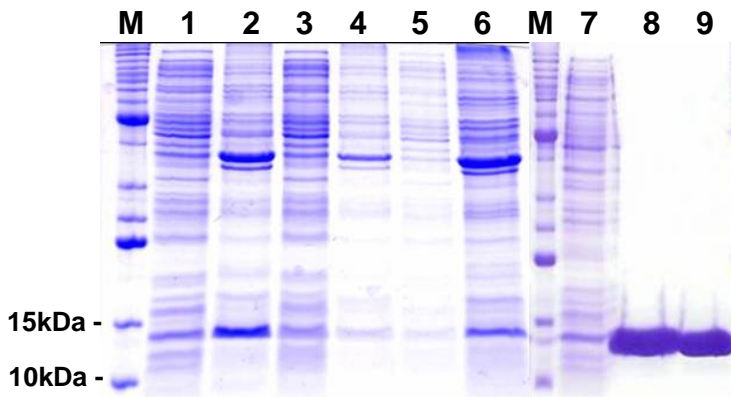


Fig 2 SDS-PAGE from steps of purification protocol. Lane M: Bench Mark Protein Ladder (Invitrogen); lane 1: sonication soluble fraction; lane 2: sonication insoluble fraction; lane 3: soluble fraction after DNase treatment; lane 4: insoluble fraction after DNase treatment; lane 5: soluble fraction after dialysis; lane 6: insoluble fraction after dialysis; lane 7: pré-SP Sepharose FF sample; lane 8: concentrated fractions eluted from SPFF; lane 9: pooled fractions from Sephacryl S100.

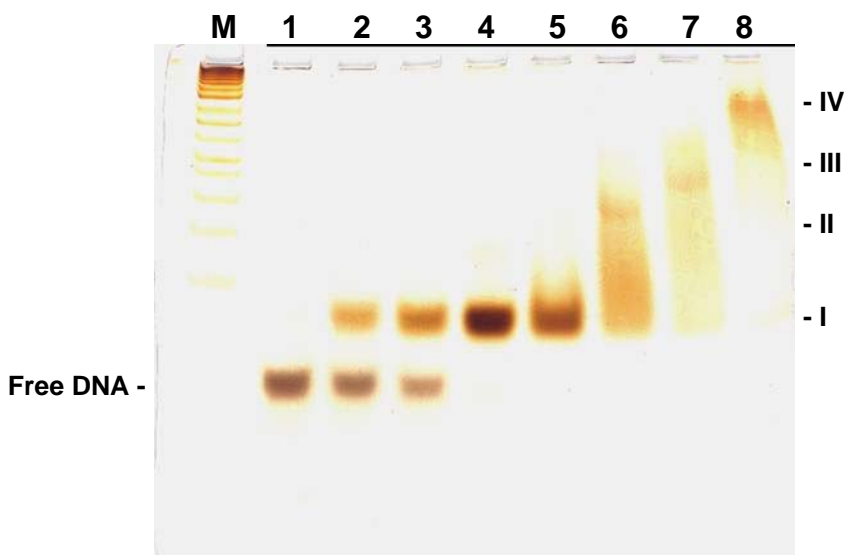


Fig 3 EMSA assay shift pattern. E74 promoter sequence. Lane M: 1Kb Plus DNA Ladder (Gibco BRL); lane 1: 500ng of free E74 (36bp); lane 2-8: 500ng E74 + Spi-C. lane2: 50ng Spi-C; lane 3: 100ng Spi-C; lane 4: 200ng Spi-C; lane 5: 400ng Spi-C; lane 6: 600ng Spi-C; lane 7: 800ng Spi-C; lane 8: 1000ng Spi-C. The different Spi-CEts-DNA complex shifts are identified as I, II, III, and IV from the smallest (less retarded) to the biggest (more retarded) respectively.

CONSIDERAÇÕES FINAIS

Uma das principais classes de células responsáveis pela imunidade é a dos linfócitos. Os linfócitos podem se apresentar como linfócitos T, que circulam no organismo e compreendem as células T citotóxicas (CD8) que destroem células infectadas por vírus e os linfócitos T auxiliares (CD4) que coordenam as respostas imunes. Os linfócitos B reconhecem as células e os materiais "estranhos" (por exemplo, bactérias que invadiram o organismo) e são responsáveis pela produção de anticorpos, que se "fixam" na superfície da célula estranha levando à sua destruição. Os anticorpos podem existir em diferentes formas conhecidas como isotipos ou classes. Nos mamíferos existem cinco isotipos diferentes de anticorpos, conhecidos como IgA, IgD, IgE, IgG e IgM³².

A sinalização por meio do complexo receptor de antígeno de célula B (BCR) é necessária para a maturação, manutenção, ativação e silenciamento de linfócitos B periféricos. Em células B maduras, o receptor de antígeno normalmente consiste de dois isotipos, IgM e IgD de membrana (mIgM, mIgD). Os sinais iniciados nos dois isotipos diferem na cinética e intensidade, mas o BCR de ambos os isotipos têm capacidade de compensar a perda um do outro, não estando clara a necessidade de expressar mIgD além de mIgM em células B maduras. Recentemente, Geisberger e colaboradores (2006) têm sugerido que IgD seja uma molécula-chave moduladora da resposta imune humoral. Além disso, sabe-se que vários defeitos nas moléculas envolvidas na sinalização do BCR podem impedir a transição de células B imaturas em maduras³⁸.

Uma classe de moléculas importantes na sinalização celular é a de fatores de transcrição. Essas moléculas estão envolvidas na regulação direta ou indireta de vários genes, podendo atuar ligando-se ao DNA, em "enhancers" ou promotores, ou

a outras proteínas reguladoras desses genes alvo. Uma família importante de fatores de transcrição é a família de proteínas Ets. Essa família tem como característica principal a presença de um domínio ligador de DNA, chamado domínio Ets, que por sua vez tem como característica principal ligar-se especificamente a seqüências com um “core” 5' GGAA/T 3'. Como essa família apresenta o domínio de ligação ao DNA muito conservado^{11,14}, mesmo entre suas subfamílias, é de extrema importância estudar os possíveis mecanismos envolvidos na ligação ao DNA, já que os diferentes membros apresentam diferenças nas suas especificidades e afinidades de ligação¹.

Como o domínio Ets é uma fração da proteína total que mantém suas características estruturais e funcionais¹⁶, tornou-se interessante estudar o tipo de seqüências reconhecidas por esse domínio, bem como estabelecer possíveis genes relacionados. A proteína Spi-C, pertencente à família de fatores de transcrição Ets, tem sido relacionada com a regulação do desenvolvimento de células B^{19,25,27,28,30}.

Com o objetivo de caracterizar possíveis alvos de ligação à Spi-C humana, foram desenhados oligonucleotídeos baseados na seqüência que codifica para o domínio Ets de ligação ao DNA de Spi-C humana. O gene sintético foi construído por sobreposição desses oligonucleotídeos, numa técnica de múltiplos passos de acordo com a patente desenvolvida pelo nosso laboratório WO2007/068068 A2³⁹. Após a síntese desse gene, o fragmento foi clonado no vetor de expressão pET23a(+), e foi realizado o seqüenciamento do mesmo para comprovar a identidade e descartar qualquer tipo de mutação ocorrida durante o processo de síntese. O plasmídeo recombinante pET23a(+):Spi-C domínio Ets foi transformado em uma série de cepas de *Escherichia coli* para a determinação do melhor perfil de expressão. Entre todas as cepas testadas a que melhor resultado apresentou foi BL21(DE3), onde a proteína se apresentou na fração solúvel numa quantidade relativa de 11 %,

tornando possível sua purificação. A purificação da proteína foi realizada por FPLC e foram testados diferentes tipos de colunas cromatográficas, como troca catiônica, interação hidrofóbica, exclusão por tamanho, entre outras. Após uma série de testes, o protocolo de purificação mais eficiente foi determinado utilizando-se as colunas de troca catiônica HiTrap SP Sepharose Fast Flow e a de exclusão por tamanho Sephacryl S-100. A identidade da proteína foi confirmada por seqüenciamento amino-terminal e espectrometria de massa.

