



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

**Ligação do domínio Ets de ESE-3 à sequência de DNA contendo as bases centrais 5'-GGAX-3' (X= A ou T): Análises cinéticas e em equilíbrio estimadas por ressonância plasmônica de superfície**

Autor

**Léia Jaskulski**

Orientador

**Prof. Dr. Luiz Augusto Basso**

Co-orientador

**Prof. Dr. Diógenes Santiago Santos**

Porto Alegre  
Março de 2009



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

**Ligação do domínio Ets de ESE-3 à sequência de DNA contendo as bases centrais 5'-GGAX-3' (X= A ou T): Análises cinéticas e em equilíbrio estimadas por ressonância plasmônica de superfície**

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

Autor

**Léia Jaskulski**

Orientador

**Prof. Dr. Luiz Augusto Basso**

Co-orientador

**Prof. Dr. Diógenes Santiago Santos**

Porto Alegre  
Março de 2008

## **AGRADECIMENTOS**

Agradeço aos meus orientadores Prof. Luiz Augusto Basso e Prof. Diógenes Santiago Santos pela oportunidade, ensinamentos e apoio dispensado na realização deste trabalho.

À Jocelei Chies, Gaby Renard e Cláudia Paiva Nunes pela ajuda, apoio, atenção e ensinamentos durante a execução deste trabalho.

Aos colegas da Quatro G Ltda. Pesquisa & Desenvolvimento Ltda e do Centro de Pesquisas em Biologia Molecular e Funcional, dos quais, o carinho, a amizade e a ajuda dada foram fundamentais para a realização deste trabalho.

Aos meus pais, avós e tios, os quais tiveram uma participação intensa na minha formação pessoal e profissional; apoiando, incentivando e ajudando em todos os momentos.

Por fim, muito obrigado a todos que de alguma forma contribuíram para a realização deste trabalho bem como para a minha formação.

## RESUMO

Tem se mostrado que a proteína ESE-3 (Fator específico de epitélio, do inglês *epithelium-specific Ets factor, family member 3*) desempenha importantes papéis na determinação do destino inicial da diferenciação do epitélio escamoso estratificado, na indução da expressão do receptor para um ligante que induz à apoptose, relacionado com necrose tumoral, e em processos inflamatórios de doenças respiratórias, tais como a asma e a fibrose cística. Tem se proposto que proteínas com domínio Ets ligam-se com maior afinidade a seqüências com núcleo central rico em purinas 5'-GGAA-3' do que em seqüências com núcleo central 5'-GGAT-3'. Neste trabalho, nós descrevemos análises cinéticas e em equilíbrio do domínio Ets de ESE-3 ligando-se ao sítio de ligação Ets presente na região promotora do gene E74 de *Drosophila*, contendo tanto as bases centrais 5'-GGAA-3' quanto 5'-GGAT-3', utilizando a técnica de ressonância plasmônica de superfície. Também descrevemos ensaios de competição. Os resultados mostraram que não há grande diferença na afinidade de ligação do domínio Ets de ESE-3 entre as seqüência contendo 5'-GGAA-3' ou 5'-GGAT-3'. No entanto, as análises dos dados cinéticos indicam que há maior valor da constante de velocidade de associação e menor valor para a constante de velocidade de dissociação, que contam para uma afinidade levemente mais alta do domínio Ets de ESE-3 por seqüências de DNA contendo as bases centrais 5'-GGAA-3'. O papel da dinâmica do reconhecimento do DNA realizado pela proteína é discutido e é proposto o mecanismo no qual o domínio Ets de ESE-3 livre existe como dois isômeros em solução que sofrem um passo lento de isomerização, que é seguido por um rápido processo de ligação ao DNA. Além disso, os dados obtidos pelos experimentos de espectrofluorimetria dão uma evidência para a inclusão de mais um passo no mecanismo proposto.

Palavras-chave: ESE-3, ressonância plasmônica de superfície, domínio Ets, cinética, ligação em equilíbrio.

## ABSTRACT

The epithelium-specific Ets factor, family member 3 (ESE-3), has been shown to play roles in determination of the early differentiated fate in stratified squamous epithelia, expression of a receptor for tumor necrosis factor-related apoptosis-inducing ligand, and airway inflammation of respiratory disorders such as asthma and cystic fibrosis. The Ets domain proteins have been proposed to bind to purine-rich central core 5'-GGAA-3' sequence with larger affinity than 5'-GGAT-3'. Here we describe surface plasmon resonance equilibrium and kinetics analyses of Ets domain ESE-3 binding to *Drosophila* E74 gene promoter sequence (a consensus Ets binding site) containing either 5'-GGAA-3' or 5'-GGAT-3' core sequence. Competition assays are also described. The results showed that there is no large difference in equilibrium binding affinity between Ets domain ESE-3 and either 5'-GGAA-3' or 5'-GGAT-3'. However, analysis of kinetic data indicates that larger association rate and lower dissociation rate constant values account for the slightly higher affinity of Ets domain ESE-3 for 5'-GGAA-3'-containing DNA sequence. The role of dynamics in protein-DNA recognition is discussed and a mechanism is proposed in which free Ets domain ESE-3 exists as two isomers in solution that undergoes a slow protein isomerization step, which is followed by a fast DNA binding process. Besides, data of spectrofluorimetric experiments give evidence to include a further step in the proposed mechanism.

Key-words: ESE-3, surface plasmon resonance, Ets domain, kinetics, equilibrium binding.

## ÍNDICE

<b>1 INTRODUÇÃO E OBJETIVOS</b>	7
<b>1.1 Introdução</b>	7
<b>1.1.1 A família Ets</b>	7
<b>1.1.2 O domínio Ets e sua interação com o DNA</b>	8
<b>1.1.3 A proteína ESE-3 humana</b>	11
<b>1.2 Objetivos</b>	19
<b>1.2.1 Objetivo Geral:</b>	19
<b>1.2.2 Objetivos específicos:</b>	19
<b>2 ARTIGO CIENTÍFICO</b>	20
<b>Abstract</b>	22
<b>Introduction</b>	23
<b>Materials and Methods</b>	25
<b>Results</b>	29
<b>Discussion</b>	32
<b>Funding</b>	34
<b>References</b>	35
<b>Figure Legends</b>	38
<b>Figures</b>	41
<b>3 CONSIDERAÇÕES FINAIS</b>	44
<b>REFERÊNCIAS BIBLIOGRÁFICAS</b>	47
<b>ANEXO I</b>	52
<b>APÊNDICE – Comprovante de submissão do artigo científico</b>	55

# 1 INTRODUÇÃO E OBJETIVOS

## 1.1 Introdução

### 1.1.1 A família Ets

Membros da família de fatores de transcrição de eucariotos, denominados Ets, foram originalmente identificados tendo por base a homologia da sequência primária com a proteína codificada pelo oncogene *v-ets* que está presente na região E26 (*E twenty six*) do vírus da eritroblastose aviária [1]. Baseado na homologia do domínio Ets da proteína Ets-1, o membro fundador da família dos fatores de transcrição Ets, outras proteínas Ets foram identificadas. No entanto, esta família de fatores de transcrição parece única na linhagem dos metazoários e está presente desde as esponjas até o homem [2, 3].

Todos os membros possuem um domínio ativador ou repressor da transcrição e também um domínio de ligação ao DNA evolutivamente conservado [4,5]. Mais de 200 genes alvos de Ets já foram descobertos, e mais de 400 genes alvos de Ets têm sido definidos baseados na presença funcional do EBS (do inglês *Ets binding site*) presentes nas suas regiões regulatórias. Incluem-se como genes alvos para as proteínas Ets: oncogenes, genes supressores tumorais, genes relacionados a apoptose, diferenciação, angiogênese, invasão e metástase [5].

O produto gênico desta família são fatores de transcrição que controlam várias funções em cooperação com outros fatores de transcrição e co-fatores resultando, portanto, na regulação da transcrição. As interações combinatórias, dependem do ambiente no qual se encontram (por exemplo: a fase embrionária apresenta ambientes diferentes do que na fase adulta de um indivíduo) e definem a especificidade para genes alvos [4,6,7,8]. Na realidade, as diferenças na especificidade de ligação resultam em alguns fatores de transcrição Ets atuando como repressores enquanto outros atuam como ativadores da transcrição [1]. Mais recentemente, evidenciou-se que alguns membros da família possuem ambas as capacidades, de repressores e ativadores, o que indica que esta distinção pode não ser absoluta.

### 1.1.2 O domínio Ets e sua interação com o DNA

A característica que unifica essa família de fatores de transcrição é a presença do domínio Ets, que é responsável por mediar a ligação ao DNA alvo, o qual contém uma sequência invariável central GGA. O domínio Ets é uma região conservada de aproximadamente 85 aminoácidos que pode estar localizado tanto na porção proteica C-quanto N-terminal (sendo na maioria das vezes presente na porção N-terminal). Os fatores de transcrição Ets são subdivididos em várias subfamílias baseados na homologia entre os domínios Ets e outros domínios, como por exemplo, o domínio Pointed responsável por realizar homo-oligomerização e heterodimerização. O domínio Ets pode ser produzido como um fragmento estável de proteína e é suficiente para ligar-se ao DNA [9]. No entanto, não é o único fator que contribui para determinar a especificidade de ligação das proteínas Ets, uma vez que interações com outros fatores de transcrição também influenciam a especificidade de ligação às sequências regulatórias específicas.

A elucidação da estrutura tridimensional dos membros da família Ets (Fli-1 [10], Ets-1 [11,12], PU.1 [13], GABP $\alpha$  [14], SAP-1 [15] e Elk-1 [16]) revelou que o domínio Ets apresenta um alto grau de conservação estrutural e adota uma estrutura secundária do tipo *winged helix-turn-helix* (*wHTH*), nos quais a segunda e a terceira hélices  $\alpha$  (H2-H3) formam o motivo *helix-turn-helix* (*HTH*), a região de reconhecimento do DNA (Fig.1). O termo *winged helix-turn-helix* é usado para descrever este domínio e chamar a atenção para a importância dos dois elementos, o *helix-turn-helix*, e as folhas  $\beta$  (*wing*) na estrutura e no contato com o DNA [9]. O domínio Ets é constituído de três hélices  $\alpha$  e quatro folhas  $\beta$ -pregueadas antiparalelas ( $\beta$ -strands), que são organizadas na seguinte forma: H1-S1-S2-H2-H3-S3-S4 (H= hélice  $\alpha$ ; S= folha  $\beta$ ) [17, 18]. Os principais contatos com o DNA são realizados pela segunda e terceira hélices alfa (*HTH*) (a região de reconhecimento do DNA), sendo que a terceira hélice  $\alpha$  está incrustada e é responsável pelos contatos com o sulco maior do DNA, e também, pelas alças (*loop*) que interagem com o sulco menor do DNA (uma localizada entre as fitas  $\beta$ -pregueadas 3 e 4 e a outra entre as hélices alfa 2 e 3) [19, 20].

Um exemplo já conhecido da estrutura terciária do domínio Ets interagindo com o DNA pode ser observado na Figura 1. O domínio Ets da proteína PU.1 contacta o DNA em três sítios: a hélice de reconhecimento  $\alpha$ -3 interage com a sequência central 5'-GGAA-3' presente na curvatura maior do DNA, na qual duas argininas (resíduos invariáveis na família Ets) promovem o contato com duas guaninas da sequência central GGA, enquanto as alças; sendo uma localizada entre a fita  $\beta$ 3 e  $\beta$ 4 e a outra situada entre a segunda e a terceira hélices



$\alpha$ ; interagem com os grupos fosfatos da dupla-fita de DNA presentes na curvatura menor do DNA estabilizando o complexo [9,19,21].

As proteínas Ets se ligam especificamente aos EBSs contendo uma seqüência central 5'-GGA (A ou T)-3' e selecionam nucleotídeos, os quais flanqueiam este centro, e contribuem pelo menos em parte para a especificidade de ligação de cada membro da família [22]. Selecionam, portanto, uma seqüência de aproximadamente 11 pares de bases [1, 9]. Têm sido propostas seqüências consenso para vários fatores de transcrição com domínio Ets. Ocorre uma significativa sobreposição entre essas seqüências consensos sobre uma região de 9 pares de bases, o que reforça a similaridade nas suas especificidades de ligação ao DNA. No entanto, inferir conclusões a respeito da especificidade de ligação baseado-se somente nas seqüências consenso pode induzir ao erro. Por sua vez, a determinação de um sítio alvo para uma determinada proteína Ets pode somente ser averiguada por experimentação direta. Portanto, a determinação da especificidade aparente de ligação ao DNA para cada proteína Ets mostra-se dependente da abordagem experimental [23].

Os fatores de transcrição com domínio Ets se ligam, por exemplo, à seqüência AACCGGAAGTA (encontrado no promotor do gene E74 de *Drosophila*) que representa um EBS de alta afinidade e é freqüentemente utilizado para detectar sua atividade de ligação [7]. Várias características são observadas na seleção das seqüências consenso e como exemplos, podemos referenciar algumas proteínas, entre elas a PU.1, que reconhece a maioria dos elementos que apresentam a seqüência central 5'-GGAA-3' mas nunca 5'-GGAT-3', enquanto outras proteínas da família Ets reconhecem tanto 5'-GGAA-3' quanto 5'-GGAT-3' como ocorre com a proteína Ets-1 [21]. No entanto, vários elementos que respondem a PU.1 apresentam uma seqüência central que é distinta das demais; 5'-AGAA-3' [24]. Ets-1 reconhece a seqüência 5'-GGAA-3' com alta afinidade e 5'-GGAG-3' com baixa afinidade [25]. A ligação das proteínas E74 e Elf-1 ao DNA é fortemente favorecida em relação ao EBS que contém no seu centro a seqüência 5'-GGAA-3' [26], uma propriedade que é determinada pela presença de um resíduo de treonina na terceira hélice alfa destas proteínas, enquanto que todas as outras proteínas Ets apresentam um resíduo de lisina na posição equivalente, e portanto, não apresentam a mesma seleção restrita [22].

A importância do nucleotídeo localizado à direita do motivo GGA, na especificidade de ligação ao DNA, pode ser constatada através da observação da sua interação com resíduos não conservados que afetam a ligação ao DNA. Por exemplo, no mutante da proteína Ets-1 (de galinha) gerado pela troca de um resíduo de lisina, presente no domínio Ets da proteína Ets-1, por um resíduo de treonina (o qual é encontrado em posição equivalente na proteína Elf-1), torna-o homólogo ao resíduo equivalente em Elf-1 e, interessante, o mutante

gerado é capaz de reconhecer somente seqüências contendo no seu centro as bases GGAA, sendo, portanto, incapaz de reconhecer a seqüência central 5'-GGA (T)-3', como faz essa proteína selvagem [22].

A conservação dos resíduos do domínio Ets indica propriedades funcionais conservadas. No entanto, uma considerável variação é observada entre os membros da família, conferindo propriedades únicas a cada uma delas. Na verdade, a alteração de um simples aminoácido na extremidade carboxi-terminal da terceira hélice  $\alpha$ , de reconhecimento ao DNA, no domínio Ets pode alterar sua especificidade de ligação ao DNA [27] e suas interações com outros fatores de transcrição [28]. Podemos então questionar como as proteínas Elk-1 e SAP-1, com alto grau de similaridade nos aminoácidos presentes na hélice de reconhecimento do DNA, podem ter diferentes propriedades de ligação ao DNA? A comparação do domínio Ets de Elk-1 ligada ao sítio de ligação presente na região promotora do gene E74 de *Drosophila* (Elk-1-E74), de alta afinidade, em comparação com o complexo SAP-1-E74 mostrou que há diferenças nas propriedades de ligação ao DNA dessas proteínas, dadas por resíduos não conservados no final da hélice de reconhecimento do DNA (ácido aspártico na posição 69 localizada na porção C-terminal da hélice  $\alpha$  do domínio Ets de Elk-1) apresenta a função de orientar os contatos realizados pelos resíduos conservados presentes na hélice de reconhecimento do DNA (especialmente da tirosina 66 do domínio Ets de Elk-1). Como resultado, quase um terço das interações entre as hélices de reconhecimento dessas proteínas com o DNA são diferentes, mesmo contendo um alto grau de similaridade, incluindo a seqüência de aminoácidos da hélice de reconhecimento do domínio Ets dessas proteínas [29].

Além disso, devemos considerar as proteínas co-regulatórias. Até o momento não está claro se as proteínas Ets funcionam sem o auxílio das proteínas co-regulatórias *in vivo*. No entanto, tais interações claramente aumentam a especificidade de reconhecimento a um dado promotor. Com isso, podemos dizer que devem existir tanto mecanismos que reforçam a especificidade de ligação para cada membro da família de fatores de transcrição Ets a um EBS presente na região promotora de um gene em particular quanto os que inibam a ligação promíscua da proteína para um sítio de ligação Ets [1].

