

UNIVERSIDADE NOVA DE LISBOA

FACULDADE DE CIÊNCIAS E TECNOLOGIA

DEPARTAMENTO DE CIÊNCIAS DA VIDA

Joana Raquel Nabais

**Localização de mRNA em *Drosophila*
melanogaster: estudo da proteína Ypsilon
Schachtel**

*Dissertação apresentada para a obtenção do Grau de Mestre
em Genética Molecular e Biomedicina, pela Universidade
Nova de Lisboa, Faculdade de Ciências e Tecnologia*

Orientador:

Prof. Doutor José Trincão (FCT/UNL)

LISBOA

2010

nº de arquivo

“Copyright”

UNIVERSIDADE NOVA DE LISBOA

FACULDADE DE CIÊNCIAS E TECNOLOGIA

DEPARTAMENTO DE CIÊNCIAS DA VIDA

Joana Raquel Nabais

**Localização de mRNA em *Drosophila*
melanogaster: estudo da proteína Ypsilon
Schachtel**

*Dissertação apresentada para a obtenção do Grau de Mestre
em Genética Molecular e Biomedicina, pela Universidade
Nova de Lisboa, Faculdade de Ciências e Tecnologia*

Orientador:

Prof. Doutor José Trincão (FCT/UNL)

LISBOA

2010

Agradecimentos

“Há gente que fica na história, na história da gente. E outras de quem nem o nome lembramos ouvir”. A todos os que fazem e fizeram parte da minha história, o meu mais sincero obrigado!

A ti Avô, pela tua existência me ter inspirado a ultrapassar todas as adversidades e pela saudade que sinto, lembrar-me de quem sempre acreditaste que eu seria... e sou.

A ti Mãe, por nunca desistires. Por teres sido pai, mãe e amiga. Por me secares as lágrimas quando o mundo é injusto e nunca me deixares esquecer que sou capaz... sempre.

A ti Pai, por tentares conhecer-me e compreender-me. À restante Família pelo carinho e incentivo.

Às minhas amigas de sempre Cátia, Saras e Leonor, por serem as irmãs que sempre desejei.

Aos Professores que me formaram e que nunca me deixaram desistir, em especial ao Professor Lampeira que esteve sempre disponível para me ajudar a ultrapassar todos os obstáculos que tive ao longo do meu percurso académico.

Ao meu Orientador, por ter acreditado em mim.

À Professora Maria João Romão por me ter acolhido no grupo XTAL.

À Diana, pela ajuda prestada ao longo deste ano.

Aos meus colegas e amigos, David, Filipe, Fábio, Catarina, Teresa e Cecília, por nunca me deixarem esquecer o porquê de ter escolhido este percurso. Obrigado pelas gargalhadas, pelas aulas de ténis e pelos conselhos. E principalmente, por me terem ajudado a continuar, quando a dois meses do prazo de entrega, perdi a pasta com todo o material para esta Tese.

A todos, obrigado por quem fui, sou e serei.

Sumário

Pretendeu-se com este trabalho, a clonagem do gene *ypsilon schachtel* (*yps*) e a consequente expressão e purificação da respectiva proteína para posterior resolução da estrutura tridimensional, pela técnica de cristalografia de raios-X. A proteína *yps* pertence à família das proteínas Y-box, conhecida pela capacidade de ligação ao elemento Y-box do DNA. Recentemente, constatou-se que a *yps* faz parte de um complexo ribonucleoproteico envolvido na localização de mRNA nas regiões anterior e posterior do oócito de *Drosophila melanogaster*. Aparenta ter um papel activo e abrangente neste processo, desconhecendo-se ainda os mecanismos através dos quais o desempenha. A resolução da estrutura tridimensional da proteína, permitiria a análise do seu arranjo e reconhecimento de regiões que poderão estar envolvidas nas hipotéticas interacções entre a proteína e o complexo que constitui, e/ou entre a proteína e o mRNA que é transportado. Atendendo essencialmente ao tamanho da proteína, a cristalografia de raios-X apresenta-se como o melhor método para o estudo pretendido. Para tal, é necessário proceder-se à clonagem do cDNA do gene *yps* em vector apropriado; sobre-expressar, isolar, purificar e cristalizar a proteína nele codificada. Em estudos anteriores, verificou-se que o grau de ordem de uma proteína afecta directamente o seu processo de cristalização: proteínas mais ordenadas constituem cristais mais ordenados e consequentemente, obtêm-se conjuntos de dados de melhor resolução. Assim, recorreu-se a diversas ferramentas bioinformáticas disponíveis com o intuito de se antever a globularidade da proteína *yps*. Verificou-se que se prevê alguma ordem apenas nos primeiros 121 aminoácidos e, deste modo, esta região constitui o objecto de estudo do trabalho aqui exposto.

Esta tese, contempla o estudo desde a clonagem do gene *yps* à purificação da proteína *yps*. Amplificaram-se dois fragmentos distintos da região globular mencionada, e ensaiou-se a sua clonagem em três vectores de clonagem e expressão: pGEX-6P1, pET-14b e pET-15b. Apenas com este último, a clonagem foi bem-sucedida. Transformou-se o vector recombinante em células competentes apropriadas e expressou-se a proteína em *Escherichia coli*, pelos métodos de indução e auto-indução. Por fim, a proteína alvo foi isolada e purificada por cromatografia de afinidade com metal imobilizado.

Índice

AGRADECIMENTOS	i
SUMÁRIO.....	ii
ÍNDICE.....	iii
INTRODUÇÃO.....	1
<i>Drosophila melanogaster</i>	1
Oogénese	2
Embriogénese	5
Ciclo de Vida.....	6
Localização de mRNA.....	7
Mecanismos de Localização de RNA.....	9
Formação do Padrão Anterior-Posterior.....	10
Gradiente Morfogénico.....	12
Complexo Ribonucleoproteico	16
Ypsilon Schachtel	18
MATERIAIS E MÉTODOS.....	24
Meios e Tampões.....	24
Clone em estudo	25
Previsão da Globularidade da Proteína.....	27
Vectores de Clonagem e Expressão.....	28
Preparação dos Vectores.....	29
Verificação dos Vectores Purificados – Electroforese em Gel de Agarose	29
Amplificação dos Fragmentos	30
Oligonucleótidos Iniciadores – <i>primers</i>	30
Simulação da Reacção de Amplificação	31

Mistura Reaccional	32
Programa de Amplificação	32
Confirmação e Extracção dos Fragmentos Amplificados – Electroforese em Gel de Agarose	33
Purificação dos Fragmentos	34
Purificação Manual – Em Coluna.....	34
Purificação Automatizada – Em Gel	35
<i>Kit</i> de Clonagem	35
Confirmação de Clonagem	37
Hidrólise Enzimática dos Fragmentos e Vectores	37
Enzimas de Restrição Utilizadas	38
Simulação das Reacções de Hidrólise	38
Liofilização	39
Reacção de Ligação Vector-Fragmento	40
Confirmação da Reacção de Ligação	40
Sequenciação de DNA.....	41
Transformação	41
Testes de Expressão.....	42
Indução com IPTG	42
Auto-indução	43
Preparação do Lisado Bacteriano	44
Análise dos Resultados dos Testes de Expressão	44
Purificação da Proteína – Cromatografia por Afinidade	45
Método de Bradford	47
RESULTADOS	48
Hidrólise e Purificação dos Vectores.....	49
Ensaio com os Vectores pGEX-6P1, pET-14b, pET-15b e pET-SUMO-28a	50

Reacções de PCR e Purificação dos Fragmentos Amplificados.....	50
Ensaio com os Vectores pGEX-6P1, pET-14b e pET-15b	51
Hidrólises Enzimáticas e Purificação dos Produtos Hidrolisados	51
Reacção de Ligação	52
Transformação dos Clones em Células NZY5 α	54
Ensaio com os Vectores pET-14b e pET-15b	54
<i>Kit</i> de Clonagem	54
Ensaio com o Vector pET-15b.....	58
Purificação do Fragmento Extraído do Clone C _{1,4}	58
Reacção de Ligação	58
Sequenciação	58
Transformação do Clone em NZYStar (NZYTech)	60
Transformação do Clone em BL21 (DE3) (NZYTech).....	61
Expressão da Proteína yps por Indução com IPTG	61
Expressão da Proteína yps por Auto-Indução	63
DISCUSSÃO	65
BIBLIOGRAFIA	75
ANEXOS	80
Anexo I - Vectores de Clonagem e Expressão	80
Anexo II – Marcadores de DNA e Proteína	83
Anexo III – E-Gel iBase Power System e E-Gel Safe Imager Real-Time Transilluminator.....	84

Introdução

Pretendeu-se com este trabalho a clonagem do gene *ypsilon schachtel* envolvido na localização de mRNA em *Drosophila melanogaster* e, a expressão da proteína nele codificada. A localização de mRNA é um mecanismo fulcral em vários processos celulares, entre os mais importantes, destaca-se o desenvolvimento dos organismos, aprendizagem e memória. No trabalho aqui explanado, pretendeu-se compreender melhor os mecanismos de transporte e localização de mRNA através do estudo bioquímico e estrutural de um conjunto de proteínas constituintes de um complexo ribonucleoproteico, nomeadamente, a proteína Yps.