Para os testes de ligação do domínio Ets de Spi-C ao DNA, decidiu-se utilizar a técnica de ressonância plasmônica de superfície (SPR), utilizando-se o aparelho Biacore X, que permite que se faça medidas das constantes de ligação (associação-dissociação) da proteína com o DNA. Para tal, foram desenhados oligonucleotídeos baseados nas seqüências de promotores de genes relacionados com o desenvolvimento de célula B. Esses oligonucleotídeos foram anelados numa reação 1:1 para a obtenção de uma DNA dupla fita. Como controle positivo, utilizou-se a seqüência promotora de E74, que é uma seqüência ligadora natural de todas as proteínas Ets, além da seqüência promotora de IgE já descrita como alvo de ligação para Spi-C. Como controles negativos, foram feitas mutações no “core” dos sítios de ligação, modificações estas descritas por outros autores^{28,40}. Como já foi dito anteriormente, as demais seqüências escolhidas já haviam sido testadas, onde proteínas da mesma subfamília que Spi-C, como PU.1, ligaram-se para a regulação da transcrição desses genes^{34,41}. Como não existem muitos dados sobre o papel da proteína Spi-C nas seqüências promotoras dos diferentes genes envolvidos no desenvolvimento de células B, e como esses genes possivelmente são regulados por muitos fatores de transcrição, foram escolhidos outros três promotores para os experimentos de ligação proteína-DNA.

Os resultados obtidos por SPR infelizmente não foram satisfatórios, pois a proteína além de ligar-se ao DNA, ligou-se também na superfície do “chip” que continha o DNA ligado. Essa característica da proteína impediu a análise dos dados, pois as curvas de ligação não poderiam ser utilizadas para obter as constantes de associação/dissociação da ligação proteína-DNA, pois a influência da ligação do domínio à dextrana do “chip” foi significativo uma vez que a regeneração total do “chip” não foi lograda.

Como uma alternativa, optou-se por um método bioquímico convencional: ensaios de retardo da migração do DNA em gel (EMSA). Esta técnica EMSA é largamente utilizada, gerando resultados confiáveis de ligação proteína-DNA por diferentes autores^{42,43,44,45}. Para a realização desses testes foram utilizados géis de poliacrilamida, com o intuito de se obter uma melhor definição desses complexos de ligação, já que os DNAs utilizados eram de 36pb. Como pôde ser visto, houve retardo da migração do DNA nos géis à medida que foi se aumentando a quantidade de proteína adicionada ao DNA. Porém, quando se compara o resultado da ligação da proteína aos DNAs mutados, onde não deveria ocorrer ligação do domínio Ets ao DNA, pode-se observar que houve o mesmo perfil de ligação de proteína-DNA, com formação de complexos similares aos dos DNAs tipo-selvagem, em todas as seqüências analisadas. O padrão de bandas em “escada” dos ensaios com EMSA sugere que a proteína está se ligando ao DNA numa proporção maior do que 1:1 (DNA:proteína), pois à medida que a concentração da proteína aumenta, vai ocorrendo um maior retardo na migração do DNA no gel com o aparecimento de outras bandas de maior tamanho. Este fenômeno pode ser explicado de duas maneiras. A primeira hipótese é que pode estar ocorrendo oligomerização de proteínas. A segunda é que com o aumento da concentração do domínio pode ocorrer o reconhecimento e ligação inespecífica de seqüências vizinhas à seqüência

consenso de ligação específica. Os resultados obtidos nesse trabalho revelam que apenas o domínio Ets de Spi-C não é suficiente para diferenciar seqüências alvo, sendo sim importante para sua ligação ao DNA mas evidenciando que os outros domínios ou partes da proteína são importantes para a especificidade e diferenciação do DNA alvo. Carlsson e colaboradores demonstraram que a proteína completa de Spi-C interage com outra proteína (STAT6) para se ligar à seqüência promotora de IgE²⁸, corroborando a hipótese de que a interação com outras proteínas é fundamental para o reconhecimento preciso da seqüência alvo.

Este trabalho tem como perspectivas utilizar novamente o sistema de ressonância plasmônica de superfície para testar outro tipo de ensaio onde o ligante será a proteína e os analitos serão diferentes seqüências alvo de DNA. Para isso as condições do experimentos deverão ser reformuladas, assim como o tipo de chip deverá ter propriedades para imobilizar a proteína.

Para obter dados estruturais do domínio Ets de Spi-C tanto na sua forma apo quanto na sua forma complexada ao DNA, protocolos para obtenção de cristais serão testados para posterior análise por cristalografia de difração de raios X.

Os resultados dos experimentos aqui descritos sugerem que sejam realizados experimentos com a proteína Spi-C completa. Para tal finalidade oligonucleotídeos serão desenhados para a construção da seqüência que codifica para a mesma. Para a expressão e purificação da proteína inteira, poderão ser testados os protocolos já estabelecidos nesse trabalho, os quais serão modificados e aperfeiçoados.

A comparação dos dados que serão obtidos utilizando ambas as formas da proteína é imprescindível para a complementação dos resultados apresentados neste trabalho, assim como para estudar a interação proteína-DNA dos fatores de transcrição Ets e mais especificamente da Spi-C. A partir desses resultados

poderemos também obter mais dados que sugiram novas informações sobre a função de Spi-C no desenvolvimento dos linfócitos B.

REFERÊNCIAS

1. Sharrocks AD. The ETS-domain transcription factor family. *Nat Rev Mol Cell Biol.* 2001 Nov;2(11):827-837.
2. Sementchenko VI, Watson DK. Ets target genes: past, present and future. *Oncogene.* 2000 Dec;19(55):6533-6548.
3. Oikawa T, Yamada T. Molecular biology of the Ets family of transcription factors. *Gene.* 2003 Jan;303:11-34.
4. Oikawa T. ETS transcription factors: possible targets for cancer therapy. *Cancer Sci.* 2004 Aug;95(8):626-633.
5. Seth A, Watson DK. ETS transcription factors and their emerging roles in human cancer. *Eur J Cancer.* 2005 Nov;41(16):2462-2478.
6. Leprince D, Gégonne J, Coll J, de Taisne C, Schneeberger A, Lagrou C, et al. A putative second cell-derived oncogene of the avian leukaemia retrovirus E26. *Nature.* 1983 Nov;306(5941):395-397.
7. Moscovici MG, Jurdic P, Samarut J, Gazzolo L, Mura CV, Moscovici C. Characterization of the hemopoietic target cells for the avian leukemia virus E26. *Virology.* 1983 Aug;129(1):65-78.
8. Nunn MF, Seeburg PH, Moscovici C, Duesberg PH. Tripartite structure of the avian erythroblastosis virus E26 transforming gene. *Nature.* 1983 Nov;306(5941):391-395.
9. Marchler-Bauer A, Anderson JB, Derbyshire MK, DeWeese-Scott C, Gonzales NR, Gwadz M et al. *CDD: a conserved domain database for interactive domain family analysis.* *Nucleic Acids Res.* 2007;35: D237-40.
10. Gunther CV, Nye JA, Bryner RS, Graves BJ. Sequence-specific DNA binding of the proto-oncoprotein ets-1 defines a transcriptional activator sequence within the long terminal repeat of the Moloney murine sarcoma virus. *Genes Dev.* 1990 Apr;4(4):667-679.
11. Karim FD, Urness LD, Thummel CS, Klemsz MJ, McKercher SR, Celada A, et al. The ETS-domain: a new DNA-binding motif that recognizes a purine-rich core DNA sequence. *Genes Dev.* 1990 Sep;4(9):1451-1453.
12. Degan MB, Degan MS, Naganuma T, Morse DE. The *ets* multigene family is conserved throughout the Metazoa. *Nucleic Acids Res.* 1993 Jul; 21(15):3479-3484.
13. Graves BJ, Petersen JM. Specificity within the ets family of transcription factors. *Adv Cancer Res.* 1998;75:1-55.
14. Bosselut R, Duvall JF, Gégonne A, Bailly M, Hémar A, Brady J et al. The product of the *c-ets-1* proto-oncogene and the related Ets2 protein act as transcriptional

activators of the long terminal repeat of human T cell leukemia virus HTLV-1. EMBO J. 1990 Oct;9(10):3137-3144.