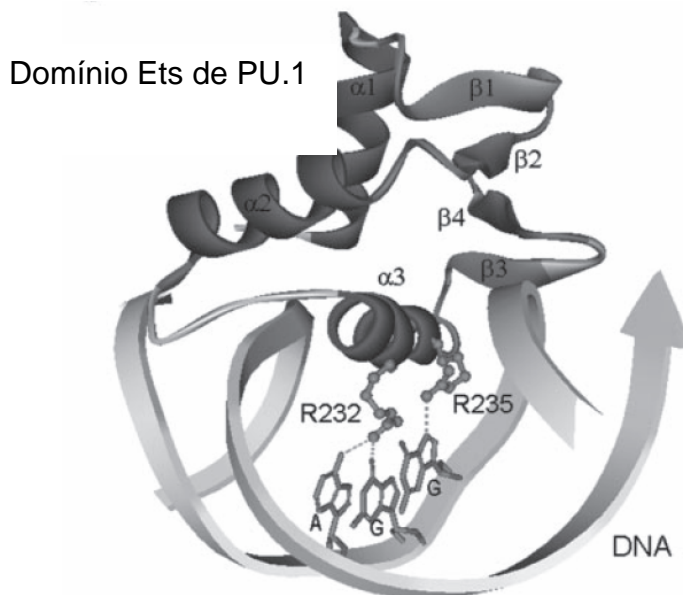


Figura 1. Estrutura terciária do domínio Ets de PU.1. A estrutura é composta por quatro folhas  $\beta$ -pregueadas antiparalelas (S) deitadas numa região de hélice-volta-hélice. A terceira  $\alpha$ -hélice (H) se acomoda no sulco maior do DNA e reconhece o consenso GGAA. O reconhecimento ocorre principalmente por meio de pontes de hidrogênio entre o consenso e as argininas. As alças localizadas entre as fitas  $\beta$ 3 e  $\beta$ 4 e entre a segunda e a terceira hélices  $\alpha$  interagem com os grupos fosfatos do sulco menor do DNA e estabilizam o complexo. Fonte: LISTMAN et al. (2005)

### 1.1.3 A proteína ESE-3 humana

#### Os domínios de ESE-3

Os genes ESE-1, ESE-2 e ESE-3 codificam para proteínas que pertencem à subfamília ESE dos fatores de transcrição Ets [30,31,32]. Estas proteínas contêm um motivo conservado de ligação ao DNA, o domínio Ets, e um domínio conservado Pointed que está envolvido em interações proteína-proteína. O domínio Ets de ESE-3 está localizado na porção carboxi-terminal da proteína [31].

A partir de um estudo realizado com células cultivadas in vitro, foi sugerido para ESE-2 e ESE-3 que o domínio Ets funciona como um repressor, e o domínio Pointed pode estar atuando como um potente domínio de ativação da transcrição [33].

ESE-3 (Fator Ets específico de epitélio, do inglês *Epithelium-specific Ets factor, family member 3*) possui duas isoformas alternativas, ESE-3a e ESE-3b. O *splicing* alternativo que gera ESE-3b ocorre devido à presença de uma seqüência de 69 nucleotídeos na região

central de ESE-3a. A comparação da seqüência deduzida de aminoácidos do domínio Ets de ESE-3 com outros membros da família Ets revelou maior nível de identidade com o domínio Ets dos fatores específicos de epitélio ESE-1 (ESX, jen, ELF3, ERT) (84%) [34] e ESE-2 (ELF5) (65%) [30]. O alinhamento realizado por Kas (2000) do domínio Ets de ESE-3 com os outros membros da família (Fig 2), exceto ESE-1 e ESE-2, revelou maior similaridade com E74 (51%), NERF (49%), MEF (48%) e ERP (46%). Este grau de identidade é muito abaixo das características similaridades existentes entre os conhecidos membros da família Ets [31].

O domínio *Pointed* (PNT) de ESE-3, localizado na porção amino-terminal, mostrou maior similaridade à ESE-2 do que à ESE-1. O fato do domínio *Pointed* não ser encontrado nos fatores de transcrição ELF1, NERF, e MEF, fatores Ets mais próximos filogeneticamente à ESE-3 (exceto ESE-1 e ESE-2), fornece apoio para a classificação das três proteínas em uma subfamília separada, a subfamília ESE. O domínio *Pointed* faz parte da classe de domínios SAM. Este domínio tem um papel funcional na mediação da homodimerização e heterodimerização [31].

A seqüência de aminoácidos deduzida de ESE-3a e ESE-3b produzem proteínas ricas em asparagina (8%), serina (7%), e leucina (11%). A porção amino-terminal de ESE-3 é caracterizada por uma abundância de resíduos ácidos, enquanto a porção carboxi-terminal por resíduos básicos. A porção central da proteína contém domínios ricos em leucina e serina [31].

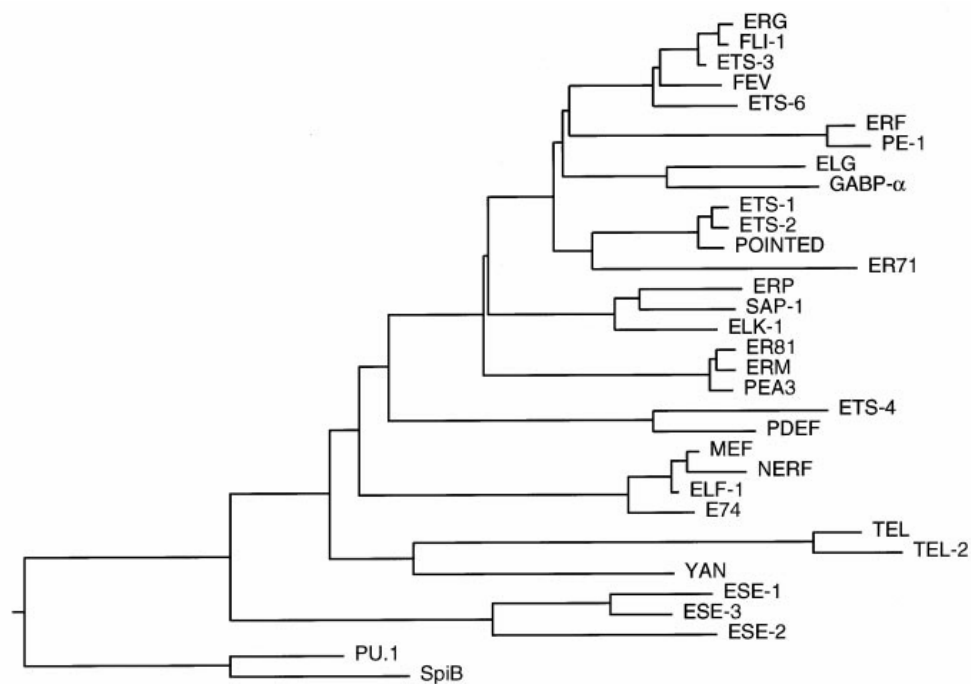


Figura 2. Análise filogenética do domínio Ets de “todos os membros conhecidos da família Ets”, relacionando a similaridade entre as seqüências de aminoácidos. Somente fatores Ets humanos e de *Drosophila* estão incluídos. Os números de acesso ao Genbank obedecem a: ERG, M21536; ERG-B/FLI-1, Y17293; ETS-3, M88473; FEV, Y08976; ETS-6, M88475; ERF, U15655; PE-1/ETV3, L16464; GABP- $\alpha$ /E4TF1-60, Q06546; ELG, M88471; ETS-1, X14798; ETS-2, AF017257; POINTED, S33167; ER71/ETV2, AC002115; ERP/NET, Z36715; SAP-1, P28323; ELK1, P19419; ER81/ETV1, U17163; ERM, X96381; PEA3/E1AF/ETV4, U18018; ETS4, M88474; PDEF, AF071538; MEF, U32645; NERF, U43188; ELF-1, P32519; E74, A53225; TEL/ETV6, U11732; TEL-2, AF116509; YAN, Q01842; ESE-1, U73844; ESE-2, AF115402; ESE-3 (AF124439); PU.1, X52056; SpiB, X66079 (adaptado de Kas, K. *et al*, 2000).

### Expressão de ESE-3 e seus papéis nas distintas células humanas

A subfamília ESE (*Epithelium-specific Ets factor*) é composta por três membros, ESE-1/ELF3/ESX, ESE-2/ELF-5/ASTH1-I, e ESE-3/EHF/ ASTH1-J, e são as únicas proteínas Ets na qual a expressão é limitada a células epiteliais [30-32,34,35]. Embora esses fatores de transcrição sejam expressos exclusivamente em células epiteliais, nem todas as células epiteliais expressam as proteínas Ets da subfamília ESE. Em adultos, o nível mais alto do RNA mensageiro de ESE-1 é encontrado no intestino e níveis muito inferiores são encontrados no pulmão [34]. Os mais altos níveis do mRNA de ESE-2 são encontrados em glândulas salivares, glândulas mamárias e na traquéia; ESE-2 apresenta assim como ESE-1, níveis inferiores no pulmão [30]. O mRNA de ESE-3 é encontrado numa variedade maior de tecidos, sendo encontrados os mais altos níveis de mRNA de ESE-3 na traquéia, no pulmão, no cólon, no pâncreas e na próstata, quando comparada a ESE-2 [31,32]. Têm-se considerado à hipótese de que os membros da família ESE, por causa da sua habilidade em regular

diferenciação e crescimento epitelial, teriam um papel em tubulogênese e em morfogênese por ramificação em órgãos contendo epitélio glandular, tais como a traquéia, e em oncogêneses de tumores derivados de epitélio, como o carcinoma brônquico [31,32].

Recentemente, Appel (2005) demonstrou que o fator de transcrição epitelial ESE-3, o qual foi previamente mostrado ser exclusivamente expresso em células de origem epitelial, é diferentemente expresso em células dendríticas derivadas de monócitos. As células dendríticas (DCs) são as células que apresentam antígenos de forma mais potente e possuem um importante papel na iniciação e manutenção da resposta imune primária. Durante a diferenciação das células dendríticas, há a expressão da proteína ESE-3 e a ativação da sua expressão se dá por meio de estímulos dados durante a maturação. Os transcritos e a proteína de ESE-3 podem ser encontrados em diferentes subpopulações, incluindo as DCs de sangue periférico e as células de Langerhans. Estudos sugerem que o fator de transcrição ESE-3 está criticamente envolvido no desenvolvimento das DCs originadas de monócitos do sangue periférico [36].

ESE-2 e ESE-3 são de interesse particular para os geneticistas por serem polimórficos e estão localizados no cromossomo 11p12-15, uma região genômica relacionada à suscetibilidade ao desenvolvimento da asma em diversas varreduras de todo o genoma humano, e por isso, são identificados como genes relacionados ao desenvolvimento da asma [5]. Estudos subsequentes, envolvendo várias populações diferentes, sustentam a associação entre o polimorfismo em ESE-2 e ESE-3 e o diagnóstico da asma; no entanto, nenhuma forma variante desses genes foi associada com o diagnóstico da asma, e também o papel dessas proteínas na biologia das vias aéreas e essa patogênese não estão estabelecidos [37,38].

Os genes que codificam para ELF5/ESE-2 e EHF/ESE-3 também foram encontrados em uma região (cromossomo humano 11p13 [39]), que foi previamente associada com LOH (*loss of heterozygosity*) em carcinomas epiteliais [40-43]. No caso de alguns carcinomas de pulmão analisados, a perda de heterozigose tem sido associada com os mais altos índices de mortalidade, resultando numa maneira direta de mensurar o potencial de metástase de tumores, sugerindo que esta região pode conter genes requeridos para o controle da migração celular [41-43].

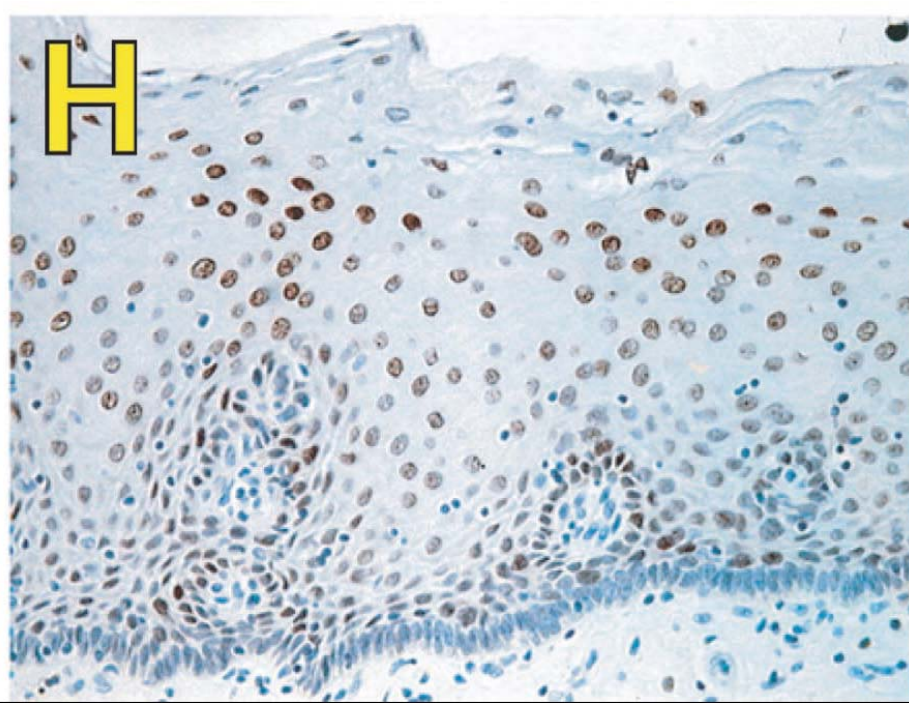


Figura 3. ESE-3 é uma proteína expressa no núcleo das células epiteliais diferenciadas. O tecido humano foi submetido à análise imuno-histoquímica usando o anticorpo monoclonal. A coloração marrom revela imuno-reatividade, e a coloração azul corresponde às células coradas com hematoxilina. H indica epitélio estratificado de esôfago (adaptado de Tugores, A. et al, 2001).

Análises imuno-histoquímicas realizadas por Tugores *et al.* (2001) [32] com anticorpo monoclonal revelaram que ESE-3 é uma proteína expressa exclusivamente no núcleo de células epiteliais diferenciadas e está ausente em carcinomas epiteliais testados. Isso é mais evidente em epitélio escamoso estratificado de esôfago, onde ESE-3 está ausente nas células-tronco (indiferenciadas), como as da camada basal, e aparece assim que elas se diferenciam (Fig. 3). A expressão de ESE-3 é perdida em carcinomas epiteliais, incluindo o da bexiga, do epitélio escamoso bucal, do ducto mamário e nos carcinomas epiteliais da próstata (Fig. 4). Certamente, pode se discutir que a diminuição da expressão de ESE-3 não é um evento patológico essencial e que outros mecanismos oncogênicos levam à perda da expressão desta proteína. Na verdade, embora todos os dados sejam sugestivos, não há evidências diretas relacionadas com a perda da expressão de ESE-3 e a progressão do câncer. Contudo, a associação restrita de ESE-3 com células epiteliais diferenciadas indica que, pelo menos, esta poderia ser um marcador para a malignidade. ESE-3 comporta-se como um repressor de ativação da transcrição induzida por Ras ou ésteres de forbol de um subgrupo de promotores que contêm sítios de ligação para AP-1 e Ets, incluindo o promotor do gene da colagenase intersticial (MMP-1; matrix metaloproteinase 1). A repressão mediada por ESE-3 depende da seqüência e do contexto do ambiente, depende da presença de um sítio de ligação de alta

afinidade para a proteína ESE-3 em combinação com cis-elementos AP-1 e do arranjo desses sítios em um dado promotor, indicando que ESE-3 é um repressor específico para um subgrupo de genes que respondem a AP-1/Ets. Tugores *et al.* (2001) propõem que ESE-3 pode ser uma proteína que determina o controle da diferenciação epitelial, ou seja, uma proteína que faz parte das células epiteliais diferenciadas, como um modulador da resposta nuclear frente à cascata de sinalização da proteína quinase ativada por mitógeno [32].