Drosophila melanogaster

Drosophila melanogaster é uma pequena mosca de aproximadamente 3 mm de comprimento e vulgarmente conhecida como “mosca da fruta” ou “mosca do vinagre” por se encontrar normalmente em torno da fruta em putrefacção.

A utilização da *Drosophila* como organismo experimental teve início em 1908 pelo investigador norte-americano Thomas Hunt Morgan, que a utilizava para estudos genéticos (Rubin e Lewis, 2000). Os estudos desenvolvidos em *Drosophila* valeram-lhe o prémio Nobel na categoria de Fisiologia/Medicina, com a demonstração de que os cromossomas são portadores de genes (The Official Website of the Nobel Prize em http://nobelprize.org/nobel_prizes/medicine/laureates/1933/morgan-bio.html). Desde então, a *Drosophila* tem sido sujeita a extensos estudos genéticos, o que leva a comunidade científica a conhecer mais sobre a genética de *Drosophila* do que de qualquer outro organismo eucarionte. Esta preferência prende-se com a facilidade de manutenção e manuseamento da mosca em laboratório: fácil de cruzar, resistente, prolífica, desenvolve-se à temperatura ambiente, poucas exigências alimentares, pequeno espaço de cultura, cromossomas politenos no estado larvar, existência de colecções de mutantes depositadas em bases de dados de livre acesso, entre outros factores (Gilbert, 2003; Goldstein e Fyrberg, 1994). Devido aos baixos

custos envolvidos para estudos laboratoriais, a *Drosophila* é também conhecida como o “modelo dos estudantes”.

Contudo, apesar de se tratar de um organismo valioso para estudos genéticos, não era passível de se estudar a sua embriologia: os embriões de mosca são demasiado complexos e instáveis, além disso, o seu reduzido tamanho não permitia a sua manipulação experimental. No entanto, o crescente conhecimento da genética de *Drosophila* através da identificação e manipulação dos seus genes e RNA (ácido ribonucleico), permitiram que se estabelecesse uma ligação entre os acontecimentos genéticos e as várias etapas do seu desenvolvimento. Esta sinergia entre genética e desenvolvimento encontrada em *Drosophila*, serve de referência à maioria dos estudos que se realizam em biologia do desenvolvimento tendo outros organismos como objecto de estudo (Gilbert, 2003).

O desenvolvimento do embrião de *Drosophila* pode ser considerado como um processo de duas fases, cada uma constituída por vários estágios. A primeira fase do desenvolvimento ocorre nos ovários do insecto fêmea e denomina-se oogénese. Nesta etapa são produzidas algumas centenas de ovos que, quando fecundados, entram na segunda fase do desenvolvimento, a embriogénese, da qual resultará o embrião.

Oogénese

A oogénese é o processo através do qual ocorre a diferenciação do óvulo, diferindo grandemente do processo análogo no macho, a espermatogénese. A diferença mais marcante entre estes dois processos é o produto de ambos: enquanto o espermatozóide é essencialmente um núcleo móvel, o óvulo é um núcleo haplóide estacionário que contém armazenado no seu interior várias enzimas citoplasmáticas, mRNA, organelos e substratos metabólicos (Gilbert, 2003). Os mecanismos de oogénese variam grandemente não só entre espécies como também dentro de uma mesma espécie. Em *Drosophila melanogaster* (Riechmann e Ephrussi, 2001), ocorre uma oogénese meroística (Bastock e St. Johnston, 2008) na qual, as células produzidas pelo oogónio (Figura 1) mantêm as ligações citoplasmáticas entre si.

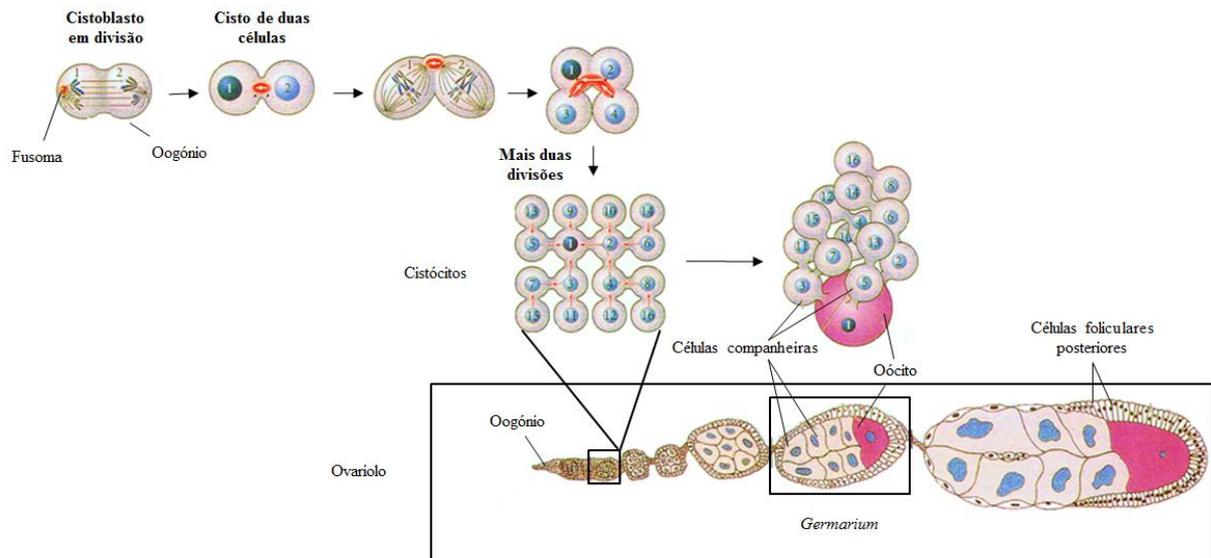


Figura 1 - Oogenese em *Drosophila*. Adaptada de Gilbert, 2003

O processo de formação do oócito, apesar de complexo, pode ser caracterizado resumidamente:

- 1) A câmara do ovo representa a unidade fundamental morfológica e fisiológica onde ocorre a oogenese. É composta por um oócito e 15 células companheiras rodeadas por uma monocamada de células foliculares, o conjunto denomina-se *germarium*. As câmaras do ovo organizam-se linearmente formando os ovariolos.
- 2) Na extremidade de cada ovariolo, o oogônio, uma linhagem germinal de células estaminais, divide-se assimetricamente para produzir duas células: uma célula que mantém as características da célula estaminal e um cistoblasto. Desde a primeira divisão, é constituída assimetricamente uma estrutura rica em espectrina, o fusoma: a sua expansão ocorre desde o pólo do fuso que permanece numa das células após a divisão, e irá atravessar os canais anel entre as células (Figura 1, célula 1 do cistoblasto).
- 3) O cistoblasto sofre quatro divisões mitóticas síncronas com citocinése incompleta, dando origem a um aglomerado de 16 células (cistócitos) interligadas por pontes intracelulares denominadas de canais anel (Figura 2). Apenas os dois cistócitos que possuem quatro ligações (Figura 1, cistócitos 1 e 2) podem vir a originar oócitos, e destes dois apenas um dará origem ao óvulo, célula que se encontra na extremidade mais posterior do ovariolo e que retém a maior parte do fusoma (Figura 1, cistócito 1).

- 4) As restantes células, tornam-se células companheiras (Gigliotti *et al.*, 2004; Hudson e Cooley, 2010; Gilbert, 2003).