15.Donaldson LW, Petersen JM, Graves BJ, McIntosh LP. Secondary structure of the ETS domain places murine Ets-1 in the superfamily of winged helix-turn-helix DNA-binding proteins. Biochemistry. 1994 Nov;33(46):13509-13516.

16.Branden C, Tooze J. Introduction to protein structure. New York: Garland Publishing, Inc; 1991.

17.Laudet V, Niel C, Duterque-Coquillaud M, Leprince D, Stehelin D. Evolution of the ets gene family. Biochem Biophys Res Commun. 1993 Jan;190(1):8-14.

18.Klemsz MJ, Maki RA. Activation of transcription by PU.1 requires both acidic and glutamine domains. Mol Cell Biol. 1996 Jan;16(1):390-397.

19.Carlsson R, Persson C, Leanderson T. SPI-C, a PU-box binding ETS protein expressed temporarily during B-cell development and in macrophages, contains an acidic transactivation domain located to the N-terminus. Mol Immunol. 2003 Jun;39(16):1035-1043.

20.Kim CA, Phillips ML, Kim W, Gingery M, Tran HH, Robinson MA et al. Polymerization of the SAM domain of TEL in leukemogenesis and transcriptional repression. EMBO J 2001 Aug;20(15):4173-4182.

21.Kodandapani R, Pio F, Ni CZ, Piccialli G, Klemsz M, McKercher S et al. A new pattern for helix-turn-helix recognition revealed by the PU.1 ETS-domain-DNA complex. Nature. 1996 Apr;380(6573):465-460. Erratum in: Nature. 1998 392(6676):630.

22.Listman JA, Wara-aswapati N, Race JE, Blystone LW, Walker-Kopp N, Yang Z et al. Conserved ETS domain arginines mediate DNA binding, nuclear localization, and a novel mode of bZIP interaction. J Biol Chem. 2005 Dec;280(50):41421-41428.

23.Shintani S, Terzic J, Sato A, Saraga-Babic M, O'hUigin C, Tichy H et al. Do lampreys have lymphocytes? The Spi evidence. Proc Natl Acad Sci USA. 2000 Jun;97(13):7417-7422. Comment in Proc Natl Acad Sci USA. 2000;97(13):6924-6926.

24.Anderson MK, Sun X, Miracle AL, Litman GW, Rothenberg EV. Evolution of hematopoiesis: Three members of the PU.1 transcription factor family in a cartilaginous fish, *Raja eleganteria*. Proc Natl Acad Sci U S A. 2001 Jan;98(2):553-558.

25.Bemark M, Mårtensson A, Liberg D, Leanderson T. Spi-C, a novel Ets protein that is temporally regulated during B lymphocyte development. J Biol Chem. 1999 Apr;274(15):10259-10267.

26.Carlsson R, Hjalmarsson A, Liberg D, Persson C, Leanderson T. Genomic structure of mouse SPI-C and genomic structure and expression pattern of human SPI-C. Gene. 2002 Oct;299(1-2):271-278.

27. Hashimoto S, Nishizumi H, Hayashi R, Tsuboi A, Nagawa F, Takemori T, et al. Prf, a novel Ets family protein that binds to the PU.1 binding motif, is specifically expressed in restricted stages of B cell development. *Int Immunol*. 1999 Sep;11(9):1423-1229.
28. Carlsson R, Thorell K, Liberg D, Leanderson T. SPI-C and STAT6 can cooperate to stimulate IgE germline transcription. *Biochem Biophys Res Commun*. 2006 Jun;344(4):1155-1160.
29. Oettgen HC. Regulation of the IgE isotype switch: new insights on cytokine signals and the functions of epsilon germline transcripts. *Curr Opin Immunol*. 2000 Dec;12(6):618-623.
30. Schweitzer BL, Huang KJ, Kamath MB, Emelyanov AV, Birshtein BK, DeKoter RP. Spi-C has opposing effects to PU.1 on gene expression in progenitor B cells. *J Immunol*. 2006 Aug;177(4):2195-2207.
31. Namen AE, Lupton S, Hjerrild K, Wignall J, Mochizuki DY, Schmierer A et al. Stimulation of B-cell progenitors by cloned murine interleukin-7. *Nature*. 1988 Jun;333(6173):571-573.
32. Janeway CA, Travers P, Walport M, Shlomchik M. *Imunobiologia: O sistema immune na saúde e na doença*. 5 ed. Porto Alegre: Artmed; 2002.
33. Kikuchi K, Lai AY, Hsu CL, Kondo M. IL-7 receptor signaling is necessary for stage transition in adult B cell development through up-regulation of EBF. *J Exp Med*. 2005 Apr;18;201(8):1197-203.
34. DeKoter R, Schweitzer BL, Kamath MB, Jones D, Tagoh H, Bonifer C, et al. Regulation of the Interleukin-7 receptor α promoter by the Ets transcription factors PU.1 and GA-Binding Protein in developing B cells. *J Biol Chem*. 2007 May;282(19):14194-204.
35. Fleming HE, Paige CJ. Cooperation between IL-7 and the pre-B cell receptor: a key to B cell selection. *Semin Immunol*. 2002 Dec;14(6):423-430.
36. Matsuuchi L, Gold MR. New views of BCR structure and organization. *Curr Opin Immunol*. 2001 Jun;13(3):270-277.
37. Hagman J, Grosschedl R. An inhibitory carboxyl-terminal domain in Ets-1 and Ets-2 mediates differential binding of ETS family factors to promoter sequences of the *mb-1* gene. *Proc Natl Acad Sci USA*. 1992 Oct;89(19):8889-8893.
38. Geisberger R, Lamers M, Achatz G. The riddle of the dual expression of IgM and IgD. *Immunology*. 2006 Aug;118(4):429-437.
39. Santos DS, Chies JM, Basso LA, Renard G, Fonseca IO. Method for the obtention of chimeric nucleotide sequences and chimeric nucleotide sequences. 2006. Patent application WO2007/068068 A2.

40. Kas K, Finger E, Grall F, Gu X, Akbarali Y, Boltax J et al. ESE-3, a novel member of an epithelium-specific Ets transcription factor subfamily, demonstrates different target gene specificity from ESE-1. *J Biol Chem.* 2000 Jan;275(4):2986-2998.
41. Sigvardsson M, Clark DR, Fitzsimmons D et al. Early B-Cell Factor, E2A, and Pax-5 cooperate to activate the early B Cell-Specific *mb-1* promoter. *Mol Cell Biol.* 2002 Dec;22(24):8539-8551.
42. Fisher RJ, Koizumi S, Kondoh A, Mariano JM, Mavrothalassitis G, Bhat NK, et al. Human ETS1 oncoprotein: purification, isoforms, -SH modification, and DNA sequence-specific binding. *J Biol Chem.* 1992 Sep;267(25):17957-17965.
43. Owen-Hughes TA, Pavitt GD, Santos DS, Sidebotham JM, Hulton CS, Hinton JC et al. The chromatin-associated protein H-NS interacts with curved DNA to influence DNA topology and gene expression. *Cell.* 1992 Oct;71(2):255-265.
44. Lipps G, Ibanez P, Stroessenreuther T, Hekimian K, Krauss G. The protein ORF80 from the thermophilic archaeon *Sulfolobus islandicus* binds highly site-specifically to double stranded DNA and represents a novel type of basic leucine zipper protein. 2001 Dec; 29(24):4973-4982.
45. Wieland G, Hemmerich P, Koch M, Stoyan T, Hegemann J, Diekmann S. Determination of the binding constants of the centromere protein Cbf1 to all 16 centromere DNAs of *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 2001;29(5):1054-1060.

ANEXO I

Testes de expressão em cepas de *Escherichia coli*

A expressão do domínio Ets recombinante da proteína Spi-C humana foi testada e comparada em diferentes cepas de *E. coli*: BL21(DE3), Rosetta-gammi(DE3), BL21(DE3)pLysS, C41(DE3)pLysS e C43(DE3)pLysS (Tabela I). Os meios LB (Luria Bertani) e TB (Terrific Broth) foram utilizados para fazer os testes.