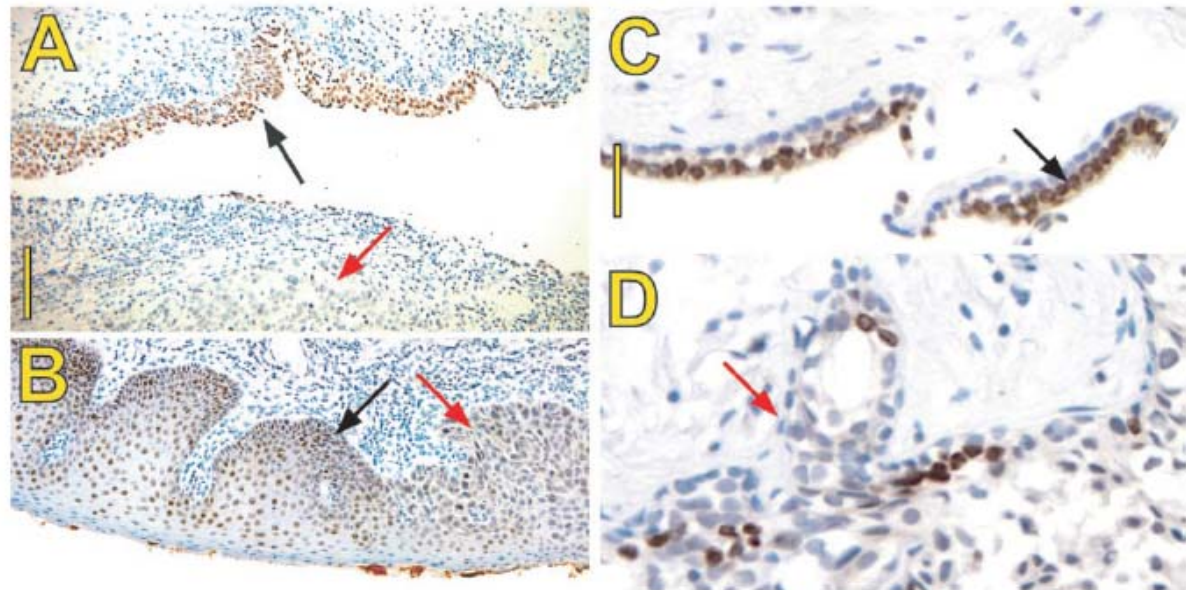


Figura 4. A imuno-reatividade de ESE-3 é perdida durante a carcinogênese. Secções de adenocarcinomas de bexiga (A) e epitélio oral (B) e carcinoma de ducto mamário (C e D), contendo tecido normal adjacente às células do carcinoma, foram sujeitas às análises imunohistoquímicas pelo uso de anticorpo monoclonal. As setas pretas indicam a expressão normal de ESE-3 no epitélio, enquanto as setas vermelhas mostram o epitélio transformado, o qual não apresenta imunoreatividade. Bar, 100 µm em A e B, 50 µm em C e D (adaptado de Tugores *et al.*, 2001).

Seth e Watson (2005) [5] sugerem um papel na supressão do crescimento celular, para os fatores de transcrição Ets (ESE-1, ESE-2, ESE-3 e PDEF), sugerindo uma possível função em supressão tumoral. A maioria dos cânceres humanos é de origem epitelial. As quatro proteínas Ets (ESE-1, ESE-2, ESE-3 e PDEF) são expressas em tecidos epiteliais incluindo mama, próstata, e cólon, tecidos frequentemente suscetíveis ao câncer. Vários estudos demonstraram que a expressão de três desses fatores está reduzida durante a tumorigênese [5]. A proteína ESE-3 está ausente em cânceres de mama, próstata, boca, e de bexiga (Fig. 4) [32]. Linhagens de células derivadas de cânceres de cólon e próstata são conhecidas por expressarem o mRNA de ESE-3, mas o nível de expressão da sua proteína não foi ainda relatado. A proteína PDEF está reduzida ou ausente em câncer de próstata e mama [44, 45] e,



significativamente, a proteína PDEF inibiu a proliferação celular, migração e invasão, consistente com seu papel como uma classe de gene supressor tumoral II [44]. Permanece por ser determinado o significado funcional da redução da expressão de ESE-2 e ESE-3 [5].

Fujikawa (2007) [46] sugere que ESE-3 têm um importante papel na indução da senescência celular como uma molécula à frente de p38 (proteína quinase ativada por mitógenos que responde a estímulos do stress, incluindo luz UV, calor, estresse osmótico, citocinas inflamatórias e fatores de crescimento). Foi identificada a expressão de ESE-3 nas células em senescência, induzidas por Ras, H<sub>2</sub>O<sub>2</sub> e encurtamento da telomerase. Experimentos *in vitro* indicam que a proteína ESE-3 liga-se ao promotor do gene p16 (supressor tumoral que atua inibindo CDK) ativando-o. Foi demonstrado recentemente que o estado de senescência celular atua como um mecanismo molecular de supressão tumoral *in vivo* [46].

Estudos realizados por Silverman (2002) [47] demonstraram que tanto a proteína ESE-3 quanto o mRNA, que codifica para a proteína ESE-3 estão constitutivamente presentes nas células epiteliais brônquicas humanas e nas células epiteliais de glândulas produtoras de muco. Coerente com esses achados, o mRNA que codifica para a proteína ESE-3 está constitutivamente presente nas células epiteliais brônquicas humanas crescidas em cultura de tecido. No entanto, o mRNA, que codifica para a proteína de ESE-2, não foi encontrado no pulmão humano ou em cultura de células epiteliais brônquicas humanas. Interessantemente, os fibroblastos e as células de músculo liso das vias aéreas não expressam a proteína ESE-3 constitutivamente; porém, após a estimulação com citocinas (interleucina-1 $\beta$  ou fator de necrose tumoral- $\alpha$ ), aumentam drasticamente os níveis do mRNA e da proteína ESE-3 por 24 h. Esta indução por citocinas é dose-dependente e é anulada por inibidores específicos da via de transdução de sinais, MEK1/2 (u0126) e P38(SB03580). A super expressão da proteína ESE-3 nas células 3T3 e nas células de músculo liso das vias aéreas humanas inibe a atividade do promotor de MMP-1, sugerindo que ESE-3 pode funcionar como um repressor da transcrição [47].

Reddy (2007) [48] demonstrou que as proteínas ESE-1 e ESE-3 ativam a transcrição do gene SPRR1B (*small proline rich protein 1B*). A proteína SPRR1B é um componente do envelope cornificado presente no epitélio escamoso estratificado, este, por sua vez, promove uma barreira física e impede a passagem da água em tecidos como a epiderme. A proteína SPRR1B não é expressa nas células epiteliais das vias aéreas, no entanto após a exposição de agentes (tabaco, carcinógenos) que causam metaplasias, há a superexpressão desta proteína no epitélio das vias aéreas. Portanto, a superexpressão da proteína SPRR1B é um biomarcador da transformação, ou seja, da presença de metaplasia nas células das vias aéreas [48,49,50].

Kas (2000) [31] demonstrou que as proteínas ESE-1 e ESE-3 comportam-se de forma diferente com respeito à ativação. No entanto, tanto a proteína ESE-3 quanto a ESE-1 ligam-se com afinidade similar a vários EBS testados. Os domínios de ligação ao DNA dessas duas proteínas são muito similares e as respectivas especificidades de ligação ao DNA são praticamente idênticas, portanto, é pouco provável que diferenças na capacidade de ativação de promotores distintos estejam relacionadas a diferenças na afinidade de ligação ao DNA. É mais provável que as diferentes capacidades de ativação estejam relacionadas com as diferenças presentes nas regiões amino-terminais dessas proteínas e da resultante interação proteína-proteína com outros fatores de transcrição e/ou co-ativadores/co-repressores que se ligam ou interagem com os elementos regulatórios únicos para cada promotor. Diferenças adicionais podem ser resultantes de modificações pós-traducionais, tais como a fosforilação de diferentes sítios por quinases [31].

Há pouquíssimas informações sobre os processos biológicos que envolvem a formação do complexo do domínio Ets da proteína ESE-3 com o DNA, fazendo com que a obtenção de dados bioquímicos seja de vital importância para o seu entendimento. Além disso, vale a pena lembrar que o conhecimento da biologia da proteína ESE-3 poderá levar à identificação de possíveis alvos moleculares de drogas que combatam o câncer.

## **1.2 Objetivos**

### **1.2.1 Objetivo Geral:**

Este trabalho tem como foco a investigação do complexo formado pelo domínio Ets, da proteína ESE-3 humana, com o DNA, e tem como objetivo principal determinar a especificidade e a cinética de ligação do domínio Ets, da proteína ESE-3, a diferentes sítios de ligação Ets. Estes dados contribuirão de forma significativa para o entendimento do mecanismo que envolve a interação do domínio Ets da proteína ESE-3 com o DNA.

### **1.2.2 Objetivos específicos:**

O presente trabalho possui os seguintes objetivos específicos que foram desenvolvidos durante o período de mestrado:

- Estabelecer o protocolo de purificação da proteína truncada, correspondente ao domínio Ets de ESE-3 humana, utilizando a técnica de Cromatografia Líquida de Alta Eficiência (*HPLC*).
- Análise da proteína por espectrometria de massa.
- Seqüenciamento N-terminal do domínio Ets da proteína ESE-3 humana.
- Caracterizar a especificidade e a cinética de ligação do domínio Ets de ESE-3 as seqüências de oligonucleotídeos contendo as bases centrais ora 5'-GGAA-3' ora 5'-GGAT-3', utilizando a técnica de espectrometria de ressonância plasmônica de superfície (*SPR*) em tempo real. Bem como inferir o mecanismo de ligação para os complexos binários formados.
- Determinar a especificidade de ligação do domínio Ets de ESE-3 as seqüências de DNA que apresentam no seu centro as bases 5'-GGAA-3', 5'-GGAT-3', 5'-GGAG-3', 5'-GGTT-3', 5'-TTAA-3' através de ensaio de competição utilizando *SPR*.

## 2 ARTIGO CIENTÍFICO

O artigo científico apresentado, *Equilibrium and kinetics of Ets domain ESE-3 binding to 5'-GGAX-3' (X=A or T) core sequence assessed by surface plasmon resonance*, foi submetido à revista *Nucleic Acids Research*, a qual, apresenta o fator de impacto de 6,954 (JCR-2007).

Equilibrium and kinetics of Ets domain ESE-3 binding to 5'-GGAX-3' (X=A or T) core sequence assessed by surface plasmon resonance.

Léia Jaskulski<sup>1,2</sup>, Cláudia Paiva Nunes<sup>3</sup>, Rodrigo Gay Ducati<sup>1</sup>, Gaby Renard<sup>3</sup>, Joclei Maria Chies<sup>3</sup>, Luiz Augusto Basso<sup>1,3,\*</sup> and Diógenes Santiago Santos<sup>1,3,\*</sup>

<sup>1</sup>Centro de Pesquisas em Biologia Molecular e Funcional, Instituto Nacional de Ciência e Tecnologia em Tuberculose, Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), 6681/92-A Avenida Ipiranga, 90619-900, Porto Alegre, RS, Brazil,

<sup>2</sup>Programa de Pós-Graduação em Biologia Celular e Molecular, PUCRS, Porto Alegre, RS, Brazil.

<sup>3</sup>Quatro G Ltda Pesquisa & Desenvolvimento, Tecnopuc-PUCRS, Porto Alegre – RS 90619-900, Brazil.

\*To whom correspondence may be addressed. Tel/Fax: +55 51 33203629; e-mail: [diogenes@pucrs.br](mailto:diogenes@pucrs.br) or [luiz.basso@pucrs.br](mailto:luiz.basso@pucrs.br).

## Abstract

The epithelium-specific Ets factor, family member 3 (ESE-3), has been shown to play roles in determination of the early differentiated fate in stratified squamous epithelia, expression of a receptor for tumor necrosis factor-related apoptosis-inducing ligand, and airway inflammation of respiratory disorders such as asthma and cystic fibrosis. The Ets domain proteins have been proposed to bind to purine-rich central core 5'-GGAA-3' sequence with larger affinity than 5'-GGAT-3'. Here we describe surface plasmon resonance equilibrium and kinetics analyses of Ets domain ESE-3 binding to *Drosophila* E74 gene promoter sequence (a consensus Ets binding site) containing either 5'-GGAA-3' or 5'-GGAT-3' core sequence. Competition assays are also described. The results showed that there is no large difference in equilibrium binding affinity between Ets domain ESE-3 and either 5'-GGAA-3' or 5'-GGAT-3'. However, analysis of kinetic data indicates that larger association rate and lower dissociation rate constant values account for the slightly higher affinity of Ets domain ESE-3 for 5'-GGAA-3'-containing DNA sequence. The role of dynamics in protein-DNA recognition is discussed and a mechanism is proposed in which free Ets domain ESE-3 exists as two isomers in solution that undergoes a slow protein isomerization step, which is followed by a fast DNA binding process.

## Introduction

Members of the Ets-domain family of eukaryotic transcription factors were originally identified on the basis of a region of primary sequence homology with the protein product of the *v-ets* oncogene encoded by the E26 (E twenty six) avian erythroblastosis virus (1). There are 25 human and 26 murine Ets family members (2). The Ets-domain is a conserved region of approximately 85 amino acids that corresponds to the DNA-binding domain of these proteins. Ets transcription factors are divided into several subfamilies based on homology within the Ets domain, in which the Ets domain may be located at the C-terminal or N-terminal end. Moreover, based on observed organization of secondary structural elements, Ets proteins are proposed to be part of a superfamily of winged helix-turn-helix (wHTH) DNA-binding proteins consisting of three alpha-helices (H) and four beta-strands (S), arranged in the order H1-S1-S2-H2-H3-S3-S4 (3). Helices 2 and 3 correspond to HTH, the DNA recognition region, and H3 is responsible for making major groove contacts with the central 5'-GGA(A/T)-3' motif, whilst additional DNA contacts are made by a loop between S3 and S4 (the wing) and a loop between H2 and H3. Residues flanking the central 5'-GGA(A/T)-3' motif also dictate whether a particular Ets-domain will bind the site and, accordingly, consensus sequences for a number of Ets-domain transcription factors have been proposed (4). Significant overlap occurs between these consensus sequences over a 9-bp region, which emphasizes the similarity in their DNA binding specificity. However, drawing conclusions only from consensus sequences may be misleading and the suitability of a site as a target for different Ets-domain proteins can only be ascertained by direct experimentation. The apparent DNA specificity appears to depend on the experimental approach (5). Multiplex analysis measures the contribution of a single base to complex stabilization without knowledge of the sequence context whereas SELEX experiments involve the isolation of individual high affinity sequences from a randomized pool of oligonucleotides (6). Based on the former approach, a mechanism has been put forward in which cooperative effects among neighboring bases flanking the central purine-rich 5'GGA(A/T)3' core sequence contribute to the formation of stable Ets-DNA complexes (6). In this indirect readout mechanism, the Ets domain recognizes a sequence-dependent structure that is either induced or already present in DNA. The indirect readout mechanism for Ets domain is supported by X-ray crystallography and NMR. This model is in contrast to a direct readout mechanism of DNA binding specificity based on SELEX results, in which protein residues recognize and interact with unique base pairs within an Ets-binding site.

The target genes for Ets-domain transcription factors include oncogenes, tumor suppressor genes, apoptosis-related genes, differentiation-related genes, angiogenesis-related genes, and invasion and metastasis-related genes (4, 7, 8). ESE-3 (Epithelium-specific Ets factor, family member 3) protein, a member of the Ets transcription factor family, presents a very restricted expression pattern, being specifically expressed in differentiated epithelial cells; this expression is lost in a number of epithelial carcinomas. ESE-3 could play a role in determination of the early differentiated fate in stratified squamous epithelia (9). More recently, ESE-3 has been shown to regulate expression of a receptor for tumor necrosis factor-related apoptosis-inducing ligand (10), and to play an important role in airway inflammation of respiratory disorders such as asthma and cystic fibrosis (11). The ESE-3 protein has been shown to bind specifically to a canonical consensus Ets binding site, the E74 site (found in the *Drosophila* E74 gene promoter) containing a GGAA core (12).

Techniques used to identify DNA-binding sites and analyze protein-DNA interactions include, to name a few, electrophoretic mobility shift assays in direct and cross-competition experiments (13), DNA-protein photo-cross-linking (14), DNA-protein co-precipitation assays (15), DNA footprinting (14), fluorescence spectroscopy (14), and atomic force microscopy (16). However, these methods present limitations such as lack of easily extractable quantitative information, inability to minimize effects of non-specific interactions, difficulty in measuring weak interactions, perturbation in DNA binding caused by covalently

attached probes, and failure in many cases to detect cooperative binding. More recently, surface plasmon resonance (SPR) spectroscopy has allowed a better characterization of protein-DNA interactions, through the study of binding affinity, kinetics, sequence specificity, and binding stoichiometry (17,18). The Ets domain has been used to study target sequence specificity because it can be produced as a stable protein fragment and is sufficient for DNA binding, being a structural and functional domain (19).