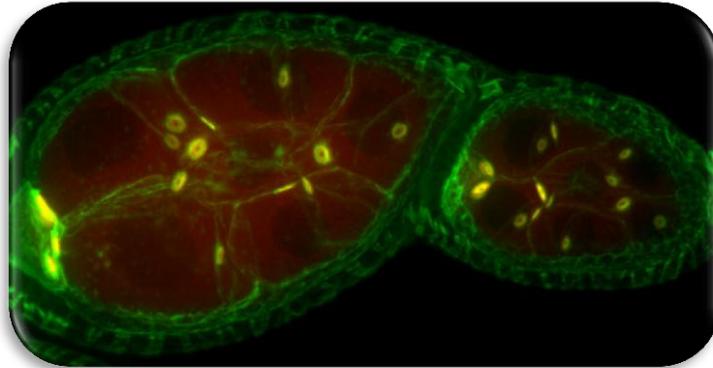


Figura 2 – Esquema de um ovaríolo de *Drosophila*. Canais anel corados a amarelo, e a actina corada a verde. Adaptada de Hudson e Cooley, 2010

A síntese de RNA encontra-se confinada às células companheiras (Cha, Koppetsch e Theurkauf, 2001), sendo que o RNA produzido por estas é transportado activamente (Arn e Macdonald, 2001) através dos canais anel para o oócito. O citoesqueleto torna-se parte fundamental no transporte dos mRNA (Bashirullah, Cooperstock e Lipshitz, 1998) das células companheiras para o citoplasma do oócito (Cooley e Theurkauf, 1994; Malgorzata e Etkin, 2005), apresentando-se a actina como um factor importante na manutenção da polaridade do transporte durante as fases finais da oogénese (Cooley *et al.*, 1992; Watson *et al.*, 1993). Se o sistema de microtúbulos for destabilizado ou mesmo interrompido, por exemplo, através de mutações que impeçam os microfilamentos de actina de revestir os canais anel, os mRNA provenientes das células companheiras são distribuídos aleatoriamente ou deixam de ser transportados. (Wang e Hazelrigg, 1994). Pensa-se que alguns dos compostos transportados das células companheiras para o oócito o façam em associação a proteínas envolvidas no transporte de outras moléculas, tais como a dineína (Schnorrer, Bohmann e Nüsslein-Volhard, 2000; Tekotte e Davis, 2002) e a quinesina (Mische *et al.*, 2007; Tekotte e Davis, 2002), permitindo que os compostos viajem ao longo da rede de microtúbulos (Nasmyth e Jansen, 1997; Saxton, 2001) disponível nas células (Theurkauf *et al.*, 1992). Deste modo, o citoesqueleto controla os movimentos dos organelos e do RNA entre as células companheiras e o oócito de tal forma que as trocas são unidireccionais e específicas (Gilbert, 2003).

Embriogénese

A etimologia da palavra embriogénese, deriva da junção das palavras embrião e génese, ou seja, é o processo através do qual o embrião é formado e se desenvolve. Este processo, contrariamente à oogénese ocorre no exterior da *Drosophila* normalmente sobre fruta ou alimentos em putrefacção, nos quais os ovos fecundados são depositados.

Clivagem

Durante os estágios iniciais da embriogénese, o núcleo zigótico divide-se sem que ocorra a formação de células, dando origem a um embrião denominado de blastoderme sincicial (Johnston e Nüsslein-Volhard, 1992). O núcleo do zigoto sofre várias divisões mitóticas no centro do ovo (Figura 3, 1-6), dando origem a 256 núcleos que depois de formados migram para a zona periférica do ovo (Figura 3, 7-8), onde a mitose continua.



Figura 3 – Clivagem sincicial no embrião de *Drosophila*. Primórdios das células do pólo circunscritos a azul. Adaptado de Gilbert, 2003

No nono ciclo de divisão (Figura 3), aproximadamente cinco núcleos vão localizar-se na superfície da região posterior do embrião, ficando delimitados por membranas celulares. Estes núcleos dão origem às células do pólo que por sua vez vão formar os gâmetas do insecto adulto. Entre os ciclos 10-13 (Figura 3) os núcleos resultantes das sucessivas divisões partilham um citoplasma comum, designando-se o embrião de blastoderme sincicial. Contudo, apesar de partilharem o mesmo citoplasma este não se apresenta uniforme: cada núcleo dispõe

de um “nicho” próprio de proteínas do citoesqueleto, formado aquando da chegada dos núcleos à periferia do ovo durante o décimo ciclo de divisão. O conjunto do núcleo e do seu respectivo “nicho” proteico, denomina-se de energídio. Por fim, após a 13ª divisão mitótica, a membrana celular do oócito dobra-se sobre os núcleos e intercala-se entre estes, dando origem à individualização dos núcleos somáticos. O blastoderme, antes sincicial, dá lugar a um blastoderme celular no qual todas as células que o constituem (aproximadamente 6 000) se encontram organizadas numa monocamada em torno do centro vitelino do ovo (Gilbert, 2003). A transcrição nos núcleos tem início na 11ª divisão e aumenta fortemente na 14ª divisão com o processo de celularização (Figura 3). A par do aumento da transcrição de RNA, ocorre o abrandamento da divisão nuclear conhecendo-se esta fase como transição da blástula intermédia.

Gastrulação

A gastrulação tem início no momento da transição da blástula intermédia e consiste na segregação dos primórdios dos futuros tecidos internos, a mesoderme e a endoderme, para o interior do embrião em desenvolvimento (Leptin, 1999; Leptin e Affolter, 2004). O plano geral do corpo da mosca adulta é igual no embrião, na larva e no adulto: duas extremidades distintas, a cauda e a cabeça, entre as quais se encontram unidades segmentadas que se repetem. Cada um destes segmentos possui uma identidade própria, que é adquirida quando é estabelecido o padrão do corpo (Alberts *et al.*, 2002). Os eixos anterior-posterior e dorsal-ventral de *Drosophila* formam-se perpendicularmente um ao outro e são ambos determinados aquando da oogénese, pela posição do oócito relativamente às células foliculares do ovário (Gilbert, 2003).

Ciclo de Vida

O período de desenvolvimento da *Drosophila melanogaster* varia com a temperatura a que ocorre, podendo ir desde 7 dias (28°C) até 50 dias (12°C). As temperaturas elevadas inibem o desenvolvimento, devido ao stresse criado no embrião pelo calor. A curta duração do ciclo de vida deste insecto representa mais uma vantagem para a sua utilização em ensaios laboratoriais.

Ao longo do seu ciclo de vida (Figura 4), a *Drosophila* passa por uma fase de ovo (0,5 mm) durante a qual ocorre a embriogênese. Passadas 24 horas eclode a primeira forma larvar ou *instar*, seguindo-se outras duas formas larvares por mudança de cutícula. Durante as fases larvares, as larvas alimentam-se dos microorganismos que decompõem a fruta e do próprio açúcar desta. Na terceira forma larvar, a mosca aumenta significativamente de tamanho (4 mm) e passa de um estado móvel para um estado estático através da segregação de uma cutícula espessa que vai formar um casulo denominado de pupa. É durante a fase de pupa, que ocorre a metamorfose: todos os tecidos larvares são degradados e são formadas as estruturas do indivíduo adulto. Ao fim de, aproximadamente, 5 dias emerge da pupa o indivíduo adulto sem pigmentação e com as asas retorcidas. Contudo, bastam apenas algumas horas para que a mosca adulta adquira a cor característica e as suas asas distendam totalmente. E, apenas 12 horas após a emergência da pupa, os indivíduos adultos atingem a maturidade sexual (Projecto *Drosophila* em <http://drosophila-m.blogspot.com/2007/11/ciclo-de-vida.html>; Fonseca, P.M.S., 1997). As fêmeas podem depositar até 400 ovos/embriões, cerca de 5 de cada vez, na fruta em putrefacção ou outros produtos alimentares em degradação. O tempo médio de vida das fêmeas é de 26 dias, enquanto nos machos é de 33 dias.