As diferentes cepas de *E. coli* foram transformadas por eletroporação com o plasmídeo recombinante pET23a(+) contendo a seqüência de nucleotídeos que codifica para o domínio Ets da proteína Spi-C humana. O mesmo procedimento foi realizado com o vetor pET-23a(+), sem o gene de interesse clonado, para controle negativo de expressão.

A seleção das células contendo o plasmídeo recombinante foi realizado em meio LB sólido contendo os antibióticos específicos para cada cepa (Tabela II) a 37°C. Uma colônia foi inoculada em 3 mL de meio líquido com os respectivos antibióticos e cultivada sob agitação de 180 rpm com temperatura controlada de 37°C, por 16-18 horas. A partir desse pré-inóculo, fez-se um inóculo diluído 1:50 em LB líquido crescido nas mesmas condições. Quando o cultivo atingiu a fase logarítmica (densidade óptica a 600nm de 0,4), foi adicionada quantidade suficiente de isopropil β -D-tiogalactopiranosídeo (IPTG) para concentração de 1 mM a um dos cultivos, sendo que outro ficou sem indução. Após a indução, 1 mL do cultivo foi coletado periodicamente a cada 3 horas, até completar 24 horas. As células coletadas foram centrifugadas a 14000 rpm (correspondentes a 20800 x g) por 1 min e conservadas a -20°C.

Tabela I – Cepas de *Escherichia coli* utilizadas nos testes de expressão

Cepa	Resistência a antibióticos	Derivação	Características
BL21(DE3)	nenhuma	B834	Deficiência de protease <i>lon</i> e <i>ompT</i> , que reduz a degradação das proteínas recombinantes expressas
BL21(DE3) pLysS	Cloranfenicol	B834	Deficiência de protease <i>lon</i> e <i>ompT</i> , que reduz a degradação das proteínas recombinantes expressas. Expressa a lisozima T7 (inibidor da T7 RNA polimerase) para reprimir a expressão basal dos genes sob o controle do promotor T7 antes da indução.
C41(DE3) pLysS	Cloranfenicol	BL21(DE3)	Deficiência de protease <i>lon</i> e <i>ompT</i> , que reduz a degradação das proteínas recombinantes expressas. Expressa a lisozima T7 (inibidor da T7 RNA polimerase) para reprimir a expressão basal dos genes sob o controle do promotor T7 antes da indução.
C43(DE3) pLysS	Cloranfenicol	BL21(DE3)	Deficiência de protease <i>lon</i> e <i>ompT</i> , que reduz a degradação das proteínas recombinantes expressas. Expressa a lisozima T7 (inibidor da T7 RNA polimerase) para reprimir a expressão basal dos genes sob o controle do promotor T7 antes da indução.
Rosetta-gami (DE3)	Kanamicina + Tetraciclina + Cloranfenicol	Origami (K-12)	Expressão de tRNAs raros, facilitando a expressão de genes que codificam códons raros de <i>E. coli</i> ; facilita formação de pontes dissulfeto

Fonte: Protocolo de células competentes - Novagen® - www.novagen.com e manual de cepas competentes - OverExpress™ - www.overexpress.com.

Tabela II – Concentrações dos antibióticos utilizadas para crescimento das diferentes cepas de *E. coli* contendo o vetor de expressão recombinante.

Cepas de <i>E. coli</i> transformadas com o plasmídeo pET23a(+)*	Concentração de uso dos antibióticos
BL21(DE3)	50 µg/mL de Ampicilina
BL21(DE3) pLysS	50 µg/mL de Ampicilina, 34 µg/mL de Cloranfenicol
C41(DE3) pLysS	50 µg/mL de Ampicilina, 34 µg/mL de Cloranfenicol
C43(DE3) pLysS	50 µg/mL de Ampicilina, 34 µg/mL de Cloranfenicol
Rosetta-gami(DE3)	50 µg/mL de Ampicilina, 34 µg/mL de Cloranfenicol, 15 µg/mL de Kanamicina e 12,5 µg/mL de Tetraciclina

* O plasmídeo pET23a(+) contém o gene para resistência a Ampicilina.

Fonte: Protocolo de células competentes - Novagen® - www.novagen.com e manual de cepas competentes - OverExpress™ - www.overexpress.com.

Para análise da expressão das diferentes cepas de *E. coli* as células foram ressuspensas em tampão Tris-HCl 50mM pH 8,0 para posterior rompimento por sonicação (3 vezes por 10 segundos a 21% de amplitude). As frações solúvel (sobrenadante) e insolúvel (sedimento) de cada coleta foram separadas por centrifugação a 20800 x g por 30 min, a 4°C e analisadas por eletroforese em géis de poliacrilamida 15% (SDS-PAGE).

Houve expressão do domínio Ets da proteína Spi-C humana recombinante nas frações solúveis e induzidas das cepas BL21(DE3) e BL21(DE3)pLysS e nas frações insolúveis das cepas BL21(DE3), Rosetta-gammi(DE3), BL21(DE3)pLysS. A cepa BL21(DE3) mostrou melhor resultado nos testes de expressão, pois a proteína estava na fração solúvel correspondendo a 11 % do total das proteínas solúveis.

Levando em consideração o resultado dos testes de expressão, foram realizados testes com a finalidade de produzir a maior quantidade de células possível expressando a proteína de interesse na melhor proporção proteína-contaminantes e de baixar o custo nesta etapa.

Os testes foram feitos com uma variável de cada vez:

- concentração de IPTG para indução: as concentrações testadas foram de 0,5 mM, 1,0 mM, 2,5 mM e 5,0 mM;

- aeração: o volume de meio de cultura em cada erlenmeyer de 2L foi de 1000 mL e 500 mL ;
- momento de indução com IPTG: induções nos momentos de densidade óptica (OD_{600}) 0,4, 0,6 e 0,8.

Protocolo de produção de células que expressam o domínio Ets de Spi-C humana

Após os testes descritos acima, o protocolo para obtenção de células que expressam o domínio Ets de Spi-C humana foi estabelecido como descrito a seguir. Uma colônia de BL21(DE3) contendo o plasmídeo recombinante (previamente selecionada em meio sólido com os antibióticos necessários) foi inoculada em 3 mL de LB líquido e cultivada a 37°C, 180 rpm e por 6 horas. 1mL deste cultivo foi inoculado em 500 mL de LB líquido em frasco de 2 L e cultivado sob as mesmas condições de 12 a 14 horas. Quantidade suficiente deste pré-inóculo foi inoculada em vários frascos contendo 500 mL de LB líquido com antibiótico para que a OD_{600} resultante fosse de 0,08. Quando o cultivo atingiu a OD_{600} 0,8 adicionou-se IPTG para concentração de 0,5 mM e foi cultivando-se por mais 3 horas. As células foram centrifugadas a 11800 x g e a 4°C e armazenadas a -20°C. Este protocolo quando realizado com 11 frascos de cultivo, rendeu aproximadamente 15 g de células.

ANEXO II

Testes de purificação do domínio Ets de Spi-C humana recombinante

Os testes de purificação do domínio Ets de Spi-C recombinante foram realizados utilizando cromatografia líquida de alta eficiência (HPLC) através do equipamento ÄKTA (GE Healthcare) para controle de fluxo, pressão da coluna cromatográfica, pH, temperatura, detecção de vários comprimentos de onda, controle de gradiente, aplicação e coleta de amostras. As células foram preparadas conforme descrito no ANEXO I e ressuspensas na proporção de 5 mL para cada grama de células. A fração solúvel das células foi tratada com DNase (Sigma-Aldrich, St. Louis, USA) para clivagem do DNA presente na amostra e dialisada contra o tampão (três vezes de 2L).

Foram utilizadas diferentes colunas cromatográficas como de troca catiônica (HiTrap SP FF, HiTrap SP XL, HiTrap CM FF, Mono S, Resorce S), de exclusão por tamanho (Sephacryl S-100), de interação hidrofóbica (HiTrap Phenyl FF low sub, Hi Trap Phenyl FF high sub, HiTrap Butyl FF, HiTrap Butyl-S FF e HiTrap Octyl FF), de troca aniônica (Mono Q) e multimodal (HiTrap MMC). A cromatografia em coluna utilizando o princípio da afinidade também foi testada com a resina DNA-cellulose (Sigma) em outro sistema de cromatografia líquida em coluna, sem pressão e a 4°C.