The binding of proteins to regulatory sequences plays a crucial role in the regulation and control of gene expression. These proteins use diverse mechanisms to discriminate among a relatively small set of DNA-binding motifs having a wide range of affinities. Kinetic analysis of biomolecular interactions reveals details of the dynamic features of molecular recognition process. Quantitative measurements of ESE-3-DNA bimolecular interaction should improve our understanding of the molecular mechanisms underlying biological processes in which ESE-3 protein is involved. Here we describe purification to homogeneity of recombinant Ets domain of ESE-3 protein, equilibrium and kinetics analyses of protein-DNA interaction by surface plasmon resonance, and competition assays. The role of dynamics in protein-DNA recognition is discussed and the difference in affinity accounted for. A mechanism is proposed in which free Ets domain ESE-3 exists as two isomers in solution, and a slow protein isomerization step is followed by a fast DNA binding process.



## Materials and Methods

*Amplification and cloning of the Ets domain ESE-3 coding sequence, and expression and purification of the recombinant protein.* The fragments corresponding to the Ets domain ESE-3 coding sequence were obtained using the polymerase chain reaction (PCR) method presented elsewhere (20). Briefly, the 297 bp nucleotide sequence of ESE-3 Ets domain (Accession number NM\_012153), was constructed by dividing it into 8 sequences of approximately 46 bp each, where the oligonucleotides contained overlapping regions of approximately 10 bp at their 5'- and 3'-ends. Two other primers, PN and PH, were designed containing restriction sites for *NdeI* and *HindIII*, respectively. Pairs of oligonucleotides were assembled and PCR-amplified to yield the fragments of interest. The primers employed in the present study are summarized in Table 1. The DNA fragment was amplified using *Pfu* DNA Polymerase (Stratagene), ligated into the pCR<sup>®</sup>-Blunt cloning vector (Invitrogen), and transformed into *Escherichia coli* DH10B electrocompetent host cells. Purified plasmid DNA was digested with *NdeI* and *HindIII*, and the agarose-gel purified insert ligated into pET-23a(+) expression vector (Novagen), which was previously digested with the same restriction enzymes. DNA sequencing of recombinant plasmid confirmed that identity and integrity were maintained, and that no mutations were introduced by the PCR amplification step.

*E. coli* BL21 (DE3) (Novagen) electrocompetent host cells were transformed with the recombinant plasmid, grown in LB medium containing 50 µg mL<sup>-1</sup> carbenicillin at 37 °C to an OD<sub>600nm</sub> of 0.4, and grown for an additional 3 hours with no IPTG induction. All subsequent steps were performed at 4 °C unless stated otherwise. Cells were harvested by centrifugation at 11800g for 30 min, resuspended in 100 mM potassium phosphate pH 7.0 (buffer A), and disrupted by addition of 0.2 mg mL<sup>-1</sup> lysozyme and sonication. The solution was centrifuged at 48000g for 30 min to remove cell debris. Streptomycin sulphate (1 % final concentration) was added to the soluble fraction to precipitate nucleic acids, centrifuged at 11800g for 30 min and the supernatant dialysed against buffer A, and centrifuged. The supernatant was loaded on an FPLC HiPrep SP XL column (GE Healthcare), pre-equilibrated with buffer A. The column was washed with 10 column volumes of buffer A, and the adsorbed material eluted with a linear gradient (0-100%) of 15 column volumes of 100 mM potassium phosphate 1.0 M KCl pH 7.0 (buffer B). The fractions containing the target protein were pooled, concentrated using an Amicon ultrafiltration membrane (molecular weight cut-off of 3,000 Da), and loaded on a Sephacryl S-200 column (GE Healthcare), pre-equilibrated with buffer A. The column was eluted with buffer A, and the fractions containing homogeneous Ets-domain ESE-3 were immediately frozen in liquid nitrogen and stored at -80 °C. Protein content of fractions was analyzed for all purification steps by 12% SDS-PAGE, and concentrations determined by Bradford's method (21) using Bio-Rad protein assay kit (Bio-Rad) and bovine serum albumin as standard. The biophysical characterization of the recombinant ESE-3 Ets domain (11,993 Da) was performed by mass spectrometry analysis in a MALDI-TOF (M@ldi-LR mass spectrometer, Micromass) and N-terminal sequence analysis by Edman degradation as described elsewhere (22).

*Oligonucleotides and sensor chip.* The oligonucleotides for competition and/or direct binding studies were from Invitrogen. DNA sequence was designed based on the consensus Ets binding site of *Drosophila* E74 described by Bredemeier-Ernst *et al* (23). This sequence was Biotin-5'-AGCTTCTCTAGCTGAATAACCGGAAGTAACTCATCG-3'. The other oligonucleotides were obtained by changing the 5'-GGAA-3' core sequence to 5'-GGAT-3', 5'-GGAG-3', 5'-GGTT-3', and 5'-TTAA-3' of the consensus sequence. Only one strand of the double-stranded DNA containing either the 5'-GGAA-3' or 5'-GGAT-3' sequence was biotinylated at the 5' end. Annealing was carried out with 10-fold molar excess of the non-biotinylated oligonucleotides to ensure that all biotinylated DNA oligonucleotides were in the form of double-stranded DNA. The sequences 5'-GGAG-3', 5'-GGAT-3', 5'-GGTT-3', and 5'-TTAA-3' that were used in competitions assays were not biotinylated and annealing was

carried out at 1:1 stoichiometry. Complementary DNA molecules were dissolved in water and after raising the temperature to 80 °C, annealing was achieved by slowly cooling down the solutions to 4 °C. The solutions containing the double-stranded sequences were prepared so that the final solution contained HBS-EP (100 mM HEPES 1.5 M NaCl 30 mM EDTA 0.5% v/v Surfactant P20 pH 7.4). The sensor chip SA with streptavidin immobilized was from GE Healthcare.

*Preparation of the sensor surface.* Immobilizations of the oligonucleotide to the sensor chip were carried out at 25 °C in HBS-EP running buffer, at 5  $\mu\text{L min}^{-1}$  flow rate. Non conjugated streptavidin was removed by three consecutive injections (10  $\mu\text{l}$ ) of conditioning fluid (1.0 M NaCl 50 mM NaOH). For immobilizations of the sequence 5'-GGAA-3', which were used for both direct (kinetics) and competition assays, dsDNA solutions were injected onto the sensor chip, giving changes of about 113 RU and 35 RU, respectively. The sequence 5'-GGAT-3', which was used for direct assays, reached changes of about 118 RU. The DNA surface concentrations may be estimated by the following approximations: 1000 RU of DNA is equivalent to 0.78 ng of DNA molecules bound at the surface (24).

*Assay procedures.* The BIAcore X equipment (GE Healthcare) was employed to study the binding process to immobilized DNA sequences. This equipment relies on Surface Plasmon Resonance (SPR) to monitor the binding process, leading to refractive index changes close to the gold surface, which results in a change in the angle at which light is absorbed that is closely correlated to the change in mass at the gold surface upon protein binding (25). Before carrying out the assays, the influence of mass transport limitation was tested by varying the flow rate at a fixed concentration of Ets domain ESE-3 (1  $\mu\text{M}$ ). We found no limited mass transfer at a fast flow rate of 50  $\mu\text{L min}^{-1}$  at the low concentration of immobilized DNA. For all assays the titration of dilution series was performed by injection of solutions, in duplicate, at flow rate of 50  $\mu\text{L min}^{-1}$  over the flow cells. In the direct assays (kinetics) and competition assays, the Ets domain of ESE-3 was diluted in HBS-EP running buffer. Kinetics and competition assays were both carried out at 25 °C. The competition assay was performed under similar experimental conditions to direct assays, except that the DNA binding domain of ESE-3 solutions at fixed concentration of 800 nM were pre-incubated with 1-, 5-, 10-, 20-, 30-, or 45-fold excesses of competitor DNAs for 30 min at room temperature previous to injection. At the end of injection, the protein solutions were replaced by running buffer at the same flow rate with a delay of 300 s or 100 s for, respectively, kinetics or competition assays. The chip surface was regenerated with 1-min injection of 2.0 M NaCl. One flow cell was used as reference surface to remove noise artifacts. The signal noise can result from bulk refractive index changes, matrix effects, nonspecific binding, injection noise, and baseline drift (25). Since the signal resulting from these nonspecific events is equally present in both flow cells, subtraction of the signal from an uncoated flow cell leads to refinement of the data.

*SPR determination of equilibrium constant for binary complex formation between Ets domain ESE-3 and 5'-GGAA-3' or 5'-GGAT-3'.* The data of the SPR titration to immobilized 5'-GGAA-3' containing oligonucleotide (113 RU) were fitted to Equation (1), where  $R_{eq_i}$  is the equilibrium resonance response unit for  $i$  concentrations of Ets domain ESE-3 ( $i = 0, 0.06, 0.08, 0.4, 0.6, 0.8, 1.0, 1.44, 2.0, 3.0, 4.0, \text{ and } 6.0 \mu\text{M}$ ),  $R_{\text{max}}$  is the resonance response unit at saturating Ets domain ESE-3 concentrations, and  $K'$  is the overall dissociation constant at equilibrium. The  $K'$  value for Ets domain ESE-3 binding 5'-GGAT-3' containing immobilized sequence (118 RU) on the sensor surface was obtained by the same manner, however at different  $i$  concentrations of Ets domain ESE-3 ( $i = 0.0, 0.2, 0.4, 0.6, 0.8, 1.0, 2.0, 4.0 \text{ and } 6.0 \mu\text{M}$ ). These binding parameters and their standard errors were obtained by the

average data of duplicate experiments, using the nonlinear regression function on SigmaPlot 2000 (SPSS, Inc.).

$$Re q_i = \frac{R_{\max} [ESE - 3]}{K' + [ESE - 3]} \quad (1)$$

Estimates of free energy change can also be obtained by substituting  $K'$  values into  $\Delta G = -RT \ln(1/K')$  equation.  $T$  is the temperature in Kelvin ( $273.15 \text{ K} = 0 \text{ }^\circ\text{C}$ ) and  $R$  is the universal gas constant ( $1.987 \text{ cal/mol K}$ ).

The stoichiometry for complex formation between Ets domain ESE-3 and Ets binding site of *Drosophila* E74 DNA containing 5'-GGAA-3' or 5'-GGAT-3' sequence was determined from fitting the data to Equation (2) (26), where  $n$  is the number of Ets domain ESE-3 molecules bound to DNA or the number of protein binding sites on DNA (assuming that a single protein molecule binds a single binding site),  $R_{\max}$  is the resonance response for saturating concentration of protein,  $R_{\text{DNA}}$  is the amount of immobilized DNA (RU),  $MW_{\text{DNA}}$  is the molecular mass of DNA (Ets binding site of *Drosophila* E74 DNA = 22,278 Da) and  $MW_{\text{ESE-3}}$  is the molecular mass of Ets domain ESE-3 (11,993 Da).

$$n = \frac{R_{\max} MW_{\text{DNA}}}{R_{\text{DNA}} MW_{\text{ESE-3}}} \quad (2)$$

*Kinetic data analysis: determination of rate constants from sensorgrams.* The apparent rate constants were determined from sensorgrams in both association and dissociation phases. The apparent association and dissociation rate constants,  $k_{\text{obs}}$  and  $k_{-1}$ , respectively, were obtained by nonlinear fitting of the sensorgram data using the BIAevaluation software version 4. The data from dissociation phase were fitted to Equation (3), where  $R_t$  is the response at time  $t$  and  $R_{t_0}$  is the amplitude of the response at the end of the injection, and  $R_{(t \rightarrow \infty)}$  is the response value time tending to infinite and represents complete dissociation of the Ets domain ESE-3-DNA complex. The apparent association rate constant,  $k_{\text{obs}}$ , can be derived from the measured  $k_{\text{off}}$  values through Equation (4), where  $[ESE-3]$  is the concentration of the analyte,  $R_{\max}$  represents the maximum analyte binding capacity, and  $RI$  is the signal at the point of analyte injection ( $t = 0$ ) (27). After final injection of analyte, when protein solutions were replaced by running buffer, there was a 30 s lasting period, which the response reflects changes in the bulk refractive index; this period was not considered in data analysis, as well the association data at 5 to 10 s after the injection point (27).

$$R_t = R_{t_0} \left( e^{-k_{-1}(t-t_0)} \right) + R_{(t \rightarrow \infty)} \quad (3)$$

$$R_t = \frac{k_{\text{obs}} [ESE - 3]}{k_{\text{obs}} [ESE - 3] + k_{-1}} R_{\max} \left( 1 - e^{-(k_{\text{obs}} [ESE - 3] + k_{-1})(t-t_0)} \right) + RI \quad (4)$$

The plot of  $k_{\text{obs}}$  values against varying  $[ESE-3]$  were fitted to an equation describing a hyperbolic decay, Equation (5), in which  $K_d$  is the intrinsic dissociation constant, and  $k_2$  and  $k_{-2}$  are, respectively, limiting forward and reverse first-order rate constants for an isomerization process between two forms of free enzyme that must occur before substrate binding (28).

$$k_{\text{obs}} = \frac{k_{-2} K_d}{[ESE - 3] + K_d} + k_2 \quad (5)$$

*Competition analysis.* The competition assay analysis was carried out as described elsewhere (18). Briefly, the level of Ets domain ESE-3 binding to immobilized Ets binding site of *Drosophila* E74 consensus DNA sequence was set at 100 %. The competition at a certain concentration of non-biotinylated competitor DNA was calculated with the relation percentage of binding value (% binding) as follows (Equation 6):

$$\%binding = \left( \frac{R_i}{R_0} \right) \times 100 \quad (6)$$

Where  $R_i$  is the signal obtained at a certain concentration of competitor, and  $R_0$  is the signal obtained in the absence of competitor DNA.

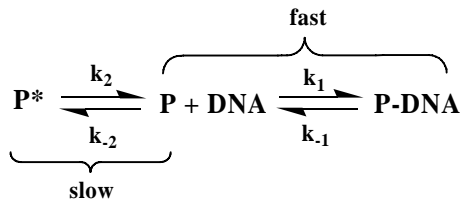
## Results

*Cloning, expression, purification, and characterization of the Ets domain ESE-3.* The PCR-amplified fragment corresponding to the Ets domain of ESE-3 coding sequence was cloned into pET-23a(+), and its identity was confirmed by sequence analysis. Recombinant protein expression in the soluble form in *E. coli* BL21(DE3) host cells was obtained after 3 h of cell growth of cell cultures reaching an OD<sub>600nm</sub> value of 0.4, with no IPTG induction. The two-step recombinant protein purification protocol (cation exchange and gel filtration) yielded homogeneous Ets domain of ESE-3. A single band of expected molecular mass was observed by SDS-PAGE stained with silver nitrate (data not shown). A total of 23 mg of homogeneous recombinant protein could be obtained from 10 g of cell paste. N-terminal amino acid sequence analysis confirmed the identity of the purified protein as Ets domain of ESE-3, and MALDI-TOF analysis demonstrated that the recombinant protein has a molecular mass of 11,993 Da (predicted molecular mass = 12,003.70 Da).

*Test of the Ets domain ESE-3 mass transport-limited binding.* To test whether mass transport limitation affected the binding process, binding rates were measured at increasing varying flow rates. The results were as follows:  $2.57 \times 10^{+5} \text{ M}^{-1} \text{ s}^{-1}$  at  $5 \mu\text{L min}^{-1}$ ,  $6.77 \times 10^{+5} \text{ M}^{-1} \text{ s}^{-1}$  at  $10 \mu\text{L min}^{-1}$ ,  $6.97 \times 10^{+5} \text{ M}^{-1} \text{ s}^{-1}$  at  $30 \mu\text{L min}^{-1}$ ,  $6.07 \times 10^{+5} \text{ M}^{-1} \text{ s}^{-1}$  at  $40 \mu\text{L min}^{-1}$ ,  $5.27 \times 10^{+5} \text{ M}^{-1} \text{ s}^{-1}$  at  $50 \mu\text{L min}^{-1}$ , and  $3.57 \times 10^{+5} \text{ M}^{-1} \text{ s}^{-1}$  at  $60 \mu\text{L min}^{-1}$  for Ets domain ESE-3 ( $1 \mu\text{M}$ ) binding to immobilized 5'-GGAA-3'-containing DNA sequence (113 RU). The association rates were thus fairly independent of the flow rate for all immobilized chips, which demonstrates the absence of mass transfer effects. At very high surface concentration, which is assumed to be proportional to the saturation response  $R_{\text{max}}$ , the binding rates will be limited by mass transfer and independent of the association rate constants (29). Since Ets domain ESE-3 has a molecular mass value of 11,993 Da and the  $R_{\text{max}}$  values were up to  $\sim 300$  RU, it could be predicted that mass transport limitation would occur only to association rate constants larger than  $10^{+5} \text{ M}^{-1} \text{ s}^{-1}$  (29). As the values for association rate constants of Ets domain ESE-3 interaction with DNAs were up to  $10^{+3}$  (Table 2), mass transport-limited binding would be expected only for  $R_{\text{max}}$  values larger than 10,000 RU (29). Therefore, mass transport limitation was not an issue for the bimolecular interaction processes between Ets domain ESE-3 and immobilized DNAs presented here. These data are an important consideration, since the presence of mass transport limitation must be considered in the kinetics analysis.