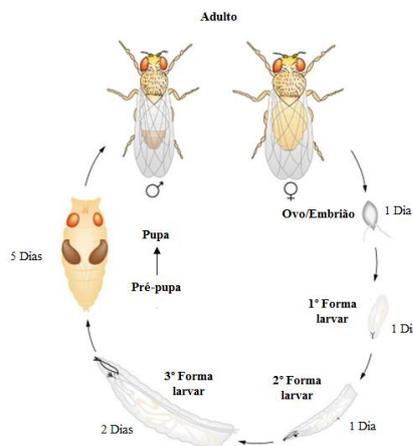


Figura 4 – Esquema do ciclo de vida da *Drosophila melanogaster*. Adaptado de Griffiths *et al.*, 2000

Localização de mRNA

Desde os primeiros estudos de embriologia experimental, sabe-se que a distribuição assimétrica (Huynh e St. Johnston, 2004) de substâncias no citoplasma do ovo (Mohr e

Richter, 2001) confere destinos específicos às células que recebem este citoplasma (Bashirullah, Cooperstock e Lipshitz, 1998). A localização de transcritos em regiões específicas da célula ocorre em todos os tipos celulares e tem diversas funções, desde o controlo da formação dos eixos do corpo até ao processo de aprendizagem e memória (Palacios, 2007; St. Johnston, 2005). De facto, a marcação incorrecta de mRNA leva a distribuições aberrantes da proteína na célula, interferindo com as vias regulatórias celulares normais. Alguns dos mRNA que codificam reguladores conhecidos da polaridade e divisão assimétrica da célula actuam também como supressores de tumores (Lécuyer *et al.*, 2007).

Em 1983, surgiu a primeira evidência da localização intracelular específica de mRNA: ficou demonstrado que o mRNA de actina se encontrava desigualmente distribuído nos embriões de ascídios (Jeffrey, Tomlinson e Brodeur, 1983; Wilhelm e Vale, 1993), uma classe de animais marinhos. O transporte de mRNA é mais eficiente que o transporte de proteínas: um único mRNA pode ser traduzido várias vezes depois de ser localizado, produzindo elevadas concentrações de proteína localizada (Wilhelm e Vale, 1993). Assim, a localização é um mecanismo que permite às células polarizadas restringirem a distribuição de uma proteína a um domínio citoplasmático específico, sendo possível enumerar algumas características comuns a todos os sistemas:

- 1) Os elementos *cis* de localização de mRNA (Jansen, 2001; Kindler *et al.*, 2005) encontram-se localizados, geralmente, na região 3' não traduzida (3'UTR) do mRNA (Bashirullah, Cooperstock e Lipshitz, 1998; Deshler *et al.*, 1998; Kuersten e Goodwin, 2003; Wilkie, Dickson e Gray, 2003);
- 2) O transporte das mensagens localizadas desde o núcleo até aos destinos finais ocorre ao longo de filamentos de actina ou de redes de microtúbulos (Nasmyth e Jansen, 1997; Saxton, 2001);
- 3) Os transcritos são ancorados nas suas regiões de localização através de ligações a elementos do citoesqueleto e de seguida são activados para tradução.

Um dos sistemas mais extensamente caracterizado para o estudo da localização de mRNA é o processo de localização em *Drosophila* (Wilhelm e Vale, 1993; Wilhelm *et al.*, 2000).

Durante a oogénese, as células companheiras sintetizam vários mRNA que são necessários no início da embriogénese e transportam-nos num processo dependente da rede de microtúbulos a locais discretos do oócito (Nasmyth e Jansen, 1997; Wilhelm *et al.*, 2000). A localização assimétrica de mRNA nas células é um dos fenómenos mais importantes da biologia do desenvolvimento, pois nas etapas iniciais do desenvolvimento, estes mRNA vão originar os determinantes localizados que vão especificar o padrão do corpo (Macdonald, Luk e Kilpatrick, 1991). Os microtúbulos são necessários para o movimento de partículas dentro das células companheiras, enquanto que o transporte destas partículas através dos canais anel que ligam as células companheiras ao oócito, aparenta ocorrer de modo independente da rede de microtúbulos e de filamentos de actina (Nasmyth e Jansen, 1997; Saxton, 2001). Contudo, depois das partículas entrarem no oócito, o movimento até ao córtex é feito com participação directa do sistema de microtúbulos, conforme representado na Figura 10 (Theurkauf, W.E. e Hazelrigg, T.I., 1998).

A localização de mRNA permite ainda o estabelecimento de gradientes de proteínas que determinam o mapa de destino celular durante a fase inicial do desenvolvimento (Wilhelm e Vale, 1993).

Mecanismos de Localização de RNA

O facto das regiões 3'UTR dos mRNA a serem localizados (Kuersten e Goodwin, 2003; Wilkie, Dickson e Gray, 2003) serem conservadas (Macdonald, 1990) sugere que os mecanismos de localização também o são. Contudo, os sinais presentes nesta região parecem marcar os transcritos para as respectivas posições na célula por diferentes vias. Como referido anteriormente, os mRNA são transportados das células companheiras para o oócito, tendo que atravessar as pontes intercelulares entre ambos (canais anel). Existem quatro mecanismos de localização de mRNA identificados (Bashirullah, Cooperstock e Lipshitz, 1998; Bullock, 2007; Lipshitz e Smibert, 2000; Tekotte e Davis, 2002; Palacios, 2007; St. Johnston, 1995):

- 1) **Controlo espacial da integridade do mRNA:** consiste na estabilização do mRNA que se encontra correctamente posicionado e na degradação dos transcritos mal localizados;

- 2) **Ancoragem a sítios de ligação localizados:** os mRNA são sequestrados por elementos de ligação localizados;
- 3) **Exportação nuclear vectorial:** também conhecida por, transporte núcleo-citoplasmático (Bashirullah, Cooperstock e Lipshitz, 1998). Tem por base a exportação dos transcritos de um local para o outro da célula;
- 4) **Transporte directo:** mecanismo mais comum de localização de mRNA. O mRNA é transportado ao longo do citoesqueleto depois de associar-se a partículas proteicas.

Formação do Padrão Anterior-Posterior

O desenvolvimento de um organismo multicelular a partir de uma única célula requer, simultaneamente, a determinação de vários tipos celulares e a organização destas células num padrão elaborado. No início do século passado, os embriologistas verificaram que os ovos de vários organismos continham regiões localizadas de citoplasma que controlavam a formação de partes específicas do padrão embrionário (St. Johnston, D. e Nüsslein-Volhard, C., 1992). O desenvolvimento embrionário, requer a constituição de três eixos cruciais que servem de base a todo o organismo, os eixos: anterior-posterior, dorsal-ventral e direita-esquerda (Gilbert, 2003).

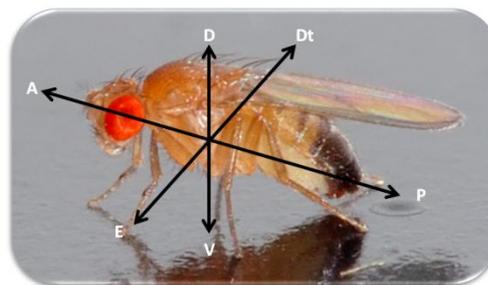


Figura 5 – Esquema em *Drosophila* dos eixos que definem o corpo de um organismo. A – anterior, P – posterior, D – dorsal, V – ventral, E – esquerda e Dt – direita.

Em *Drosophila*, os eixos embrionários são especificados através da localização assimétrica (Huynh e St. Johnston, 2004) de proteínas reguladoras e RNA no oócito em desenvolvimento (Cha, Koppetsch e Theurkauf, 2001). Verificou-se que os genes de efeito materno expressos no ovário da mãe, produzem mRNA que são colocados em diferentes locais do ovo. Por sua vez, estes mRNA codificam proteínas envolvidas na regulação da

transcrição e da tradução que se difundem pela blastoderme sincicial levando à activação ou repressão de genes zigóticos (Gilbert, 2003). Através de estudos genéticos, concluiu-se existirem 4 sinais maternos localizados (Figura 6) que definem a organização básica e a polaridade do eixo embrionário anterior-posterior (AP) (Johnstone e Lasko, 2001; Macdonald, 1990). A caracterização destes sinais impulsionou o estudo de vários fenómenos do desenvolvimento, tais como: determinantes localizados, indução e gradientes morfogénicos.