A proteína de interesse interagiu com todas as colunas HiTrap de troca catiônica, sendo a HiTrap SP FF a que melhor separou-a das outras proteínas presentes na amostra. Nas interações hidrofóbicas não houve interação da proteína alvo com a resina, pois a mesma foi observada em todos os lavados e nunca no gradiente. Como segunda coluna, sempre foi utilizada a Sephacryl S-100 sendo útil não só para retirar algumas proteínas contaminantes da amostra, mas também para trocar o tampão fazendo possível a utilização de uma terceira coluna.

Na tentativa de eliminar alguns contaminantes visíveis quando o gel era corado com prata foram testadas a Mono S, a Mono Q e a Resorce S que apresentaram padrões parecidos da proteína-alvo eluída, sendo que não havia separação do contaminante quando analisadas da mesma forma. Na coluna HiTrap MMC, que apresenta propriedades de interação hidrofóbica e troca catiônica ao mesmo tempo, a proteína permaneceu fortemente ligada na coluna após o gradiente e não foi eluída nem mesmo com NaCl 2M, desligando-se da resina somente com

NaOH 1M. A resina DNA-celulose foi testada com e sem adição de BSA no tampão. O BSA serviria para bloquear possíveis espaços onde a celulose poderia estar exposta e impedir interações da proteína com a celulose. O resultado desse teste foi negativo, já que não houve interação da proteína alvo com a resina com nenhum dos tampões.

Contudo, apenas a coluna de troca catiônica HiTrap SP FF (Figura 1) e a de exclusão por tamanho Sephacryl S-100 (Figura 2) mostraram eficiência na eliminação dos contaminantes durante a eluição da proteína de interesse. O protocolo de purificação da proteína está descrito no manuscrito.

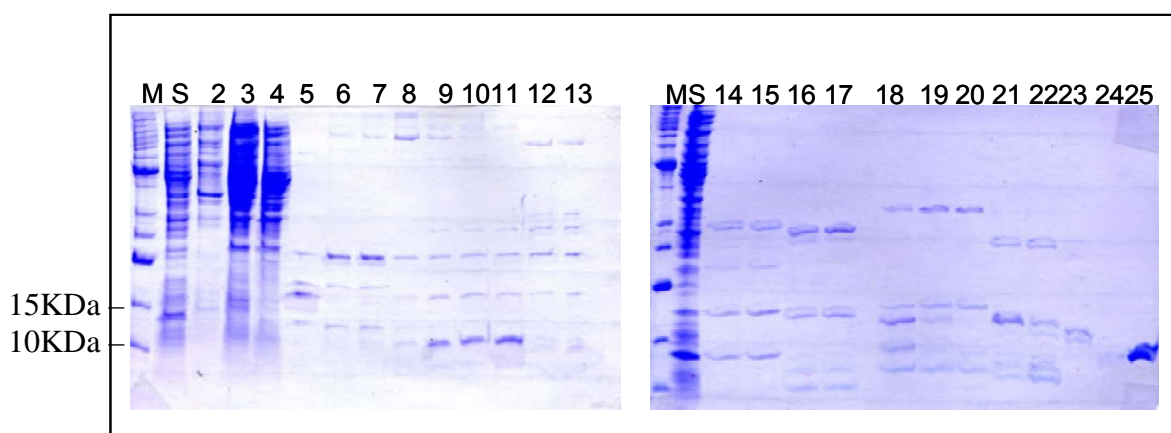


Figura 1. Eletroforese em gel de poliacrilamida das frações eluídas da coluna de troca catiônica SP Sepharose Fast Flow utilizando FPLC. Canaleta M: marcador Bench Mark Protein Ladder; canaleta S: amostra injetada na coluna; canaletas 2-25: frações eluídas da coluna. As canaletas 24 e 25 representam as frações que foram utilizadas para o próximo passo.

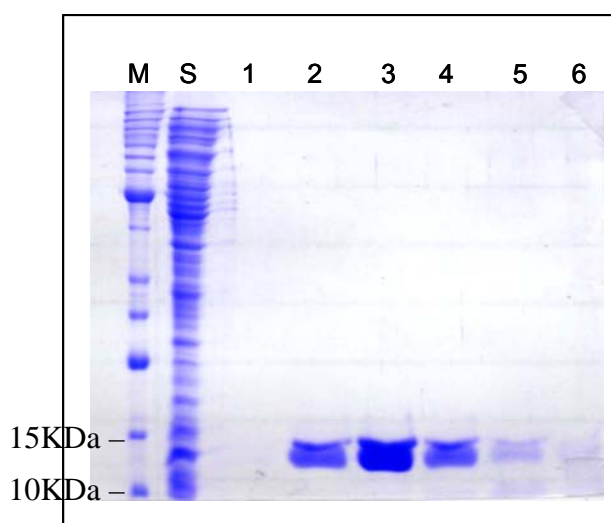


Figura 2. Eletroforese em gel de poliacrilamida das frações eluídas da coluna de exclusão por tamanho Sephacryl S100 utilizando FPLC. Canaleta M: marcador Bench Mark Protein Ladder; canaleta S: amostra injetada na SPFF; Canaletas 1-6: frações eluídas da coluna. As canaletas 2, 3, 4 e 5 representam as frações utilizadas para o próximo passo.

ANEXO III

Testes de ligação à DNA

Para obter constantes de associação e dissociação do domínio Ets de Spi-C com o DNA foram feitos ensaios de ligação utilizando ressonância plasmônica de superfície e espectrofluorimetria. Em ambos os casos foi utilizado o oligonucleotídeo de 36 pares de base referente à seqüência do promotor de E74 de *Drosophila sp.*

Ressonância plasmônica de superfície

A técnica de ressonância plasmônica de superfície baseia-se na detecção de mudança de massa na superfície de um chip quando, após a imobilização de um ligante, passa-se (injeta-se) um analito contido em tampão específico (HBS) com fluxo constante. Havendo ligação, a luz polarizada muda de ângulo e é detectada. Cada chip tem duas superfícies, uma é utilizada como célula de referência (flow-cell 1 – FC1) e a outra suporta a molécula imobilizada, onde deve ocorrer a associação e a dissociação entre o ligante e o analito durante o ensaio (flow-cell 2 – FC2). O sistema gera os resultados em forma de sensorgramas, ou seja, gráficos que mostram a alteração da resposta em Unidades de Ressonância (RU) versus o tempo. O sensorgrama que é utilizado para visualizar a ligação e para retirar as constantes de associação e de dissociação é obtido pela subtração do valor do sinal em FC2 pelo de FC1. Neste caso, o oligonucleotídeo foi a molécula escolhida para ser imobilizada no chip. Para isso foi encomendado um oligonucleotídeo igual ao descrito no artigo submetido, porém com uma molécula de biotina ligada na extremidade 5'. O chip utilizado contém uma camada de dextran carboximetilada e uma camada de estreptavidina ligada covalentemente ao dextran, na sua superfície, assim o oligonucleotídeo biotinilado liga-se com grande afinidade à estreptavidina imobilizando o DNA na superfície do chip.

Os métodos e resultados estão descritos no artigo submetido. Em suma, não foi possível a utilização do sistema porque a proteína em questão ligou-se não somente ao DNA imobilizado, mas também ao dextran presente na superfície do chip. Assim, houve ligação tanto na célula FC1 quanto na célula FC2, o que impossibilitou que os dados necessários para a análise da ligação fossem obtidos, já

que o sensorgrama FC2-FC1 não representava somente a ligação com o DNA. Além disso, mesmo após a injeção de altas concentrações de cloreto de sódio, utilizado para dissociar a proteína do DNA, a proteína permaneceu ligada ao dextran e com isso a resposta de ligação foi cada vez menor.

Para uma demonstração gráfica do que ocorreu nos ensaios descritos, seguem as Figuras 1, 2, 3 e 4 com breve detalhamento.