*Equilibrium dissociation constants of Ets domain ESE-3 binding to either 5'-GGAA-3' or 5'-GGAT-3' containing oligonucleotide by direct assay.* Equilibrium dissociation constant ( $K'$ ) values were determined to evaluate the influence of the fourth base (A or T) on affinity of Ets domain ESE-3 for GGA core consensus sequence of immobilized *Drosophila* E74 (a global ligand for all Ets proteins). The hyperbolic increase in equilibrium resonance responses ( $\text{Req}_i$ ) upon binding of increasing concentrations of Ets domain ESE-3 to either 5'-GGAA-3' (Fig 1A) or 5'-GGAT-3' containing (Fig 1B) *Drosophila* E74 sequence were plotted to Equation (1). This analysis yielded values for equilibrium dissociation constant ( $K'$ ) of  $0.4 \pm 0.1 \mu\text{M}$  and  $R_{\text{max}}$  of  $219 \pm 19$  RU for 5'-GGAA-3', and  $K'$  of  $0.8 \pm 0.1 \mu\text{M}$  and  $R_{\text{max}}$  of  $312 \pm 16$  RU for 5'-GGAT-3' (Table 3). Since there is only one site on Ets Binding Site (EBS) of *Drosophila* E74 gene for Ets domain ESE-3 binding, a 1:1 stoichiometry was expected with a  $R_{\text{max}}$  value of 61 RU. However, the  $R_{\text{max}}$  values suggest a protein:DNA stoichiometry of 4:1 ( $219\text{RU} \times 22,278\text{Da} / 113\text{RU} \times 11,993\text{Da}$ ) for 5'-GGAT-3' and 5:1 ( $312\text{RU} \times 22,278\text{Da} / 118\text{RU} \times 11,993\text{Da}$ ) for 5'-GGAT-3'. The larger  $K'$  value demonstrate that binding of Ets domain ESE-3 to GGAT-containing *Drosophila* E74 sequence is two-fold weaker than to GGAA-containing DNA.

*Kinetics of Ets domain ESE-3 binding.* The sensorgrams for binding of increasing concentrations of Ets domain ESE-3 to immobilized 5'-GGAA-3' (Fig. 2A) and 5'-GGAT-3' (Fig. 2B) DNAs allowed to determine values for both apparent association and dissociation rate constants. Data from association phases (Fig. 2, insets present the sensorgrams) were best fitted to a single exponential equation [Equation (4)], yielding values for apparent association rate constants ( $k_{obs}$ ). Plots of  $k_{obs}$  against increasing Ets domain ESE-3 concentrations showed a hyperbolic decrease (Fig. 2A and B), which are consistent with a mechanism in which there are two forms of free protein (Ets domain ESE-3) in solution followed by a bimolecular binding to only one form of free protein (28,30). This mechanism predicts that the free protein undergoes a slow isomerization reaction followed by a fast bimolecular process. Accordingly, the data were fitted to an equation [Equation (5)] that describes a system in which P and P\* represent two isomers of free protein, and only P can bind to the immobilized DNA (Scheme I). This analysis yielded values of  $K_d$ , the intrinsic dissociation constant of Ets domain ESE-3-DNA binary complex formation, and the forward ( $k_2$ ) and reverse ( $k_{-2}$ ) rate constants (Table 2) for the isomerization of P (Ets domain ESE-3 that binds EBS of *Drosophila* E74 gene) and P\* (Ets domain ESE-3 that does not bind EBS of *Drosophila* E74 gene).



Scheme I

In order to inspect the data for internal consistency, we assessed the overall dissociation constant,  $K_{D(overall)}$ , for Ets-domain ESE-3-DNA binding mechanism, which is given by the following relationship in Equation (7):

$$K_{D(overall)} = K_d \left( \frac{1}{K_2} + 1 \right) \quad (7)$$

In this equation,  $K_2$  represents the equilibrium isomerization constant for the conversion of P\* to P of Ets domain ESE-3 ( $K_2 = [P]/[P^*] = k_2/k_{-2}$ ), and  $K_d$  is the equilibrium dissociation constant for the Ets domain ESE-3-DNA binary complex defined by  $K_d = k_{-1}/k_1$ . The values  $k_2$ ,  $k_{-2}$ , and  $K_d$  (Table 2) given by Equation (5) allow to calculate  $K_{D(overall)}$  values of  $2 (\pm 1) \mu\text{M}$  for immobilized 5'-GGAA-3' and  $10 (\pm 5) \mu\text{M}$  for 5'-GGAT-3' sequence. These values are larger than the ones derived from equilibrium SPR data ( $K'$ ), which were  $0.4 \mu\text{M}$  for 5'-GGAA-3' and  $0.8 \mu\text{M}$  for 5'-GGAT-3' (Table 2). To reconcile these discrepancies, a conformational change of Ets domain ESE-3-DNA binary complex should be invoked in the proposed mechanism (Scheme I). However, further experimental evidence is needed to confirm or refute this proposal since no SPR signal would result from conformational changes of the binary complexes on the Sensor Chip. Notwithstanding, Gibbs free energy change ( $\Delta G$ ) can be obtained by substituting  $K'$  values into  $\Delta G = -RT \ln(1/K')$  equation, in which T is the temperature in Kelvin ( $298.15 \text{ K} = 25 \text{ }^\circ\text{C}$ ) and R is the universal gas constant ( $1.987 \text{ cal/mol K}$ ). This analysis (Table 2) of the equilibrium SPR data indicates that the energetic difference between Ets domain ESE-3 recognition of 5'-GGAA-3' and 5'-GGAT-3' is only  $0.4 \text{ kcal mol}^{-1}$  more favorable for the former interaction. If, on the other hand, the values of the dissociation constant derived from kinetic data analysis ( $K_d$ ) were employed, the Gibbs free energy changes would be  $-9.7 \text{ kcal mol}^{-1}$  for 5'-GGAA-3' and  $-8.1 \text{ kcal mol}^{-1}$  for 5'-GGAT-3', suggesting that binding to the former is  $1.6 \text{ kcal mol}^{-1}$  more favorable. The energies of hydrogen bonds have been variously estimated to be between  $2.9$  and  $9.1 \text{ kcal mol}^{-1}$  (31).

Accordingly, disruption or formation of a single hydrogen bond would suffice to account for the difference in Gibbs free energy change in this biomolecular interaction.

The dissociation rate constant value for dissociation of DNA from the Ets domain ESE-3-DNA binary complex ( $k_{-1}$ ) calculated from SPR kinetic data fitted to Equation (3) was  $0.61 \times 10^{-3} \text{ s}^{-1}$  for 5'-GGAA-3' and  $1.8 \times 10^{-3} \text{ s}^{-1}$  for 5'-GGAT-3' (Table 2). These data suggest that DNA containing the 5'-GGAT-3' sequence dissociates faster from Ets domain ESE-3 protein. Estimates for the "true" association rate constant ( $k_1$ ) can be calculated from the dissociation constant value ( $K_d$ ) estimated from the apparent association rate constant ( $k_{\text{obs}}$ ) versus [ESE-3] plots (Fig. 2 and Scheme I) using the following relationship:  $K_d = k_{-1}/k_1$ . This analysis yielded values for the "true" association rate constants of  $8 \times 10^{+3} \text{ M}^{-1} \text{ s}^{-1}$  for 5'-GGAA-3' and  $1.7 \times 10^{+3} \text{ M}^{-1} \text{ s}^{-1}$  for 5'-GGAT-3' (Table 2). These data show that the rate of bimolecular interaction between Ets domain ESE-3 and 5'-GGAA-3' is faster than the recognition of 5'-GGAT-3'. Moreover, since transport limitation will likely affect binding kinetics to a dextran-based sensor surface if association rate constant is larger than  $10^{+5} \text{ M}^{-1} \text{ s}^{-1}$  (32), the values here reported for the "true" association rate constants for either 5'-GGAA-3' or 5'-GGAT-3' are not transport limited. The larger value for the association rate constant and the lower dissociation rate constant for Ets domain ESE-3 binding to 5'-GGAA-3'-containing DNA sequence can account for the higher affinity of the latter as compare to 5'-GGAT-3' interaction.

*Competitions assays by SPR.* Competition assays were carried out to investigate the influence of substitution of the fourth base in 5'-GGAX-3' core sequence (X=A, T, or G) on Ets domain ESE-3-DNA binary complex formation. This competition assay approach provides the assessment to relative affinity of a large number of DNA sequences with one surface preparation (18). Accordingly, a new SA Sensor Chip having immobilized Ets binding site of *Drosophila* E74 containing the 5'-GGAA-3' core DNA sequence (35 RU) was employed in the competition assays. Considering that 0.78 ng of a DNA molecule bound at the surface gives a response of 1000 RU (24), we can expect  $\sim 0.0273 \text{ ng}$  or  $6.3 \times 10^{-16} \text{ moles}$  of DNA ( $22,278 \text{ g mol}^{-1}$ ) distributed within the dextran layer and, considering a flow cell total volume of  $\sim 0.06 \mu\text{L}$ , a concentration of 20 nM of immobilized DNA. The non-biotinylated competitor DNAs were Ets binding site of *Drosophila* E74 containing either 5'-GGAT-3', 5'-GGAG-3', 5'-GGTT-3', or 5'-TTAA-3' sequences. The 5'-GGTT-3' and 5'-TTAA-3' core sequences were included as negative controls. Ets domain ESE-3 protein was fixed at 800 nM and was incubated with 1, 5, 10, 20, 30, 45-fold excess of non-biotinylated competitor DNAs. The extent of binding of Ets site of ESE-3 to immobilized 5'-GGAA-3' depends on the affinity of competitor DNA for this protein in solution: the larger the affinity of competitor DNA, the lower concentration of free protein in solution to bind immobilized DNA. A plot of percentage of binding value (% binding) calculated by Equation (6) against concentration of competitor DNAs (Fig. 3) yielded estimates for  $\text{IC}_{50}$  values, which were as follows:  $\sim 190, 300, 350, 510,$  and  $650 \text{ nM}$  for, respectively, GGAA, GGAT, GGAG, GGTT, TTAA containing sequences. These data suggest the following Ets domain ESE-3 affinity order:  $\text{GGAA} > \text{GGAT} \geq \text{GGAG} > \text{GGTT} \geq \text{TTAA}$ . These results are in agreement with the consistently larger dissociation constant values for 5'-GGAT-3' as compared to 5'-GGAA-3' determined from equilibrium and kinetics of binding (Table 2).

## Discussion

The seminal work by Seeman et al. (33) concluded that interactions involving two hydrogen bonds would be required to uniquely distinguish each base-pair from the other, and inferred that two hydrogen bonds from a single functional group would specify a site with higher precision than from two independent groups (analogous to the “chelate effect” in which formation of one bond favours formation of additional bonds by an increase in effective concentration). It has been pointed out that even though no simple “recognition code” exists for protein-DNA complexes, some common interaction patterns between amino acids and nucleic acid bases can be found in a given structural context such as helix-turn-helix motif (34). The guanidinium group of Arg donates two hydrogen bonds to the O6 and N7 acceptor groups of guanine and a second in which the carboxamide group of Asn (or Gln) hydrogen bonds to the N7 acceptor and N6 donor groups of adenine (33,35). The other frequent bidentate interaction utilizes the amino group of Lys to hydrogen bond to the O6 and N7 acceptors of guanine (35). These residues are conserved in ESE-3 transcription factor (12). When hydrogen bonds, water-mediated bonds and van der Waals contacts were analysed in a wide range of protein-DNA complexes, it was found that two-thirds of all protein-DNA interactions involve van der Waals contacts as compared to one-sixth for hydrogen bonds and one-sixth for water-mediated bonds (35). Moreover, over two-thirds of contacts are made with the sugar-phosphate backbone of DNA, which do not directly depend on the underlying DNA sequence. Accordingly, it has been concluded that most protein-DNA interactions can be said to stabilise the complex or aid the indirect read-out of the bases through recognition of the DNA structure (35). As mentioned above, an indirect mechanism has been proposed for Ets-DNA complex formation in which Ets domain proteins recognize an induced or pre-existing sequence-dependent structure (6). Computer simulations have been used to analyze the effects of point mutations in the DNA core sequence on Ets domain and DNA binary complex formation, and the results prompted a proposal of a direct readout mechanism in the Ets-1 Ets domain-DNA complex formation (36). In addition, combinatorial interactions of Ets factors with other proteins, including other sequence specific transcription factors, result in context-dependent transcriptional regulation and define target gene specificity (2,37). The results presented here show that even though there is not a dramatic difference in equilibrium binding affinity between Ets domain ESE-3 and either 5'-GGAA-3' or 5'-GGAT-3', analysis of kinetic data indicate that larger association rate constant and lower dissociation rate constant values account for the slightly higher affinity of Ets domain ESE-3 for 5'-GGAA-3'-containing DNA sequence. These data would suggest that a direct read-out mechanism (33) appears to play, at least in part, a role in Ets domain ESE-3 protein-DNA recognition. Differences in dissociation rate constants for protein-DNA interactions have been reported (38,39). Analyses of specific and nonspecific binding have indicated that the mechanisms of protein-DNA interactions could involve a number of steps (40). Firstly, long-range electrostatic forces bring the protein and the DNA into proximity and the counter-ions on free DNA are released into the solvent. The nonspecific binding ends this step, which is predominantly driven by the favourable increase in entropy. The weak binding and fast dissociation rates would permit that the proteins slides along the DNA. As the protein recognizes its consensus base sequence, specific hydrogen bonds and van der Waals contacts are formed, and high-affinity binding is achieved with slow dissociation rates. This is in agreement with the lower dissociation rate constant value determined for Ets domain ESE-3 protein interaction with 5'-GGAA-3'-containing DNA sequence. However, it should be pointed out that neighbouring sequences have also been shown to play a role in promoter recognition by ESE-3 (10). At any rate, our results demonstrate that dynamic features should also be taken into consideration when trying to understand protein-DNA interactions. Accordingly, it has been pointed out that in some cases, such as in dynamic regulation of signal transduction processes, kinetic control could have a more important role than affinity control (32). Interpretation at molecular level will have to await three-dimensional structure



determination to provide further insight on how specific structural features contribute to kinetics of Ets domain ESE-3 and DNA bimolecular recognition.

## **Funding**

This work was supported by National Institute of Science and Technology on Tuberculosis (Decit/SCTIE/MS-MCT-CNPq-FNDCT-CAPES) and Millennium Initiative Program (CNPq), Brazil, to DSS and LAB. DSS and LAB also acknowledge grants awarded by FINEP. DSS (304051/1975-06) and LAB (520182/99-5) are research career awardees from the National Council for Scientific and Technological Development of Brazil (CNPq).