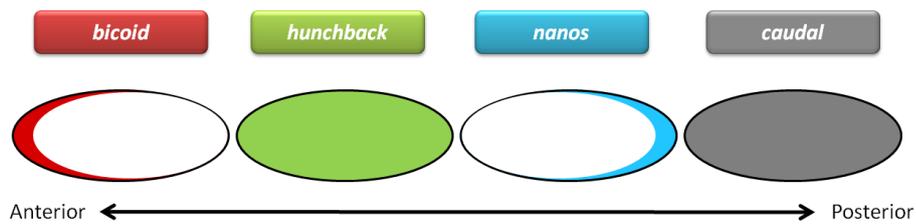


Figura 6 – Distribuição anterior-posterior dos 4 sinais maternos que são depositados no ovo de *Drosophila*.

Assim, estes 4 mRNA são fundamentais para a formação do eixo anterior-posterior: *bicoid* (Lasko, 1999; Macdonald, 1990; Palacios, 2007) e *hunchback* (Lasko, 1999) são necessários para a formação da cabeça e do tórax; *nanos* e *caudal* fulcrais para a formação dos segmentos abdominais. Tal como representado na Figura 6, o mRNA *bicoid* localiza-se na região anterior do óvulo não fecundado (St. Johnston, 1995), ancorado aos microtúbulos anteriores; o gene *bcd* codifica a respectiva proteína que contém um homeodomínio capaz de iniciar uma série de eventos transcripcionais responsáveis pela formação dos segmentos anteriores do corpo. Na extremidade oposta encontra-se ancorado o mRNA de *nanos*, este gene codifica uma proteína de ligação ao RNA que promove a formação do plano posterior do corpo (Wilhelm e Vale, 1993). Os mRNA de *hunchback* (factor de transcrição induzido pela cascata *bcd*) e *caudal* encontram-se uniformemente distribuídos por todo o oócito (Gilbert, 2003; Lasko, 1999; Wilhelm e Vale, 1993). Estas distribuições são executadas no oócito em desenvolvimento, pelos microtúbulos. Como referido, as células companheiras localizadas na extremidade anterior da câmara do ovo sintetizam os mRNA que viajam para o oócito através do citoesqueleto. Neste local, os mRNA são ligados aos microtúbulos por uma série de proteínas motoras (Gilbert, 2003). Depois destas proteínas se encontrarem localizadas, os genes zigóticos por elas regulados são expressos em domínios determinados e parcialmente justapostos e denominam-se genes *gap*. Estes genes constituem o início da actividade transcripcional do zigoto e a expressão das respectivas proteínas cria um gradiente ao longo

dos vários domínios (Schroeder *et al.*, 2004). Assim, o pré-padrão anterior-posterior é formado pela regulação espacial da transcrição dos genes *gap*.

Gradiente Morfogénico

O gradiente morfogénico equivale ao sinal posicional, segundo a teoria da informação posicional proposta por Wolpert em 1969. De acordo com o conceito de informação posicional, a concentração de morfogénio informa as células acerca da sua posição num campo embrionário. Cabe às células, interpretarem esta informação através de um programa de diferenciação apropriado (Wolpert, 1969). Em 1991, Slack definiu morfogénio como um factor de indução capaz de invocar mais do que uma resposta positiva a partir do tecido em que se encontra (Slack, 1991). A existência de gradientes no desenvolvimento dos organismos é desde há muito considerada como um mecanismo para a formação de diversidade espacial a partir de estágios, aparentemente uniformes. Verificam-se em vários modelos animais (ouriços do mar, anfíbios e insectos) que as diferentes propriedades celulares ao longo dos eixos embrionários eram alteradas de um modo quantitativo e não qualitativo, sendo a melhor explicação para esta constatação uma alteração gradual da concentração de uma substância morfogénica (Sander, 1976). Através de ensaios de embriologia experimental e de análise genética, verificou-se que o padrão antero-posterior em *Drosophila* é determinado por dois gradientes opostos, com origem nos pólos anterior e posterior do ovo, respectivamente (Driever e Nüsslein-Volhard, 1988; Wilhelm *et al.* 2000). Assim, os eixos do embrião de *Drosophila* são estabelecidos durante a oogénese à medida que diferentes mRNA e proteínas são localizados nos pólos do oócito em desenvolvimento (Theurkauf e Hazelrigg, 1998).

Para que um sistema materno actue na especificação de parte do padrão embrionário, tem de corresponder a dois critérios: alguns componentes do sistema têm de estar localizados de modo a providenciarem um sinal inicial assimétrico (Mohr e Richter, 2001); e este sinal tem de levar à produção (directa ou indirectamente) de um factor de transcrição activo que regule os genes zigóticos alvo (St. Johnston, D. e Nüsslein-Volhard, C., 1992). No caso do sistema anterior, os critérios necessários para desempenhar o papel de sistema maternal, são cumpridos pelos produtos do gene *bcd* (Wilhelm e Vale, 1993). A sua tradução ocorre após fertilização do ovo e resulta na produção de um gradiente de concentração da proteína *bcd* que

se estende da região anterior para a região posterior. Pensa-se que este gradiente surja por difusão a partir da fonte localizada anteriormente: à medida que o número de cópias maternas do gene *bcd* aumenta, aumenta também a produção de RNA e proteína, resultando na expansão do gradiente da proteína *bcd* em direcção à extremidade posterior. Assim, as estruturas anteriores formam-se em regiões de elevada concentração de *bcd*, enquanto baixas concentrações desta proteína levam ao desenvolvimento de elementos do padrão posterior. A presença de um homeodomínio na sequência de *bcd* sugere que esta proteína se liga ao DNA, ou seja, a *bcd* determina o padrão anterior através da regulação directa dos genes zigóticos alvo. O sistema posterior difere significativamente do sistema anterior. Apesar do sinal inicial localizado no pólo posterior ser uma RNA maternal, o produto deste RNA não regula directamente a expressão de genes zigóticos. Em alternativa, o sistema posterior actua impedindo a tradução de um factor de transcrição codificado por uma RNA maternal ubíquo, o gene *caudal* (*cad*). Assim, ao contrário da proteína *bcd* que desempenha um papel instrutivo na formação do padrão anterior, o sinal posterior desempenha somente um papel permissivo. (St. Johnston, D. e Nüsslein-Volhard, C., 1992).

Na Figura 7, constata-se que no pólo anterior, o mRNA *bcd* é traduzido na proteína *bcd*, formando assim um gradiente, com o máximo de concentração na região anterior (Driever e Nüsslein-Volhard, 1988; Micklem, 1995).

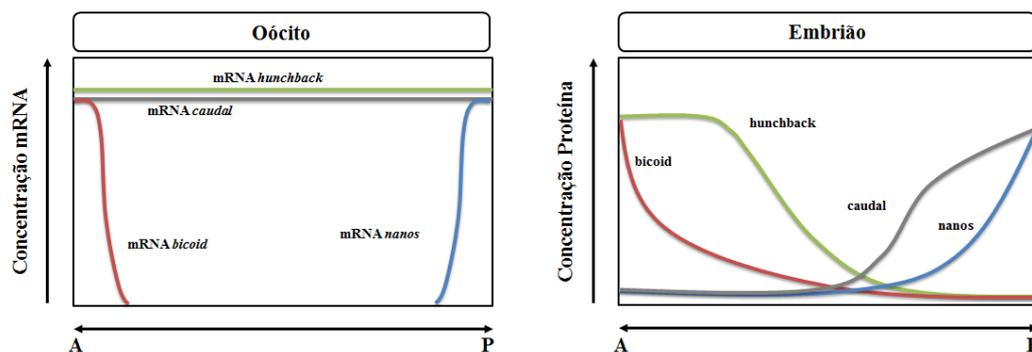


Figura 7 – Variação da concentração ao longo do eixo AP: de mRNA (no óócito, à esquerda) e de proteína (no embrião, à direita). Adaptada de Gilbert, 2003

No pólo oposto, o mRNA *nanos* é traduzido na respectiva proteína, que forma também um gradiente com concentração máxima na extremidade posterior. A regulação genética que leva à formação dos gradientes encontra-se esquematizada na Figura 8: a proteína *bcd* inibe a

tradução do mRNA *caudal*, por interacção com a região 3'UTR de *cad* (Bashirullah, Cooperstock e Lipshitz, 1998), levando a que a proteína *cad* seja apenas traduzida na parte posterior da célula. Antagonicamente, a proteína *nanos* em conjunto com a proteína *pumilio*, liga-se ao mRNA *hunchback*, impedindo a sua tradução na parte posterior do embrião num mecanismo de ligação de mRNA via controlo espacial da integridade do mRNA (Lasko, 1999; St. Johnston, 1995). O resultado destas interacções é a criação de 4 gradientes proteicos no embrião. As proteínas *bcd*, *hb* e *cad* são factores de transcrição cujas concentrações relativas conseguem activar ou reprimir determinados genes zigóticos. O zigoto inicia desta forma, a sua actividade transcripcional (Gilbert, 2003).