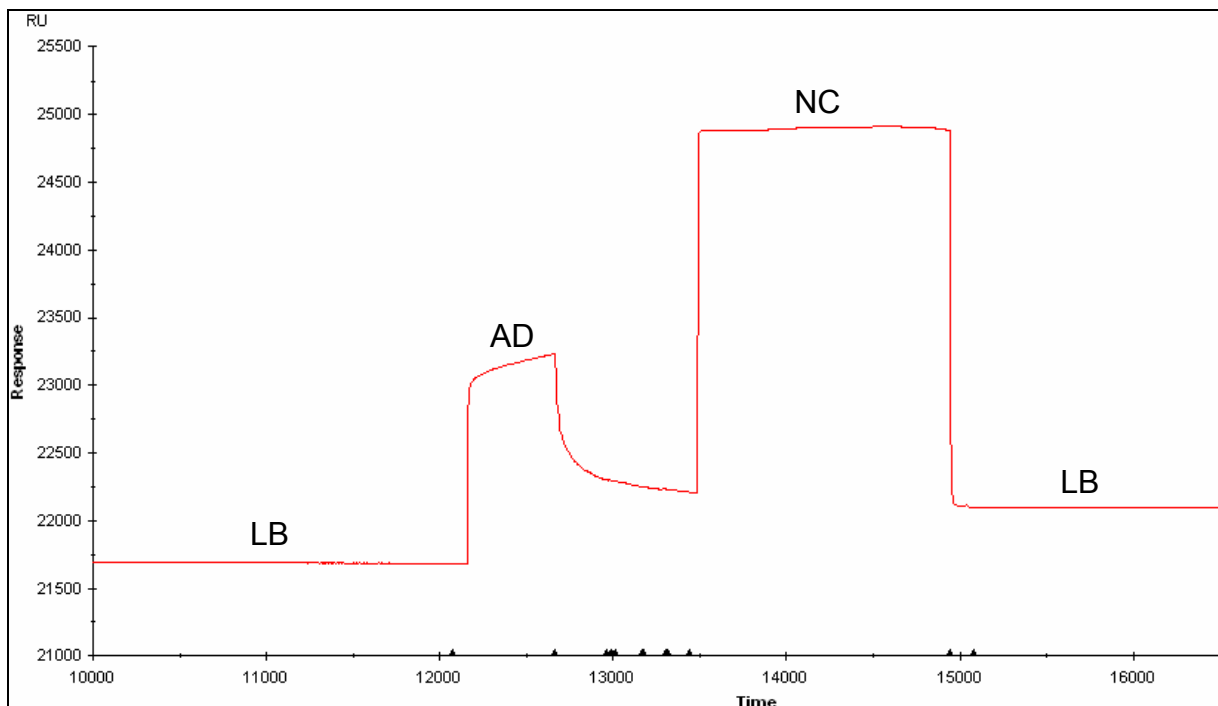


Figura 1. Sensorgrama da resposta em FC1 (célula sem DNA imobilizado) após injeção do domínio Ets de Spi-C. A primeira curva (AD) corresponde à associação e dissociação da proteína com o dextran. A linha de base (LB), que é a resposta sem o analito, não voltou ao valor original nem após a injeção de NaCl 2M (NC) significando que a superfície não foi totalmente regenerada e que a proteína permaneceu ligada mesmo após o NaCl.

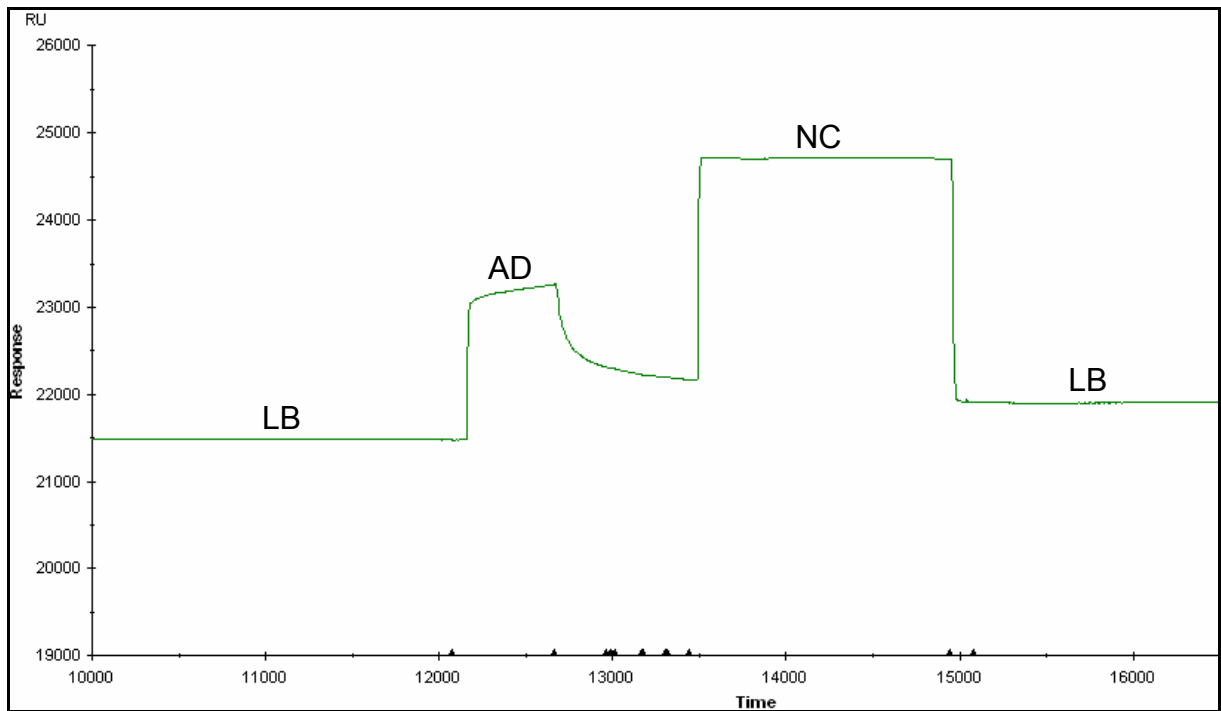


Figura 2. Sensorgrama da resposta em FC2 (célula com DNA imobilizado) após a mesma injeção do domínio Ets de Spi-C da Figura 1. A primeira curva (AD) corresponde à associação e dissociação da proteína com o DNA e com o dextran. A linha de base (LB) não voltou ao valor original nem após a injeção de NaCl 2M (NC). O comportamento foi muito similar ao de FC1, somente a diferença ou variação de resposta é que foi maior por haver ligação não só no dextran, mas também no DNA.