## References

1. Karim,F.D., Urness,L.D., Thummel,C.S., Klemsz,M.J., McKercher,S.R., Celada,A, Van Beueren,C., Maki,R.A., Gunther,C.V. and Nye,J.A., *et al.* (1990) The ETS-domain: a new DNA-finding motif that recognizes a purine-rich core DNA sequence. *Genes Dev.*, 4, 1451-1453.
2. Hsu,T., Trojanowska,M. and Watson,D.K. (2004) Ets proteins in biological control and cancer. *J. Cell. Biochem.*, 91, 896-903.
3. Donaldson,L.W., Petersen,J.M., Graves,B.J. and McIntosh,L.P. (1994) Secondary structure of the ETS domain places murine Ets-1 in the superfamily of winged helix-turn-helix DNA-binding proteins. *Biochemistry*, 33, 13509-16.
4. Sharrocks,A.D. (2001) The Ets-domain transcription factor family. *Nature Reviews*, 2, 827-837.
5. Caceres,R.A., Nunes,C.P., Azevedo,W.F.Jr., Basso,L.A. and Santos,D.S. (2007) Molecular dynamics and structural studies of the Ets domain-DNA complexes. *Current Bioinformatics*, 2, 222-228.
6. Szymczyna,B.R. and Arrowsmith,C.H. (2000) DNA binding specificity studies of four ETS proteins support an indirect read-out mechanism of protein-DNA recognition. *J. Biol. Chem.*, 275, 28363-28370.
7. Oikawa,T. and Yamada,T. (2003) Molecular biology of the Ets family of transcription factors. *Gene*, 303, 11-24.
8. Oikawa,T. (2004). ETS transcription factors: Possible targets for cancer therapy. *Cancer Sci.*, 95, 626-633.
9. Tugores,A., Le,J., Sorokina,I., Snijders,A.J., Duyao,M., Reddy,P.S., Carlée,L., Ronshaugen,M., Mushegian,A., Watanaskel,T., Chu,S., Buckler,A., Emtage,S. and McCormick, M.K. (2001) The epithelium-specific Ets protein EHF/ESE-3 is a context-dependent transcriptional repressor downstream of MAPK signaling cascades. *J. Biol. Chem.*, 276, 20397-20406.
10. Lim,H.J., Cho,J.Y., Park,Y.B., Park,J.W. and Kwon,T.K. (2006) ESE-3 transcription factor is involved in the expression of death receptor (DR-5) through putative Ets sites. *Biochem. Biophys. Res. Commun.*, 350, 736-741.
11. Wu,J., Duan,R., Cao,H., Field,D., Newnham,C.M., Koehler,D.R., Zamel,N., Pritchard,M.A., Hertzog,P., Post,M., Tanswell,A.K. and Hu,J. (2008) Regulation of epithelium-specific Ets-like factors ESE-1 and ESE-3 in airway epithelial cells: potential roles in airway inflammation. *Cell Res.*, 18, 649-663.
12. Kas,K., Finger,E, Grall,F., Gu,X., Akbarali,Y., Boltax,J., Weiss,A., Oettgen,P., Kapeller,R., Libermann,T. (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.
13. Brenowitz,M., Seneor,D.F., Shea,M.A. and Acker,G.K. (1986) Quantitative DNAase footprint titration: a method for studying protein-DNA interactions. *Methods Enzymol.*, 130, 132-181.
14. Guille,M.J. and Kneale,G.G. (1997) Methods for the Analysis of DNA-Protein Interactions. *Mol. Biotechnol.*, 8, 35-52.
15. Franza,B.R., Josephs,S.F., Gilman,M.Z., Ryan,W. and Clarkson,B. (1987) Characterization of cellular proteins recognizing the HIV enhancer using a microscale DNA-affinity precipitation assay. *Nature*, 330, 391-395.
16. Yokota,H., Fung,K., Trask,B.J., van den Engh,G., Sarikaya,M. and Aebersold,R. (1999) Sharp DNA bends as landmarks of protein-binding sites on straightened DNA. *Anal. Chem.*, 71, 1663-1667.
17. Brockman,J.M, Frutos,A.G. and Corn,R.M. (1999) A multistep chemical modification procedure to create DNA arrays on gold surfaces for the study of protein-DNA

- interactions with surface plasmon resonance imaging. *J. Am. Chem. Soc.*, 121, 8044-8051.
18. Teh,H.F., Peh,W.Y.X., Su,X. and Thomsen,J.S. (2007) Characterization of Protein-DNA Interactions Using Surface Plasmon Resonance Spectroscopy with Various Assay Schemes. *Biochemistry*, 46, 2127-2135.
  19. Graves,B.J. and Petersen,J.M. (1998) Specificity within the ets family of transcription factors. *Adv. Cancer Res.*, 75, 1-55.
  20. Santos,D.S., Chies,J.M., Basso,L.A., Renard,G. and Fonseca,I.O. (2006) Method for the obtention of chimeric nucleotide sequences and chimeric nucleodite sequences. Patent application WO2007/068068 A2.
  21. Bradford,M.M., McRorie,R.A. and Williams,W.L. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.*, 72, 248-254.
  22. Azzoni,A.R., Tada,S.F., Rosselli,L.K., Paula,D.P., Catani,C.F., Sabino,A.A., Barbosa,J.A., Guimarães,B.G., Eberlin,M.N., Medrano,F.J. and Souza,A.P. (2004) Expression and purification of a small heat shock protein from the plant pathogen *Xylella fastidiosa*. *Protein Expr. Purif.*, 33, 297-303.
  23. Bredemeier-Ernst,I., Nordheim,A. and Jaknecht,R. (1997) Transcriptional activity and constitutive nuclear localization of the ETS protein Elf-1. *FEBS Lett.*, 12, 47-51.
  24. Bouffartigues,E., Leh,H., Anger-Leroy,M., Rimsky,S. and Buckle,M. (2007) Rapid coupling of surface plasmon resonance (SPR and SPRi) and proteinChip based mass spectrometry for the identification of proteins in nucleoprotein interactions. *Nucleic Acids Res.*, 35, e39.
  25. Katsamba,P.S., Park,S., and Laird-Offringa,I.A. (2002). Kinetic studies of RNA-protein interactions using surface plasmon resonance. *Methods*, 26, 95-104.
  26. Majka,J. and Speck,C. (2007) Analysis of protein-DNA interactions using surface plasmon resonance. *Adv Biochem. Engin. Biotechnol.*, 104, 13-36.
  27. O'Shannessy,J.D., Brigham-Burke,M., Soneson,K.K., Hensley,P. and Brooks,I. (1993) Determination of Rate and Equilibrium Binding Constants for Macromolecular Interactions Using Surface Plasmon Resonance: Use of Nonlinear Least Squares Analysis Methods. *Analytical Biochemistry*, 212, 457-468.
  28. Basso,L.A., Engel,P.C. and Walmsley,A.R. (1995) The mechanism of substrate and coenzyme binding to clostridial glutamate dehydrogenase during reductive amination. *Eur. J. Biochem.*, 234, 603-615.
  29. Karlsson,R., Roos,H., Fågerstam,L. and Persson,B. (1994) Kinetics and Concentration Analysis Using BIA Technology. *Methods: A companion to Methods in Enzymology*, 6, 99-110.
  30. Nakatani,H. and Hiromi,K. (1980) Analysis of signal amplitude in stopped-flow method for enzyme-ligand systems. *J. Biochem.*, 87, 1805-1810.
  31. Fersht,A. (1999) *Structure and Mechanism in Protein Science*. W. H. Freeman and Company, NY.
  32. De Mol,N.J. and Fischer,M.J.E. (2008) Kinetic and thermodynamic analysis of ligand-receptor interactions: SPR applications in drug development. In Schasfoort,R.B.M. and Tudos,A.J. (ed.), *Handbook of Surface Plasmon Resonance*. The Royal Society of Chemistry Publishing, The Netherlands, pp. 123-172.
  33. Seeman,N.C., Rosenberg,J.M. and Rich,A. (1976) Sequence-specific recognition of double helical nucleic acids by proteins. *Prot. Natl. Sci. USA*, 73, 804-808.
  34. Cheng,A.C., Chen,W.W., Fuhrmann,C.N. and Frankel,A.D. (2003) Recognition of nucleic acid bases and base-pairs by hydrogen bonding to amino acid side-chains. *J. Mol. Biol.*, 327, 781-796.

35. Luscombe,N.M., Laskowski,R.A. and Thornton,J.M. (2001) Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Res.* 29, 2860-2874.
36. Obika,S., Reddy,S.Y., Bruice,T.C. (2003) Sequence specific DNA binding of Ets-1 transcription factor: molecular dynamics study on the Ets domain-DNA complexes. *J. Mol. Biol.*, 331, 345-359.
37. Buchwalter,G., Gross,C., Wasylyk,B. (2004) Ets ternary complex transcription factors. *Gene*, 324, 1-14.
38. Hart,D.J., Speight,R.E., Cooper,M.A., Sutherland,J.D. and Blackburn,J.M. (1999) The salt dependence of DNA recognition by NF- $\kappa$ B p50: a detailed kinetic analysis of the effects on affinity and specificity.
39. Oda,M., Furukawa,K., Sarai,a. And Nakamura,H. (1999) Kinetic analysis of DNA binding by the c-Myb DNA-binding domain using surface plasmon resonance. *FEBS Lett.*, 454, 288-292.
40. Oda,M. and Nakamura,H. (2000) Thermodynamic and kinetic analyses for understanding sequence-specific DNA recognition. *Genes Cells*, 5, 319-326.

## Figure Legends

**Figure 1.** Equilibrium SPR titration. (A) Equilibrium resonance responses ( $Req_i$ ) plotted as a function of increasing Ets domain ESE-3 concentration upon binding to immobilized *Drosophila* E74 containing 5'-GGAA-3' core sequence. (B) Equilibrium resonance responses ( $Req_i$ ) plotted as a function of increasing Ets domain ESE-3 concentration upon binding to immobilized *Drosophila* E74 containing 5'-GGAT-3' core sequence.

**Figure 2.** Kinetics of Ets-domain ESE-3 to *Drosophila* E74 containing either 5'-GGAA-3'(A) or 5'-GGAT-3'(B) core sequence. Analysis of Ets domain ESE-3 concentration dependence of the  $k_{obs}$  values calculated from the slow phase by Equation (5). Insets give SPR sensorgrams at increasing concentrations of Ets domain ESE-3 after baseline correction.

**Figure 3.** Competition profile of the binding of Ets domain ESE-3 to immobilized *Drosophila* E74 containing 5'-GGAA-3' core sequence. Non-biotinylated competitor DNAs were at 1, 5, 10, 20, 30, 45-fold excesses over Ets domain ESE-3 protein concentration. The Ets domain ESE-3 binding signal in the absence of competitors is set to 100 %.

**Table 1.** The primers employed to obtain the fragments that encode the Ets domain ESE-3\*.

PN - 5' CCCATATGCACAACCCGAGAGGGACT 3'  
P1f - 5' CACAACCCGAGAGGGACTCACTTATGGGAATTCATCCGCGACATCC 3'  
P2r - 5' CATTTTATTAATCCTGGGTTCTTGTCTGGGTTCAAGAGGATGTCGCG 3'  
P3f - 5' TAATAAAATGGGAAGACCGATCTGAGGGCGTCTTCAGGTTCTTGAA 3'  
P4r - 5' TTCTTTTTACCCCATAGCTGAGCCACTGCCTCTGATTTC AAGAACC 3'  
P5f - 5' GTAAAAAGAAGAACAACAGCAGCATGACCTATGAAAAGCTCAGCCGA 3'  
P6r - 5' GTAAAAAGAAGAACAACAGCAGCATGACCTATGAAAAGCTCAGCCGA 3'  
P7f - 5' AAATTCTGGAGCGTGTGGATGGACGAAGACTGGTATATAAATTTGG 3'  
P8r - 5' TCAGTTTTCATTTTCTCTCCATCCTCGGGCATTCTTCCCAAATTTAT 3'  
PH -5' CTAAGCTTTCAGTTTTCATTTTCTCTC 3'

\*The oligonucleotide primer sequences for construction of the Ets domain ESE-3 coding sequence (5'-**atg** cacaaccga gagggactca cttatgggaa tcatccgcg acatcctctt gaaccagac aagaaccag gattaataaa atgggaagac cgatctgagg gcgtcttcag gttcttgaaa tcagaggcag tggetcagct atggggtaaa aagaagaaca acagcagcat gacctatgaa aagctcagcc gagctatgag atattactac aaaagagaaa ttctggagcg tgtggatgga cgaagactgg tatataaatt tgggaagaat gcccgaggat ggagagaaaa tgaaac **tga**-3'). The restriction sites, *Nde*I and *Hind*III, are underlined in primers PN and PH, respectively.

**Table 2.** Equilibrium parameters and kinetic constants estimated from SPR measurements of Ets domain ESE-3 binding to EBS *Drosophila* E74 gene containing 5'-GGAA-3' or 5'-GGAT-3' core sequence.

	GGAA	GGAT
$K'$ ( $\mu\text{M}$ ) <sup>a</sup>	0.4 $\pm$ (0.1)	0.8 $\pm$ (0.1)
$k_{-2}$ ( $\text{s}^{-1}$ ) <sup>b</sup>	6 ( $\pm$ 3) $\times 10^{+6}$	1.3 ( $\pm$ 0.07) $\times 10^{+6}$
$k_2$ ( $\text{s}^{-1}$ ) <sup>b</sup>	2.4 ( $\pm$ 0.3) $\times 10^{+5}$	1.6 ( $\pm$ 0.9) $\times 10^{+5}$
$K_d$ ( $\mu\text{M}$ ) <sup>b</sup>	0.08 ( $\pm$ 0.05)	1.1 ( $\pm$ 0.4)
$K_2$ <sup>c</sup>	0.04 ( $\pm$ 0.02)	0.12 ( $\pm$ 0.07)
$K_{D(\text{overall})}$ ( $\mu\text{M}$ ) <sup>d</sup>	2 ( $\pm$ 1)	10 ( $\pm$ 5)
$k_{-1}$ ( $\text{s}^{-1}$ ) <sup>e</sup>	0.61 ( $\pm$ 0.04) $\times 10^{-3}$	1.8 ( $\pm$ 0.6) $\times 10^{-3}$
$k_1$ ( $\text{M}^{-1}\text{s}^{-1}$ ) <sup>f</sup>	8 ( $\pm$ 5) $\times 10^{+3}$	1.7 ( $\pm$ 0.6) $\times 10^{+3}$
$\Delta G$ (kcal/mol) <sup>g</sup>	-8.7 ( $\pm$ 0.1)	-8.3 ( $\pm$ 0.1)

<sup>a</sup> Equilibrium constant ( $K'$ ) obtained from fitting equilibrium resonance response units data to Equation (1).

<sup>b</sup> Calculated by fitting apparent association rate ( $k_{\text{obs}}$ ) data to Equation (5).

<sup>c</sup> Equilibrium constant for isomerization of P to P\* calculated with the relationship,  $K_2 = k_2/k_{-2} = [\text{P}]/[\text{P}^*]$ .

<sup>d</sup> Calculated by substituting the  $K_d$  and  $K_2$  values into Equation (7).

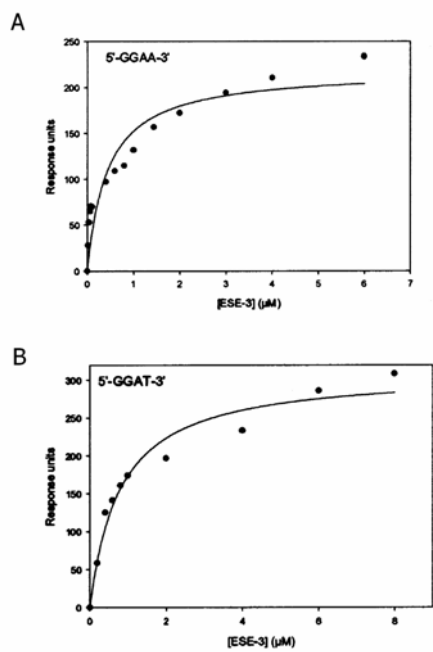
<sup>e</sup> Calculated from fitting of SPR sensorgrams to Equation (3).

<sup>f</sup> Determined by the equation  $K_d = k_{-1}/k_1$ .

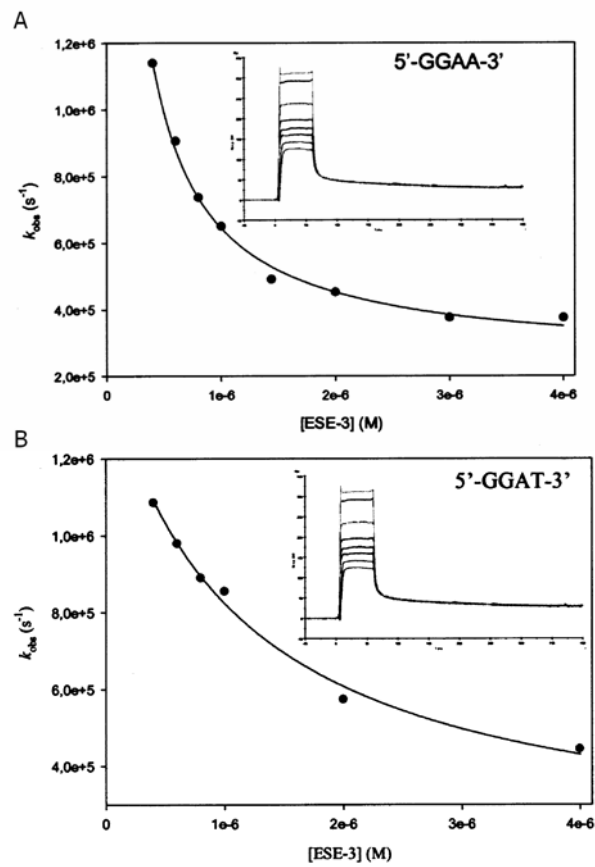
<sup>g</sup> Gibbs free energy change ( $\Delta G$ ) calculated by  $\Delta G = -RT \ln(1/K')$ .