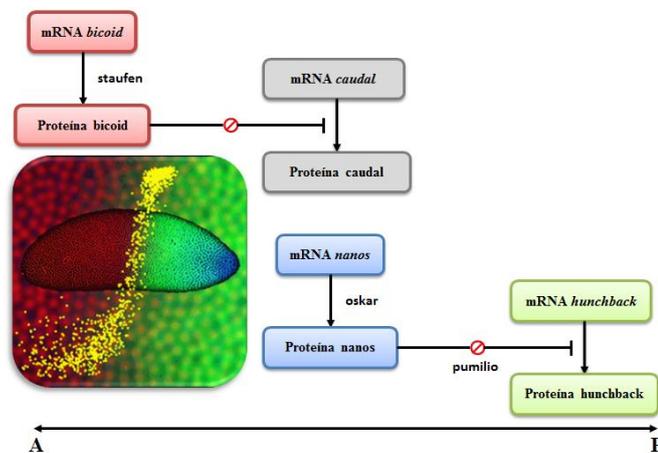


Figura 8 – Regulação genética que leva à formação do padrão anterior-posterior em *Drosophila*. Adaptada de Gilbert, 2003

Como foi referido anteriormente, a formação do gradiente antero-posterior do morfogénio *bcd* (Mickleth, 1995) ocorre em consequência da pré-localização do mRNA *bcd* no pólo anterior do oócito. Esta localização requer pelo menos 3 genes, *swallow* (*sww*), *staufen* (*stau*) e *exuperantia* (*exu*) (Bashirullah, Cooperstock e Lipshitz, 1998; Schnorrer, Bohmann e Nüsslein-Volhard, 2000; Wilsch-Bräuninger, Schwarz e Nüsslein-Volhard, 1997) sendo que, o gene *exu* actua no último passo deste processo (Macdonald, Luk e Kilpatrick, 1991); e de uma rede de microtúbulos (Cha, Koppetsch e Theurkauf, 2001). A proteína codificada por este gene, *exu*, colocaliza-se com o mRNA *bcd* durante a fase inicial da localização, quando o mRNA *bcd* se posiciona nas regiões apicais das células companheiras. Contudo, na fase final da localização de *bcd*, quando este mRNA é transportado das células companheiras para a região anterior do oócito, a proteína *exu* deixa de estar colocalizada, não

havendo detecção da expressão de *exu* nos embriões. Pressupõe-se que esta proteína seja vital no estabelecimento da localização anterior do mRNA *bcd*, mas não participa na manutenção desta localização (Bashirullah, Cooperstock e Lipshitz, 1998). A *exu* possui 2 domínios de fosforilação (pela cinase de serina/treonina Par-1), que influenciam directamente a localização do mRNA *bcd*: a fosforilação destes locais activa a *exu* para mediar a localização deste mRNA (Riechmann e Ephrussi, 2004). Deste modo, verifica-se que o gene *exu* é iniciador do processo de localização do mRNA *bcd*: o padrão anterior do embrião de *Drosophila* depende da localização do mRNA de *bcd* no pólo anterior do oócito em desenvolvimento, e a localização deste mRNA depende do gene *exu* (Macdonald, Luk e Kilpatrick, 1991) e de um citoesqueleto de microtúbulos intacto (Theurkauf e Hazelrigg, 1998; Wang e Hazelrigg, 1994;).

Apesar de inicialmente pensar-se que a proteína *exu* não se ligava directamente ao mRNA *bcd* (Macdonald, Luk e Kilpatrick, 1991) verificou-se mais tarde que o mRNA *bcd* contém uma sequência na sua extremidade 3'UTR (Wilhelm e Vale, 1993) que interage directamente com a proteína *exu*, denominada de elemento de ligação 1 (BLE1 – *Bicoid Localization Element 1*). Este, por sua vez, medeia a ligação deste mRNA ao motor molecular dineína (Januschke *et al.*, 2002; Tekotte e Davis, 2002) que é mantida no centro organizador do microtúbulo (a extremidade (-)) que está ancorado no lado anterior do oócito (Cha, Koppetsch e Theurkauf, 2001). A extremidade (+) do microtúbulo está projectada para o pólo posterior (Wilhelm e Vale, 1993). Enquanto a mensagem *bicoid* é ligada à extremidade ancorada dos microtúbulos via uma proteína motora dineína (motor molecular dirigido à extremidade (-)), os determinantes posteriores são transportados através da proteína motora quinesina I (Januschke *et al.*, 2002; Tekotte e Davis, 2002). Esta proteína move-se em direcção à extremidade (+) do microtúbulo (Gilbert, 2003; Wilhelm e Vale, 1993). A quinesina I vai ligar o mRNA *oskar* (*osk*) (Gunkel *et al.*, 1998; Lasko, 2003; Munro *et al.*, 2006) e a proteína *stau* (Irion e St. Johnston, 2007; Irion *et al.*, 2006). A *stau* permite a tradução da mensagem *osk* (Dollar *et al.*, 2002), sendo a proteína *oskar* resultante capaz de ligar a mensagem *nanos* (Kugler e Lasko, 2009). Deste modo, o gene *stau* é essencial para se dar início ao transporte do mRNA *oskar* (Bashirullah, Cooperstock e Lipshitz, 1998; Lin *et al.*, 2006). No final da oogénese, o mRNA *bicoid* fica ancorado à extremidade anterior do oócito e o mRNA *nanos* fica ligado à extremidade posterior (Kugler e Lasko, 2009). Após fecundação, estes mRNA podem ser traduzidos a proteínas.

Complexo Ribonucleoproteico

A primeira sugestão de que o RNA poderia ser transportado através de uma grande partícula ribonucleoproteica (RNP) surgiu com base no estudo do RNA de BC1, um transcrito de 152 pares de base (pb) de polimerase tipo III que se encontra localizado nas dendrites (prolongamentos do corpo celular dos neurónios) de mamíferos (Wilhelm e Vale, 1993), e tem como principal função a repressão da tradução ao impedir o recrutamento da pequena unidade ribossomal para o RNA mensageiro (Wang *et al.*, 2002). Contudo, o RNA de BC1 não é passível de tradução, ao contrário da maioria dos RNA localizados, o que levantou a questão se a formação de RNP seria ou não um requerimento universal para o mecanismo de localização de RNA (Wilhelm e Vale, 1993). O surgimento de novas técnicas de biologia molecular e a sua aplicação no estudo da localização de mRNA permitiram verificar, através de ensaios de hibridação *in situ*, que os mRNA localizados apresentam um padrão granular no citoplasma (Ainger *et al.*, 1993), confirmando o papel inequívoco das partículas RNP na localização de mRNA de acordo com uma via de localização de mRNA por transporte directo (Johnstone e Lasko, 2001; St. Johnston, 1995). Sabe-se hoje que o mRNA de *bcd* se associa a partículas RNP (Figura 9, passo 1), e que esta ligação é, provavelmente, mediada pela proteína exu (Cha, Koppetsch e Theurkauf, 2001; Wilhelm *et al.*, 2000). Esta proteína actua como factor *trans* (Kindler *et al.*, 2005) ao interagir com a região 3'UTR do mRNA *bcd*, o factor *cis* (Jansen, 2001; Wilsch-Bräuninger, Schwarz e Nüsslein-Volhard, 1997).