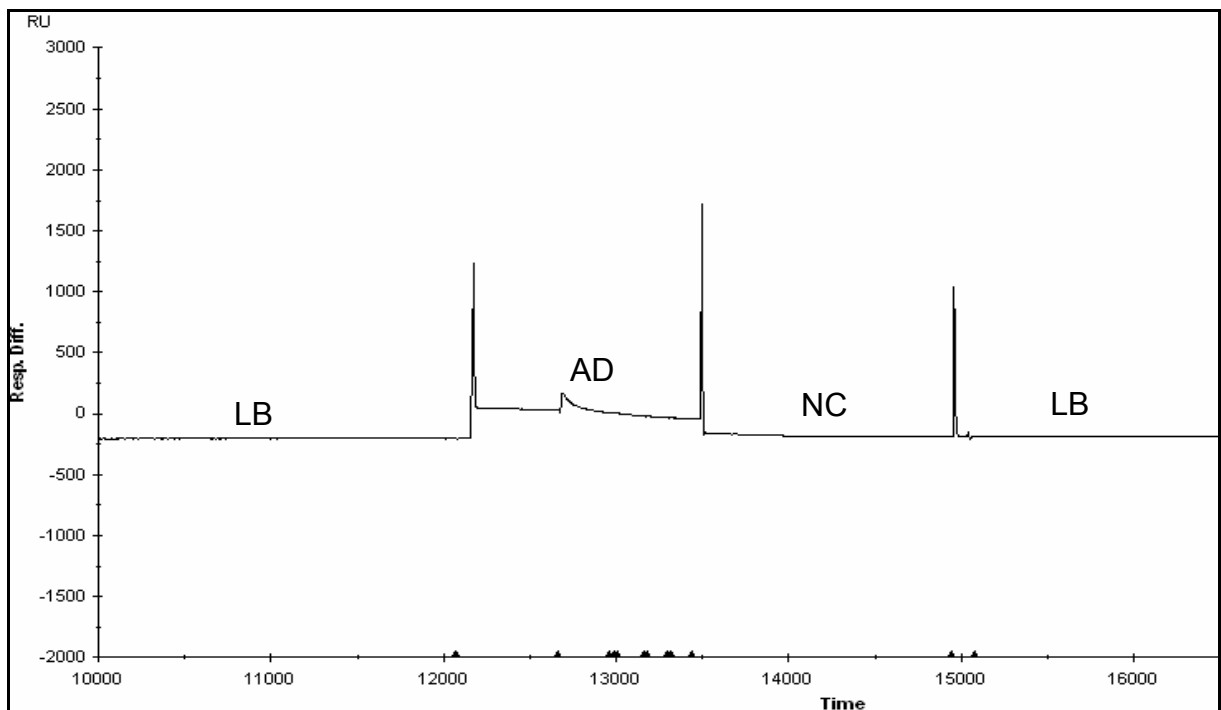


Figura 3. Sensorgrama da subtração de FC2-FC1 após a mesma injeção do domínio Ets de Spi-C da figura 1. Este seria o sensorgrama de onde seriam retiradas as constantes de associação e de dissociação. O perfil de resposta esperado neste sensorgrama teria que mostrar uma curva de associação e dissociação parecida com as das células separadas (Figuras 1 e 2) e o nível da linha de base (LB) deveria voltar ao normal após a regeneração. Neste caso a linha de base diminuiu após a injeção de NaCl 2M (NC), pois FC2 sempre ficou maior que FC1. O mesmo foi observado a cada injeção de proteína.

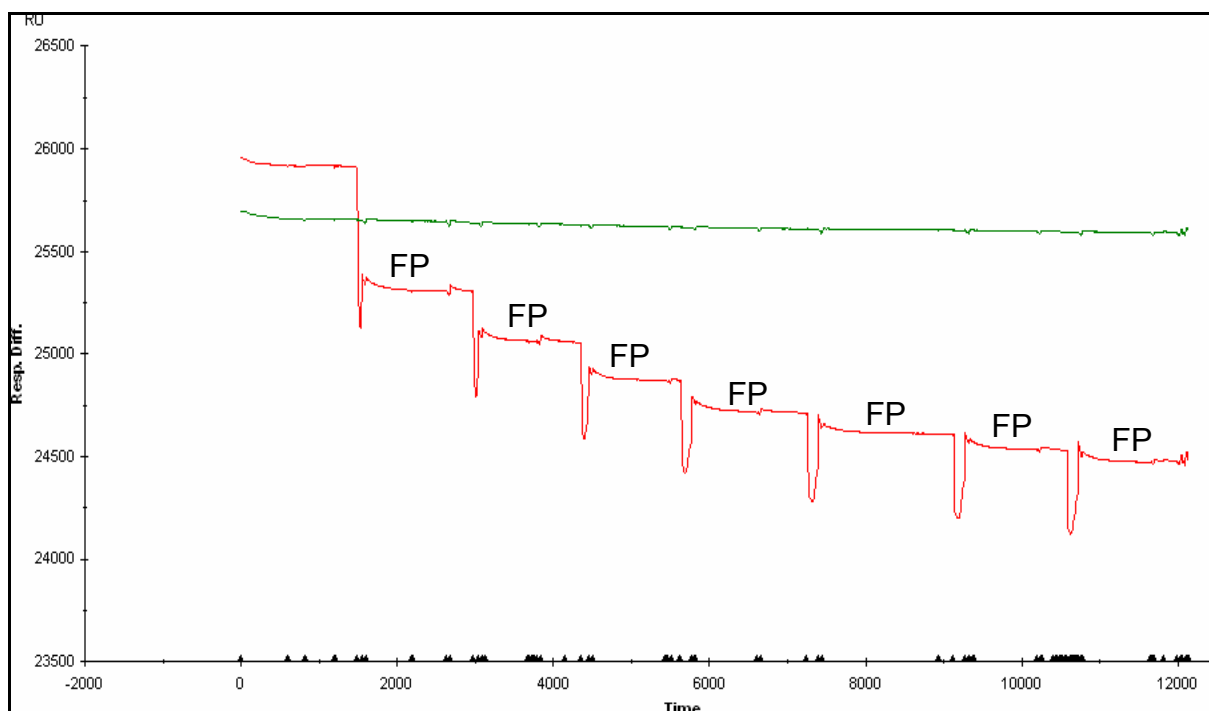


Figura 4. Sensorgrama mostrando FC1 (vermelho) e FC2 (verde) separadamente após teste de regeneração de FC1. Cada curva de dissociação observada em FC1 (FP) ocorreu após injeção de tampão fosfato de potássio pH 2,0, o que carregou positivamente as carboxilas do dextran e fez com que a proteína (carregada positivamente) não tivesse mais afinidade pela superfície do chip. Este teste foi feito para confirmar que o acúmulo de massa no chip foi devido à ligação da proteína com o dextran e mesmo com essa diminuição drástica de pH a linha de base em FC1 não voltou ao valor original.

Espectrofluorimetria

A proteína recombinante possui dois resíduos de triptofano em sua estrutura primária, o que pode levar à emissão de fluorescência quando excitada em determinados comprimentos de onda. Primeiramente foi retirado o espectro de absorção na área do UV-visível do domínio Ets de Spi-C a $2,0\mu\text{M}$ em tampão definido em espectrofotômetro (modelo UV-2550, Shimadzu, Kyoto, Japan). O espectro de absorção do DNA em solução já estava definido (não mostrado). Em espectrofotofluorímetro (modelo RF-5301PC, Shimadzu, Kyoto, Japan) com a temperatura constante (25°C), a emissão de fluorescência do domínio Ets de Spi-C foi testada em vários comprimentos de onda incidentes pois os comprimentos de onda com maior absorção da proteína se sobrepõem à faixa de absorção do DNA. Definiu-se que o melhor comprimento de onda incidente para o experimento de ligação ao DNA foi 297nm. Todos os experimentos foram realizados na mesma temperatura.

Foi observada a mudança na emissão de fluorescência da proteína com a adição de DNA, porém os dados obtidos não são conclusivos, visto que a proteína não emite fluorescência de forma constante ao longo do tempo. Com o passar do tempo, a emissão de fluorescência vai diminuindo aos poucos, tanto na sua forma livre (Gráfico 1) quanto após a adição de uma quantidade suficiente e predeterminada de DNA (Gráfico 2). Isto foi comprovado após a verificação constante da fluorescência máxima emitida ao longo do tempo durante 5 horas para ambas as situações. Ao analisar a variação das diferenças de fluorescências ao longo do tempo obtêm-se uma curva com o perfil de associação, mas como a adição do DNA também diminui a absorção de luz incidente e, assim, diminui a emissão de fluorescência, não há como mensurar o quanto o aumento de concentração de DNA interfere na diminuição da emissão de fluorescência, nem mesmo se a formação do complexo aumenta ou diminui a fluorescência da proteína em determinado comprimento de onda (Gráfico 3). Levando em consideração as informações descritas anteriormente não foi possível retirar qualquer tipo de constante de ligação ao DNA pelo domínio Ets da proteína Spi- C.