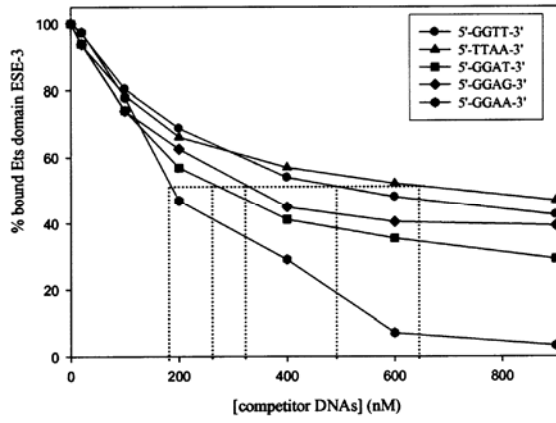
## Figures



**Figure 1.** Equilibrium SPR titration.



**Figure 2.** Kinetics of Ets-domain ESE-3 to *Drosophila* E74 containing either 5'-GGAA-3'(A) or 5'-GGAT-3'(B) core sequence.



**Figure 3.** Competition profile of the binding of Ets domain ESE-3 to immobilized *Drosophila* E74 containing 5'-GGAA-3' core sequence.

### 3 CONSIDERAÇÕES FINAIS

ESE-3 é um membro da família de fatores de transcrição Ets conhecido por apresentar expressão em células epiteliais. A proteína ESE-3 é constitutivamente expressa em células epiteliais diferenciadas, porém ocorre a perda da sua expressão em uma série de carcinomas epiteliais, sugerindo uma possível função como supressora tumoral [5]. Tem sido demonstrado que ESE-3 pode apresentar um importante papel no controle processo inflamatório de doenças respiratórias tais como a asma e a fibrose cística [51].

As proteínas Ets atuam regulando e controlando a expressão de genes. Os fatores de transcrição Ets se ligam em seqüências consenso de aproximadamente 9 pb contendo invariavelmente a seqüência central 5'-GGA (A ou T)-3' encontradas em regiões regulatórias de genes específicos [1,22,52]. Há uma significativa sobreposição entre as seqüências consenso, o que ressalta a existência de similaridade nas especificidades de ligação ao DNA, entre as proteínas Ets. No entanto, não se pode inferir conclusões a respeito da especificidade de ligação tomando por base somente a seqüência de oligonucleotídeos. Com isso, para cada proteína Ets a determinação de um sítio alvo, assim como a especificidade para diferentes sítios somente pode ser averiguada por experimentação direta [54].

Com o propósito de entender como determinadas seqüência de DNA são reconhecidas pelos fatores de transcrição Ets, tem se proposto dois mecanismos para a formação do complexo entre a proteína Ets e o DNA. O mecanismo de leitura indireta (*indirect readout mechanism*) [53], no qual as proteínas com domínio Ets reconhecem uma seqüência dependente da estrutura, que pode tanto ser pré-existente no DNA ou induzida devido à ligação a proteína. O outro chamado de mecanismo de leitura direta (*direct readout mechanism*) [54], no qual uma única base presente na seqüência central do sítio de ligação ao DNA, contribui para a estabilização do complexo binário. As interações com outras proteínas, sendo elas fatores de transcrição ou não, dependentes tanto do ambiente como do contexto celular nas quais se encontram, também determinam a especificidade de ligação a seqüências regulatórias presentes em genes alvos das proteínas Ets [4,7,8].

Contudo, a aquisição de dados quantitativos que permitam a obtenção de detalhes das características dinâmicas do processo de reconhecimento molecular é de extrema relevância. A técnica de espectroscopia por ressonância plasmônica de superfície permite realizar estudos de afinidade e cinética de ligação, determinar especificidade por diferentes seqüências de oligonucleotídeos e também determinação da estequiometria da reação, sendo portanto, uma excelente técnica para a obtenção dos dados pretendidos neste trabalho [55,56].

Conseqüentemente, os dados concedidos pela técnica de *SPR* resultaram na melhora da nossa compreensão a cerca do mecanismo biomolecular que envolve a interação entre a proteína ESE-3 e o DNA bem como do entendimento dos processos moleculares no quais ESE-3 participa.

Assim, com o objetivo de realizar os estudos da cinética de ligação do domínio Ets de ESE-3 às seqüências de oligonucleotídeos, correspondendo ao sítio de ligação Ets, contendo as bases centrais 5'-GGAA-3' ou 5'-GGAT-3', bem como para determinar a especificidade de ligação desta proteína para as seqüências DNAs contendo as diferentes bases centrais: 5'-GGAA-3', 5'-GGAT-3', 5'-GGAG-3', 5'-GGTT-3', 5'-TTAA-3' (ver as seqüências de oligonucleotídeos contidas no artigo). Foram realizadas as seguintes etapas para obtenção da proteína truncada, conforme previamente descrito no trabalho de conclusão de curso realizado por Jaskulski, L. [57]: (1) Construção da seqüência de DNA que codifica para o domínio Ets da proteína ESE-3 humana (Gen Bank NM\_012153); (2) Clonagem do fragmento de interesse no vetor pCR®-Blunt (Invitrogen); (3) Subclonagem no vetor de expressão pET23a(+) (Invitrogen). (4) Confirmação da identidade e da integridade do fragmento de DNA clonado, por seqüenciamento automático da seqüência do plasmídeo recombinante; (5) testes de expressão em diferentes cepas de *Escherichia coli*. Dentre as cepas testadas, foi observada a melhor expressão do domínio Ets recombinante nas frações solúvel e insolúvel da cepa *E. coli* BL21(DE3) crescidas à 37°C no meio de cultura LB, na condição não induzida. Portanto, estas condições de crescimento foram utilizadas para a produção da proteína truncada.

As proteínas contaminantes foram eliminadas utilizando sistema de purificação de Cromatografia Líquida de Alta Eficiência (*HPLC*) ÄKTA (GE Healthcare). O protocolo de purificação definido incluiu dois passos, sendo que o primeiro utiliza uma coluna cromatográfica que realiza troca catiônica e o segundo gel filtração. Uma única banda de massa molecular esperada foi observada em análise por SDS-PAGE corado com nitrato de prata, o que representa que o protocolo descrito é um método eficiente, no qual o domínio Ets de ESE-3 é obtido com pelo menos 98% de pureza. Um total de 23 mg da proteína recombinante homogeneia pôde ser obtido a partir de 10 g de células *E. coli* BL21(DE3), apresentando um bom rendimento. Análise do seqüenciamento de aminoácidos confirmou a identidade da proteína como sendo o domínio Ets de ESE-3 e a análise por espectrômetro de massas revelou que a proteína recombinante possui uma massa molecular de 11,993 Da, muito próximo do valor pré determinado, teórico: 12,003.70 Da.

Os resultados obtidos da ligação em equilíbrio determinada por *SPR*, apresentados no artigo, mostraram que não há grande diferença na afinidade de ligação do domínio Ets de ESE-3 entre as seqüências contendo 5'-GGAA-3' ou 5'-GGAT-3'. A afinidade para DNA

contendo 5'-GGAT-3' é duas vezes menor em relação ao DNA contendo 5'-GGAA-3'. No entanto, a análise dos dados cinéticos justifica a maior afinidade do domínio Ets para o DNA contendo as bases centrais 5'-GGAA-3', pois mostra que há na formação do complexo binário, um valor maior para constante de velocidade de associação e um valor menor para a constante de velocidade da dissociação do complexo formado. Estes dados sugeriram, pelo menos em parte, que o mecanismo de leitura direto [56] parece ter um papel no reconhecimento realizado pelo domínio Ets de ESE-3 ao DNA, já que a mutação da quarta base de A para T na seqüência central 5'-GGAX-3' (sendo X=A ou T) influenciou na estabilidade do complexo binário formado. Além disso, sugeriu-se um mecanismo de interação entre o domínio Ets e o DNA no qual há duas formas isoméricas da proteína em solução e somente uma dessas formas pode ligar-se ao DNA. Este passo caracteriza a fase lenta da reação, enquanto a fase rápida da reação é caracterizada pela formação do complexo binário.

O valor de  $K''$  (constante de dissociação em equilíbrio) para o DNA 5'-GGAA-3', obtido por espectrofluorimetria em estado estacionário (ver anexo I), forneceu evidências experimentais que sustentam a existência de mais um passo no mecanismo proposto, no qual após a rápida associação bimolecular, ocorre um passo de mudança conformacional do complexo formado entre o domínio Ets de ESE-3 e o DNA (P-DNA) para (P-DNA\*).

Os experimentos de competição sugerem uma ordem de afinidade do domínio Ets de ESE-3 pelos DNAs contendo as seguintes bases centrais: GGAA > GGAT  $\geq$  GGAG > GGTT  $\geq$  TTAA. Esses resultados estão de acordo com os obtidos pela constante de dissociação em equilíbrio, a qual apresentou valores maiores para o DNA contendo 5'-GGAT-3' quando comparado a 5'-GGAA-3'.

Deve-se levar em consideração, para o domínio Ets de ESE-3, que as bases que flanqueiam a seqüência central 5'-GGAA-3' também apresentam um papel no reconhecimento de regiões promotoras [58]. Além disso, devemos sempre lembrar que a ativação ou não da transcrição de genes específicos é influenciada pelo ambiente e pelo contexto celular [52,59]. Ainda assim, nossos resultados demonstram que quando se busca o entendimento das interações entre proteínas e o DNA também devem ser levadas em consideração as características dinâmicas dessas interações.

## REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Sharrocks AD, Brown L, Ling Y, Yates PR. The Ets-domain transcription factor family. *Int. J. Biochem. Cell Bio.* 1997;29:1371–1387.
- [2] Degnan BM, Degnan SM, Naganuma T, and Morse D E. The *ets* multigene family is conserved throughout the Metazoa. *Nucleic Acids Res.* 1993;21:3479–3484.
- [3] Laudet V, Niel C, Duterque-Coquillaud M, Leprince D, and Stehelin D, *et al.* Evolution of the *ets* gene family. *Bioch Bioph Res Comm.* 1993;190:8–14.
- [4] Oikawa T. ETS transcription factors: Possible targets for cancer therapy. *Cancer Sci.* 2004;95: 626-633.
- [5] Seth A, Watson DK. ETS transcription factors and their emerging roles in human cancer. *Europ J of Canc.* 2005;41:2462-2478.
- [6] Sharrocks AD. The ETS-Domain transcription factor family. *Nat rev Mol Cell Biol.* 2001;2:897-837.
- [7] Li R, Pei H, Watson DK. Regulation of ETS function by protein-protein interactions. *Oncog.* 2000;19:6514–6523.
- [8] Wasylyk B, Hagman J. Gutierrez-Hartmann A. Ets transcription factors: nuclear effectors of the Ras-MAP-kinase signaling pathway. *Trends Biochem Sci.* 1998;23:213–216.
- [9] Raves BJ, Petersen JM. Specificity within the *ets* Family of Transcription Factors. *Cancer Res.* 1998;75:1-55.
- [10] Liang H, Mao X, Olejniczak ET, Nettesheim DG, Yu L, Meadows RP, *et al.* Solution structure of Fli-1 when bound to DNA. *Nature Struct. Biol.* 1994;1:871–876.
- [11] Donaldson LW, Petersen JM, Graves BJ, McIntosh, LP, *et al.* Solution structure of the ETS-domain from murine Ets-1: a winged helix–turn–helix motif. *EMBO J.* 1996;15:125–134.
- [12] Werner MH, Clore GM, Fisher CL, Fisher RJ, Trinh L, Shiloacj J, *et al.* Correction of the NMR structure of the ETS1/DNA complex. *J Biomol NMR.* 1997;10:317–328.
- [13] Kodandapani R. A new pattern for helix–turn–helix recognition revealed by the PU.1 ETS-domain DNA complex. *Nature.* 1996;380:456–460.
- [14] Batchelor AH, Piper DE, Brousse FC, *et al.* The structure of GABP $\alpha/\beta$ : an ETS-domain-ankyrin repeat heterodimer bound to DNA. *Science.* 1998;279:1037–1041.
- [15] Mo Y, Vaessen B, Johnston K, Marmorstein R. Structures of SAP-1 bound to DNA sequences from the E74 and *c-fos* promoters provide insights into how ETS proteins discriminate between related DNA targets. *Mol. Cell.* 1998;8:210–212.

- [16] Mo Y, Vaessen, Johnston K, Marmorstein R. Structure of the elk-1-DNA complex reveals how DNA distal residues affect ETS domain recognition of DNA. *Nature Struct. Biol.* 2000;7:292–297.
- [17] 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;33:13509-13516.
- [18] Liang H, Olejniczak ET, Mao C, Nettesheim DG, Yu L, Thompson CB, *et al.* The Secondary Structure of the Ets Domain of Human Fli-1 Resembles that of the Helix-Turn-Helix DNA-Binding Motif of the Escherichia coli Catabolite Gene Activator Protein. *Proc. Natl. Acad. Sci. U.S.A.* 1994;91:11655-11659.
- [19] 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;280(50):41421-41428.
- [20] Pio F, Kondandapani R, Ni C, Ni C-Z, Shepard W, Klemszi M, *et al.* News insights on DNA Recognition by ets Proteins from Crystal Structure of the PU.1 ETS Domain-DNA. *J Biol Chem.* 1996;271:23329-37.
- [21] Pio F, Assa-Munt N, Yguerabide J, and Maki R A, *et al.* Mutants of ETS domain PU.1 and GGAA/T recognition: Free energies and Kinetics. *Protein Science.* 1999;8:2098-2109.
- [22] Bosselut R, Levin J, Adjadj E, and Ghysdael J. A single amino acid substitution in the ETS domain alters core DNA binding specificity of ETS-1 to that of the related transcription factors ELF-1 and E74. *Nucleic Acids Res.* 1993;21:5154-5191.
- [23] Caceres RA, Nunes CP, Azevedo WFJ, Basso LA and Santos DS. Molecular dynamics and structural studies of the Ets domain-DNA complexes. *Curr Bioinf.* 2007;2:222-228.
- [24] Pahl HL, Scheibe SRJ, Zhang D, Chen H-M, Galson DL, Maki RA, *et al.* The proto-oncogene PU.1 regulates expression of the myeloid-specific CD11b promoter. *J Biol Chem.* 1993;268(7):5014-20.
- [25] Garvie CW, Hagman J, Wolberger C. Structural Studies of Ets-1/Pax Complex Formation on DNA. *Molecular Cell.* 2001;8:1267-1276.
- [26] Urness LD, Thummel CS. Molecular interactions within the ecdysone regulatory hierarchy: DNA binding properties of the Drosophila ecdysone-inducible E74A protein. *Cell.* 1990;63:47-61.
- [27] Shore P, Whitmarsh AJ, Bhaskaran R, Davis RJ, Waltho JP, *et al.* Determinants of DNA-binding specificity of ETS-domain transcription factors. *Mol Cell Biol.* 1996;16: 3338–3349.
- [28] Fitzsimmons D, Hodsdon W, Wheat W, Maira S-M, Wasylyk B, and Hagman J, *et al.* Pax-5 (BSAP) recruits Ets protooncogene family proteins to form functional ternary complexes on a B-cell-specific promoter. *Genes Dev.* 1996;10:2198–2211.
- [29] Mo Y, Vaessen B, Johnston K, Marmorstein R. Structure of the Elk-1-DNA complex reveal how DNA-distal residues affect ETS domain recognition of DNA. *Nature structural biology*, 2000;7:292-297.