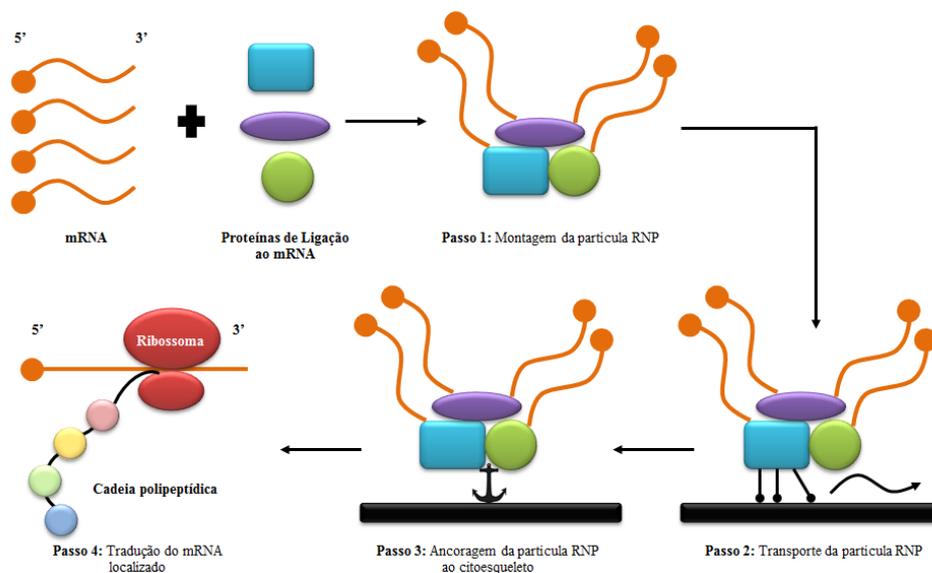


Figura 9 – Modelo do mecanismo de localização de mRNA. Adaptado de Wilhelm e Vale, 1993

Depois de formado, o complexo RNP-mRNA tem de ser transportado das células companheiras para a região anterior do oócito. Este transporte é feito através de filamentos do citoesqueleto por proteínas motoras (Ainger *et al.*, 1993; Bullock, 2007), como representado no passo 2 da Figura 9. A participação de elementos do citoesqueleto (Figura 10) foi devidamente comprovada através de ensaios com destabilizadores destes elementos (colchicina, citocalasina, entre outros) verificando-se a interrupção do transporte destas partículas (Mische *et al.*, 2007).

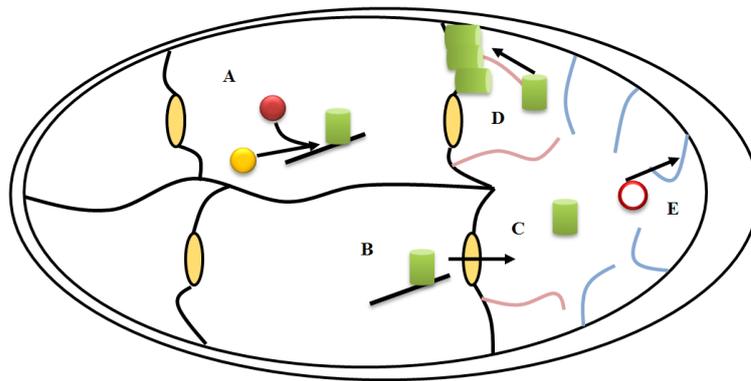


Figura 10 – Modelo da localização do mRNA *bcd* dependente da rede de microtúbulos. A - no citoplasma das células companheiras o mRNA *bcd* forma partículas RNP (a vermelho) em associação com proteína exu (a amarelo), sendo esta associação mediada por microtúbulos; B - estas partículas são transportadas pelos canais anel ao longo de uma via independente de microtúbulos; C - ao entrarem no oócito, as RNP encontram uma rede de microtúbulos organizada por todo o córtex, sendo transportadas por elementos dos microtúbulos (a rosa); D - acumulando-se na zona anterior do oócito; E - o mRNA *bcd* que não teve proveniência no citoplasma das células companheiras (contorno a vermelho) é transportado por todos os elementos da rede produzindo uma distribuição cortical apolar. Adaptada de Cha, Koppetsch e Theurkauf, 2001 e Saxton, 2001

De forma a assegurar-se uma distribuição restricta de proteína, é reprimida a tradução do mRNA que é transportado, sendo activada quando este atinge a sua localização. Quando as partículas RNP chegam ao seu destino final, são ancoradas ao citoesqueleto (Figura 9, passo 3) através de um mecanismo independente de microtúbulos e do transporte citoplasmático (Mohr e Richter, 2001; Wilhelm e Vale, 1993). Pensa-se que a proteína sww (Schnorrer *et al.*, 2002) medeia esta ancoragem (Schnorrer, Bohmann e Nüsslein-Volhard, 2000; Wilsch-Bräuninger, Schwarz e Nüsslein-Volhard, 1997). Por fim, depois de localizado, o mRNA é finalmente traduzido na respectiva proteína (Figura 9, passo 4). A dependência da tradução de mRNA da localização correcta do mRNA a ser traduzido é deveras importante, pois previne a

expressão de determinada proteína fora do seu “nicho” o que poderia resultar em efeitos danosos para a embriogénese (Wilhelm e Vale, 1993).

Em 1994, Wang e Hazelrigg, através de ensaios com um gene quimérico que codificava uma fusão entre a proteína verde fluorescente (GFP, *Green Fluorescent Protein*) *Acquorea victoria* e a proteína exu, conseguiram visualizar o mapa de localização celular desta última. Com base nesses estudos, falaram pela primeira vez na possível existência de um complexo ribonucleoproteico, ou vesículas, dos quais a proteína exu fazia parte, e que transportaria o mRNA *bcd* (Bashirullah, Cooperstock e Lipshitz, 1998; Macdonald, 1990) ao longo de microtúbulos, marcando-o para ser enviado para o córtex anterior do oócito (Theurkauf e Hazelrigg, 1998). Na altura ainda não era claro se o papel da proteína exu seria formar estas partículas ou mediar a sua localização subcelular. Contudo, estes investigadores não acreditavam que a exu estivesse envolvida na ligação directa ao mRNA de *bcd*. Supuseram que outra proteína existente no mesmo complexo ribonucleoproteico se ligasse directa e especificamente ao mRNA *bcd* (Wang e Hazelrigg, 1994). Através de ensaios de microscopia electrónica com marcação por anticorpos também se constatou que a proteína exu fazia parte de umas estruturas de elevada densidade electrónica, denominadas de corpos esponjosos. Após várias experiências, concluiu-se que estes corpos serviriam de “oficina” para a montagem dos complexos RNP-mRNA, ou seja, constituiriam o compartimento intracelular necessário para a ligação dos elementos *cis* e *trans*. (Nakamura, *et al.*, 2001; Wilsch-Bräuninger, Schwarz e Nüsslein-Volhard, 1997).

Poucos anos mais tarde, investigadores conseguiram isolar bioquimicamente a proteína exu e demonstraram que esta fazia parte de um complexo sensível a RNase e que este último seria composto por, pelo menos, outras seis proteínas. Uma das proteínas identificadas como constituinte deste complexo, foi a proteína Ypsilon Schachtel (*yps*), uma proteína Y-box (Mansfield, Wilhelm e Hazelrigg, 2002; Wilhelm *et al.*, 2000)

Ypsilon Schachtel

As proteínas Y-box são uma família de proteínas de ligação ao DNA altamente conservada desde bactérias a humanos. Estas proteínas têm a capacidade de se ligarem a uma

sequência específica de DNA (CGTATTGGCCAA), conhecida como elemento Y-box; que faz destas proteínas reguladoras multifuncionais da expressão genética (Matsumoto e Wolffe, 1998). A maioria das proteínas Y-box não apresenta homologia significativa nas regiões C-terminal, no entanto nos vertebrados, esta região das proteínas contém um elevado número de resíduos de arginina (Matsumoto e Bay, 2005). A primeira proteína Y-box a ser clonada, foi isolada a partir de células-B humanas, ficando conhecida como YB-1. Desde então, muitas outras proteínas Y-box foram identificadas noutros organismos: *Xenopus*, rato, galinha, coelho, vaca, moluscos marinhos, etc. Verificou-se que todas as Y-box possuem uma região com aproximadamente 70 resíduos de aminoácidos que apresenta elevada homologia com as proteínas bacterianas *cold-shock*. Como tal, esta região foi denominada de domínio *cold-shock* (CSD – *Cold Shock-Domain*) (Matsumoto e Wolffe, 1998). A estrutura tridimensional da região CSD, foi determinada por espectroscopia de Ressonância Magnética Nuclear (RMN) e por cristalografia de raios-X: consiste em 5 folhas β ordenadas antiparalelamente, formando uma estrutura em barril- β .