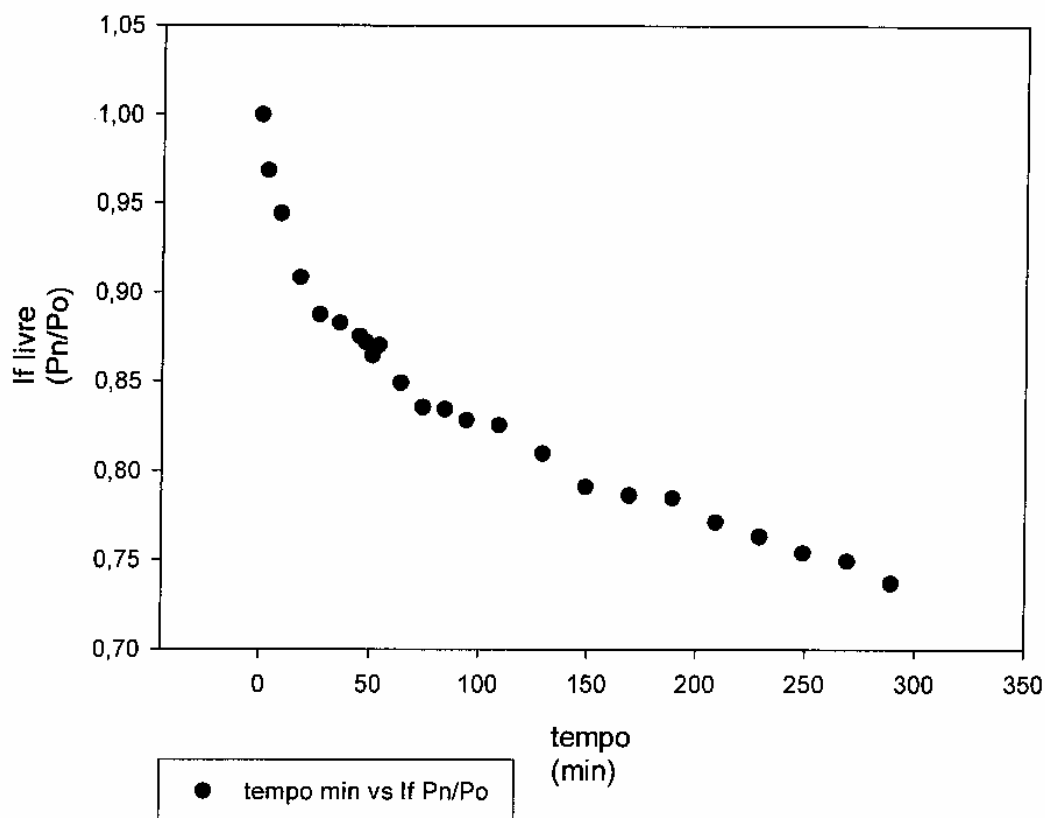


Gráfico 1. Diferença da intensidade de fluorescência da proteína livre a 2 μ M versus o tempo. Medidas de comprimento de onda de emissão 332 nm.

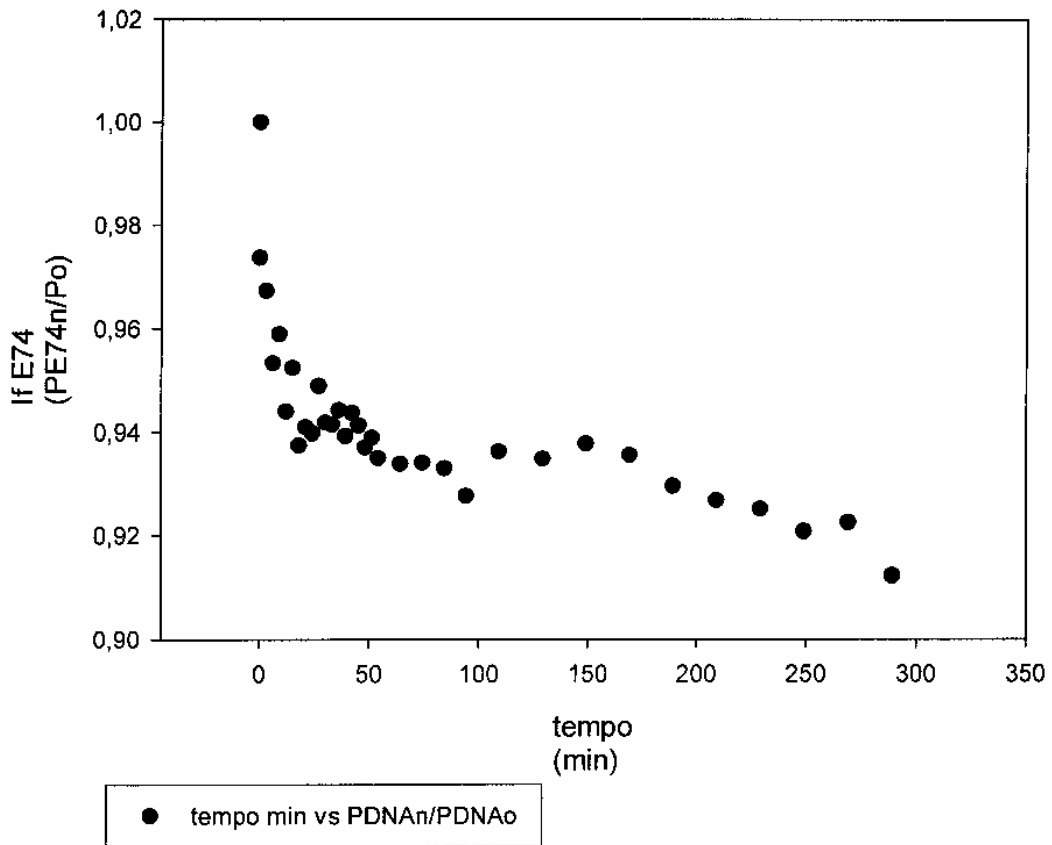


Gráfico 2. Diferença da intensidade de fluorescência da proteína a 2,0 μM com adição de 0,338 μM DNA versus o tempo. Medidas de comprimento de onda de emissão 332 nm.

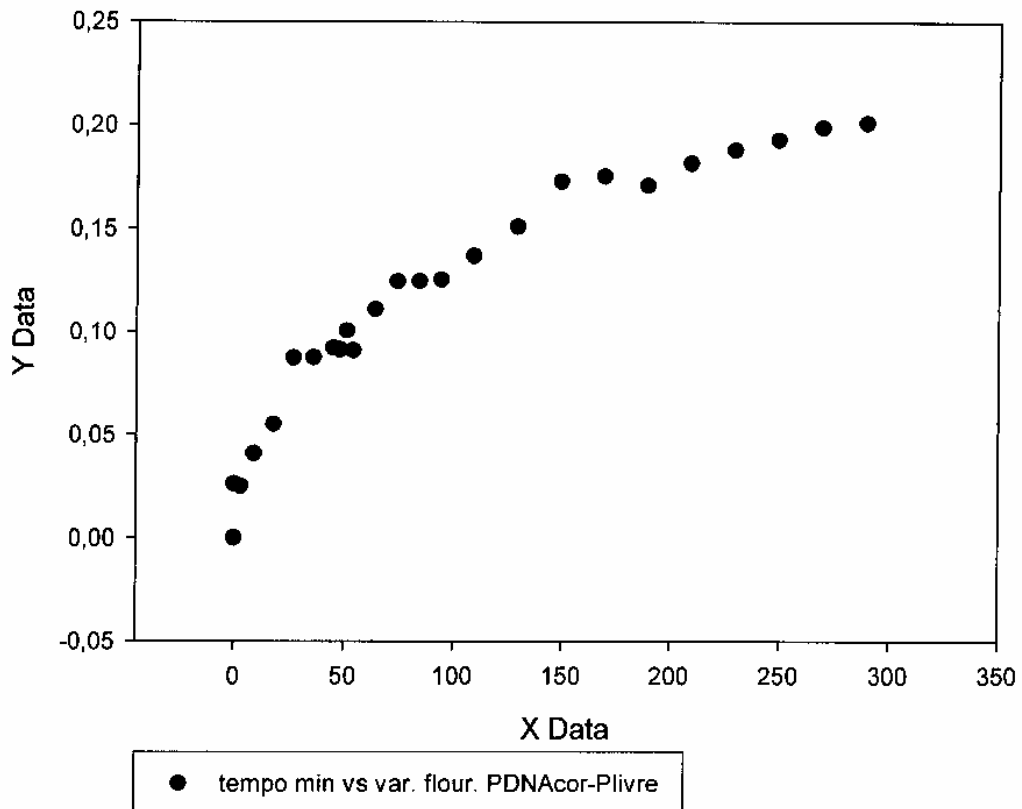


Gráfico 3. Variação das diferenças de fluorescências ao longo do tempo. Eixo das abscissas: tempo em minutos; eixo das ordenadas: diferença da fluorescência da proteína com DNA dividido pela diferença da fluorescência da proteína livre nos mesmos momentos. Medidas de comprimento de onda de emissão 332 nm.