- [30] Oettgen P, Kas K, Dube A, Gu X, Grall F, Thamrongsak U, *et al.* Characterization of ESE-2, a Novel ESE-1-related Ets Transcription Factor That Is Restricted to Glandular Epithelium and Differentiated Keratinocytes. *J Biol Chem.* 1999;274:29439–29452.
- [31] 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;275:2986–2998.
- [32] Tugores A, Le J, Sorokina I, Snijders AJ, Duyao M, Reddy OS, *et al.* The epithelium-specific ETS protein EHF/ESE-3 is a context-dependent transcriptional repressor downstream of MAPK signaling cascades. *J Biol Chem.* 2001;276:20397–20406.
- [33] Choi YS, Shina S. Determination of the consensus DNA-binding sequence and a transcriptional activation domain for ESE-2. *Bioch J.* 2006;389(3):497-507.
- [34] Oettegen P, Alani RM, Barcinski MA, Brown L, Akbarali Y, Boltax J, *et al.* Isolation and characterization of a novel epithelium-specific transcription factor, ESE-1, a member of the ETS family. *Mol Cell Biol.* 1997;17:4419–4433.
- [35] Kleinbaum LA, Duggan C, Ferreira E, Coffey GP, Buttice G, Burton FH, *et al.* Human chromosomal localization, tissue/tumor expression, and regulatory function of the ETS family gene EHF. *Biochem Biophys Res Com.* 1999;264:119–126.
- [36] Appel S, Bringmann A, Grünebach F, Weck MM, Bauer J, Brossart P. Epithelial-specific transcription factor ESE-3 is involved in the development of monocyte-derived DCs. *Blood.* 2005;107:8.
- [37] Slutsky AS, Zamel N. Genetics of Asthma: the University of Toronto Program. *Am J Respir Cri. Care Med.* 1997;156:S130–S132.
- [38] Sreekumar GP, Binnie CG, Carden EG, Sharma S, Sprankle CS, Sherman PA, *et al.* Association of asthma and related pheontypes with ASTH-I/J genes on chromosome 11p in Caucasians. *J Am Soc Hum Genet.* 2000;A354:1974.
- [39] Brooks-Wilson AR, Buckler A, Cardon L, Carey A, Galvin M, Miller A, *et al.* Asthma related genes. International patent publication number: W0 99/37809. 1999.
- [40] Weinberg RA. Tumor suppressor genes. *Science.* 1991;254:1138–1146.
- [41] Reeve AE, Sih SA, Raizis AM, and Feinberg AP, *et al.* Loss of allelic heterozygosity at a second locus on chromosome 11 in sporadic Wilms' tumor cells. *Mol Cel. Biol.* 1989;9:1799–1803.
- [42] Fong KM, Zimmerman PV, Smith PJ. Correlation of loss of heterozygosity at 11p with tumour progression and survival in non-small cell lung cancer. *Gen Chrom Canc.* 1994;10:183–189.
- [43] Shipman R, Schraml P, Colombi M, Ludwing CU. Allelic deletion at chromosome 11p13 defines a tumour suppressor region between the catalase gene and D11S935 in human non-small cell lung carcinoma. *Int J Oncol.* 1998;12:107–111.

- [44] Feldman RJ, Sementchenko VI, Gayed M, Fraig MM, and Watson DK. Pdef expression in human breast cancer is correlated with invasive potential and altered gene expression. *Cancer Res.* 2003;63:4626–4631.
- [45] Nozawa M, Yomogida K, Kanno N, Nonomura N, Miki T, Okuyama A, *et al.* Prostate-specific transcription factor hPSE is translated only in normal prostate epithelial cells. *Cancer Res.* 2000;60:1348–1352.
- [46] Fujikawa M, Katagiri T, Tugores A, Nakamura Y, Ishikawa F, *et al.* ESE-3, an Ets family transcription factor, is up-regulated in cellular senescence. *Cancer spi.* 2007;98(9):1468-75.
- [47] Silverman ES, Baron RM, Palmer LJ, Le L, Hallock A, Subramaniam V, *et al.* Constitutive and cytokine-induced expression of the ETS transcription factor ESE-3 in the lung. *Am. J. Respir. Cell. Mol. Biol.* 2002;27: 696-704.
- [48] Reddy SPM, Vuong H, Adisheshaiah PP. Interplay between Proximal and Distal Promoter Elements Is Required for Squamous Differentiation Marker Induction in the Bronchial Epithelium. *The J of Biol Chem.* 2003;278(24):21378-21387.
- [49] Reddy SPM, Adisheshaiah P, Shapiro P, and Vuong H. BMK1 (ERK) Regulates Squamous Differentiation Marker *SPRR1B* Transcription in Clara-like H441 Cells. *Am J Respir Cell Mol Biol.* 2002;27:64-70.
- [50] Deng J, Pan R, Wu R. Distinct Roles for Amino- and Carboxyl-terminal Sequences of SPRR1 Protein in the Formation of Cross-linked Envelopes of Conducting Airway Epithelial Cells. *J. Biol. Chem.* 2000;275:5739-5747.
- [51] Wu J, Duan R, Cao H, Field D, Newnham CM, Koehler DR, Zamel N, *et al.* Regulation of epithelium-specific Ets-like factors ESE-1 and ESE-3 in airway epithelial cells: potential roles in airway inflammation. *Cell Res.* 2008;18:649-663.
- [52] Hsu T, Trojanowska M. and Watson DK. Ets proteins in biological control and cancer. *J. Cell. Biochem.* 2004;91:896-903.
- [53] Luscombe NM, Laskowski RA, and Thornton JM. Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nuc Acids Res.* 2001;29:2860-2874.
- [54] Seeman NC, Rosenberg JM and Rich A. Sequence-specific recognition of double helical nucleic acids by proteins. *Prot. Natl. Sci. USA.* 1976;73,804-808.
- [55] Brockman JM, Frutos AG, and Corn RM. (1999) A multistep chemical modification procedure to create DNA arrays on gold surfaces for the study of protein-DNA interactions with surface plasmon resonance imaging. *J. Am. Chem. Soc.*, 121, 8044-8051.
- [56] The HF, Peh WYX, Su X, and Thomsen JS. (2007) Characterization of Protein-DNA Interactions Using Surface Plasmon Resonance Spectroscopy with Various Assay Schemes. *Biochemistry*, 46, 2127-2135.

[57] Jaskulski L. Clonagem, expressão e caracterização do domínio Ets de ESE-3. Monografia (Graduação em Farmácia)-Faculdade de Farmácia, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre; 2006.

[58] Lim HJ, Cho JY, Park YB, Park JW and Kwon TK. ESE-3 transcription factor is involved in the expression of death receptor (DR-5) through putative Ets sites. *Biochem. Biophys. Res. Comm.* 2006;350:736-741.

[59] Buchwalter G, Gross C, Wasylyk B. Ets ternary complex transcription factors. *Gene.* 2004;324:1-14.

[60] Bredemeier-Ernst I, Nordheim A, and Jaknecht R. Transcriptional activity and constitutive nuclear localization of the ETS protein Elf-1. *FEBS Lett.* 1997;12:47-51.

## ANEXO I

### **Determinação da especificidade de ligação do domínio Ets de ESE-3 às seqüências de DNA (E74 selvagem, E74 mutado) através da técnica de espectrofluorimetria em estado estacionário:**

Foram realizados experimentos de ligação em equilíbrio utilizando a técnica de espectrofluorimetria em estado estacionário, a qual permite observar a variação da fluorescência intrínseca de aminoácidos triptofano presentes no domínio Ets de ESE-3 devido à ligação dos fragmentos de DNA.

Duas seqüências de oligonucleotídeos foram desenhadas baseadas na seqüência consenso do sítio de ligação Ets do gene E74 de *Drosophila* descrita por Bredemeier-Ernst *et al* [60], sendo que uma é constituída do tipo selvagem (5'-AGCTTCTCTAGCTGAATA ACCGGAAGTAACTCATCG-3') enquanto a outra é um mutante gerado pela mudança da seqüência central de 5'-GGAA-3' deste EBS para 5'-GGTT-3'. As seqüências de oligonucleotídeos foram sintetizadas pela Invitrogen. As moléculas complementares de DNA foram dissolvidas em água e aneladas com temperatura inicial de 80°C por 1' e a cada ciclo diminuindo 1 grau, propiciando a temperatura ideal de anelamento das fitas, até atingir 4°C. As soluções contendo as seqüências dupla fitas foram preparadas de tal forma que a solução final continha HBS-EP (100 mM HEPES 1.5 M NaCl 30 mM EDTA 0.5% v/v Surfactant P20 pH 7.4).

Os dois experimentos de fluorescência foram conduzidos da mesma maneira. Estes ensaios foram realizados em um espectrofluorímetro Shimadzu, modelo RF-5301, conectado a um banho termostático a 25°C. Foi utilizado para irradiar a proteína (domínio Ets de ESE-3) o comprimento de onda de 298nm com *slit* (de excitação) de 3nm, e os espectros de emissão de fluorescência da proteína foram registrados na faixa de 300-450 nm, com *slit* (de emissão) de 10 nm. Utilizou-se para a titulação, cuvetas de quartzo contendo 2 mL da solução da proteína na concentração de 2 µM em tampão HBS-EP e uma cuveta extra no caminho da luz de excitação somente com tampão HBS-EP. Foi realizada a titulação da proteína (o domínio Ets de ESE-3) nas faixas de concentração [0,042-1,2 µM], [0,103-2,2 µM] para os DNAs, contendo no seu centro as bases 5'-GGAA-3' e 5'-GGTT-3', respectivamente, sendo que, a diluição máxima não ultrapassou a 3,15 %. Os espectros de emissão da proteína foram realizados na ausência de DNA (obtido somente pela fluorescência intrínseca do domínio Ets de ESE-3) e também imediatamente após a mistura com as moléculas de DNA. Curvas

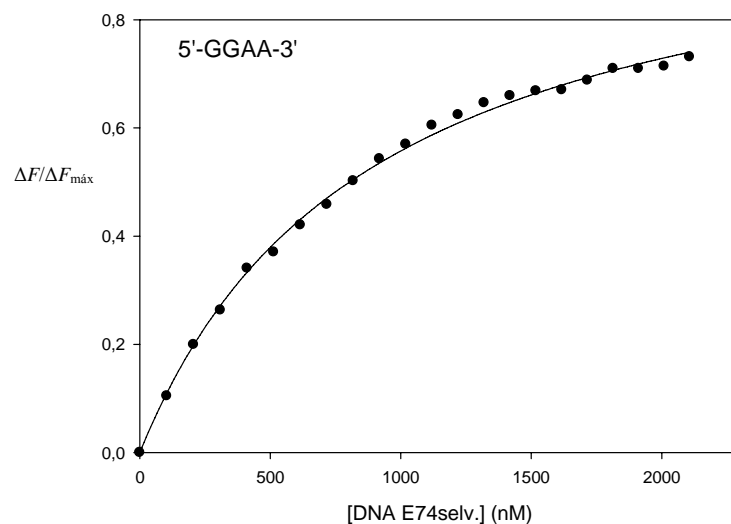
controle foram realizadas com a presença da solução de proteína (onde não foi adicionada solução de DNA) e uma cubeta extra no caminho da luz de excitação que continha HBS-EP, a qual foi adicionada o DNA na mesma concentração utilizada para titular a proteína e também acrescentando o mesmo volume de HBS-EP à solução de proteína. Estas curvas controle permitiram a subtração do efeito de filtro interno daqueles valores obtidos da titulação da proteína com as soluções de DNA, com isso, resultando somente em valores da variação da fluorescência do domínio Ets de ESE-3 devido à ligação às moléculas de DNA.

A análise dos dados experimentais foi realizada através de regressão não-linear, utilizando o programa SigmaPlot 9.0. Nas três análises se observou um aumento hiperbólico da variação da fluorescência devido à ligação, com aumento da concentração do DNA (Figura 5A e 5B). Os dados foram adequados à equação (a), onde a  $(\Delta F)$  representa a amplitude da fluorescência da proteína para cada concentração do DNA,  $F_{max}$  a amplitude de fluorescência máxima na concentração do DNA que satura todos os sítios disponíveis da proteína e  $K''$  a constante de dissociação em equilíbrio, que corresponde à metade do valor da variação de fluorescência máxima, enquanto  $[P]_0$  e  $[DNA]_0$  representam as concentrações iniciais da proteína e do DNA, respectivamente.

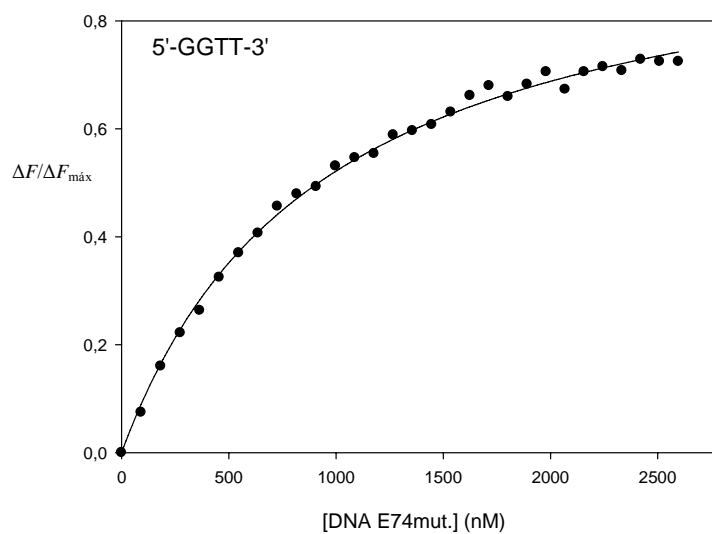
$$\frac{\Delta F}{\Delta F_{max}} = \frac{1}{2[P]_0} \left\{ (K'' + [P]_0 + [DNA]_0) \pm \sqrt{((K'' + [P]_0 + [DNA]_0)^2 - (4[P]_0[DNA]_0))} \right\} \quad (a)$$

Os valores das constantes de dissociação em equilíbrio obtidos por espectrofluorimetria em estado estacionário foram determinados com o propósito de avaliar a afinidade do domínio Ets de ESE-3 às seqüências de DNA que apresentam no seu centro as bases 5'-GGAA-3' e 5'-GGTT-3'. Através da análise dos dados foram obtidos os valores da  $K''$  (constante de dissociação em equilíbrio) para as moléculas de DNA 5'-GGAA-3' e 5'-GGTT-3' de  $880 \pm 30$  nM e  $940 \pm 30$  nM, respectivamente. Os valores  $K''$  demonstram que o domínio Ets de ESE-3 se liga à seqüência E74 de *Drosophila* contendo as bases centrais 5'-GGAA-3' sem grandes variações na afinidade em relação à seqüência mutante 5'-GGTT-3'.

Sendo que, o valor calculado para a  $K_{D(overall)}$ , constante de dissociação em equilíbrio global, ( $2 \pm 1$   $\mu$ M) para o DNA contendo as bases 5'-GGAA-3' é maior do que o valor da constante de dissociação em equilíbrio ( $K''$ ) obtido pela técnica de espectrofluorimetria em estado estacionário, o valor da  $K''$  fornece evidência experimental de que há mais um passo no mecanismo proposto, no qual a rápida associação bimolecular é seguida por uma mudança conformacional do complexo binário formado entre o domínio Ets de ESE-3-DNA, ou seja do complexo P-DNA para P-DNA\*.



**Figura 5A.** Espectrofluorimetria em estado estacionário, ligação em equilíbrio do DNA E74 selvagem (5'-GGAA-3') ao domínio Ets de ESE-3. O aumento da variação da fluorescência da proteína devido à ligação ao DNA depende do aumento da concentração deste DNA.



**Figura 5B.** Espectrofluorimetria em estado estacionário, ligação em equilíbrio do DNA E74 mutante (5'-GGTT-3') ao domínio Ets de ESE-3. O aumento da variação da fluorescência da proteína devido à ligação ao DNA depende do aumento da concentração do DNA.

## APÊNDICE – Comprovante de submissão do artigo científico

O comprovante de submissão do artigo científico: *Equilibrium and kinetics of Ets domain ESE-3 binding to 5'-GGAX-3' (X=A or T) core sequence assessed by surface plasmon resonance*, é mostrado abaixo:

Manuscript Central Página 1 de 1

---

OXFORD JOURNALS Edit Account | Instruc

# Nucleic Acids Research

[Main Menu](#) → [Corresponding Author Dashboard](#) → Submission Confirmation

## Submission Confirmation



Thank you for submitting your manuscript to *Nucleic Acids Research*.

Manuscript ID: NAR-00176-2009

Title: Equilibrium and kinetics of Ets domain ESE-3 binding to 5'-GGAX-3' (X=A or T) sequence assessed by surface plasmon resonance

Authors: Jaskulski, Leia  
Nunes, Claudia  
Ducati, Rodrigo  
Renard, Gaby  
Chies, Jocelei  
Basso, Luiz  
Santos, D.

Date Submitted: 28-Jan-2009

 Print  Return