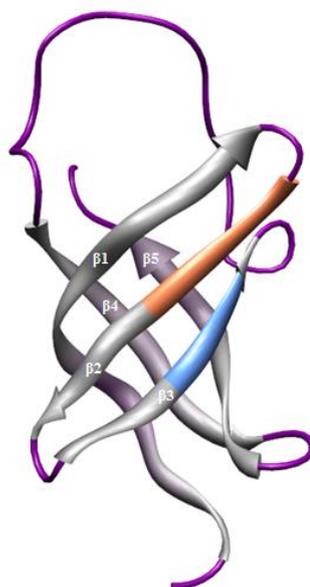


Figura 11 – Estrutura tridimensional do domínio CSD. A sequência de aminoácidos deste domínio foi obtida a partir do Protein Data Bank (PDB) no endereço <http://www.pdb.org/pdb/home/home.do>. O ficheiro PDB foi manipulado no programa Chimera (Pettersen *et al*, 2004). Motivos RNP-1 e RNP-2 assinalados a laranja e a azul, respectivamente.

Dentro do domínio CSD existem 2 elementos de ligação ao RNA já identificados, os motivos RNP-1 e RNP-2 (Thieringer, H.A., 1997). Estes domínios encontram-se presentes em

muitas proteínas de ligação ao RNA e constituem motivos octaméricos (RNP-1) e hexaméricos (RNP-2), conservados nesta região (Landsman, D., 1992; Matsumoto e Wolffe, 1998). Em todas as proteínas Y-box eucariotas, a região CSD localiza-se na extremidade N-terminal.

Em *Drosophila* foi também identificada uma proteína Y-box. O gene que codifica esta proteína foi isolado através de reacções de PCR com *primers* degenerados desenhados para hibridarem na região CSD. O fragmento amplificado foi isolado e clonado no vector pUC19. A sequenciação do cDNA resultante permitiu a identificação da ORF (*Open Reading Frame*) que contém o CSD, que se estende da posição 213 à posição 1 289, sugerindo que codifica para uma proteína de 359 aminoácidos com um peso molecular previsto de 39 kDa. O gene desta ORF foi denominado de *ypsilon schachtel*, e verificou-se estar localizado no braço esquerdo do cromossoma 3, na divisão 69-70a-c, conhecida pela região onde foram mapeados vários genes letais. De modo a estudar-se a expressão de *yps* no desenvolvimento de *Drosophila*, fizeram-se vários ensaios de protecção contra RNase (RPA – *Rnase Protection Assay*) (Thieringer *et al.*, 1997). Esta técnica permite a determinação dos níveis de mRNA de um gene de interesse em diferentes tecidos e estágios do desenvolvimento em função do tempo (Emery, 2007). Os transcritos de *yps* foram detectados em todos os estágios do desenvolvimento, incluindo em embriões de 2-3h, embriões de 14-22h, larvas no 3º estágio larvar e moscas adultas. Os níveis de expressão mais significativos verificaram-se em moscas adultas com um valor aproximadamente 3 vezes superior aos outros estágios estudados (Figura 12).

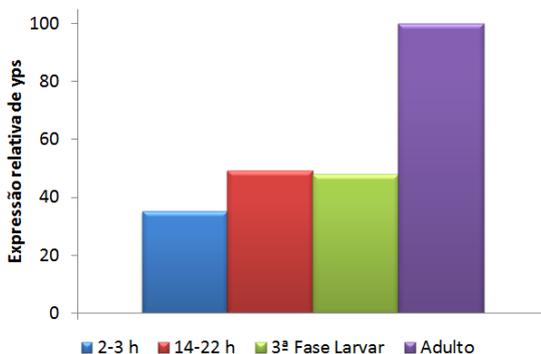


Figura 12 – Expressão relativa da proteína *yps* em vários estágios do desenvolvimento de *Drosophila*. Adaptada de Thieringer, 1997

A hibridação *in situ* deste gene em embriões de *Drosophila* nos estágios iniciais do desenvolvimento permitiu a observação do padrão de expressão do gene *yps* (Figura 13). A presença de transcritos de *yps* no estágio 2 é clara, ou seja, é possível concluir-se que estes são depositados no citoplasma do ovo antes de este ser fertilizado. Verifica-se pelas imagens que o mRNA de *yps* se encontra distribuído por todo o citoplasma embrionário e a sua expressão também se verifica no tecido mesodermal (Thieringer, 1997).

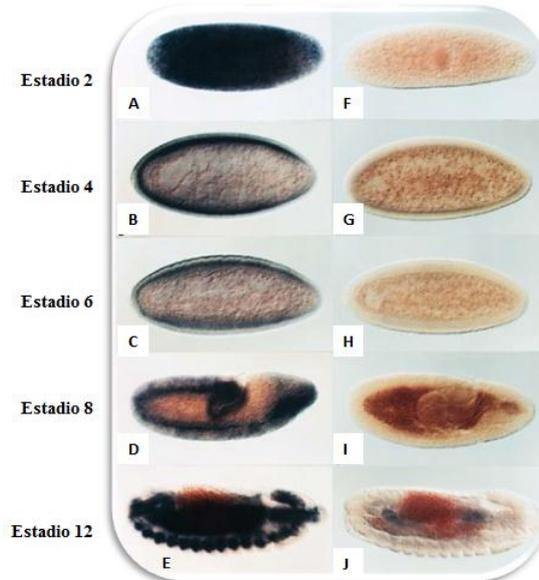


Figura 13 – Padrão de expressão do gene *yps* no início da embriogénese. Embriões de A-E foram hibridados com sondas de RNA *antisense*; os embriões F-J serviram de controlo tendo sido hibridados com sondas *sense* de RNA marcadas com digoxigenina, nas mesmas condições que os embriões anteriores. Adaptada de Thieringer, 1997

Por último, construiu-se um plasmídeo com a região codificante do CSD de *yps* e procedeu-se à sua sobre-expressão num sistema bacteriano. A proteína expressa foi purificada por cromatografia de troca iónica com um rendimento de 95%. A análise da proteína pura pela técnica de *gel-shift* mostrou que o domínio CSD da proteína *yps* era capaz de se ligar a RNA de cadeia simples, sugerindo que a *yps* poderá estar envolvida em mecanismos de ligação ao RNA. Outra particularidade da proteína *yps* é a sua aparente basicidade, com um ponto isoeléctrico (pI) de 11,2. A sequência que precede o codão de iniciação ATG, AACGATG, corresponde a 5 das 7 posições da sequência *consensus* de iniciação da tradução (C/A)AA(C/A)ATG em *Drosophila*. Na posição 1 289 existe um codão STOP, TAA, seguido de um sinal putativo de poli-adenilação. Existem, ainda, duas regiões ricas em glutamina que

flanqueiam o domínio CSD: na extremidade N-terminal do CSD encontram-se 10 resíduos de glutamina, e na extremidade C-terminal existem repetições de 6 e 5 resíduos do mesmo aminoácido. Outra observação importante, é o facto do nível de transcritos de *yps* não se alterar quando sujeitos a tratamentos a baixas e altas temperaturas, o que poderá indicar que o gene *yps* não é um gene indutível por choque térmico (Thieringer, 1997).

Em 2000, Wilhelm e colaboradores determinaram o tamanho da proteína *yps* a partir de extractos de *Drosophila*, enquanto estudavam as propriedades bioquímicas da proteína *exu*. Ao fazê-lo, constataram que a *exu* apresentava um tamanho bastante superior ao esperado, levando a crer que se deveria à associação da *exu* com outras proteínas e /ou RNA. Optaram então, por tratar o isolado com RNase A e verificaram por ensaios de imunoprecipitação a existência de um padrão de 7 bandas: 57, 74, 76, 78, 82, 88 (proteína *exu*) e 147 kDa. Ao sujeitarem o complexo composto por estas proteínas a um tratamento com elevadas concentrações de RNase, a proteína com o peso molecular de 57 kDa permanecia ligada à *exu*, o que permitiu concluir, que a ligação entre estas proteínas é independente de RNA. Posteriormente, procederam à identificação deste polipeptídeo por microsequenciação e verificaram que se tratava do produto do gene *yps* (Wilhelm *et al.*, 2000).

Este estudo permitiu ainda verificar que no complexo constituído pela *yps* e pela *exu*, também se encontrava presente o mRNA de *osk*. Foi assim alterado o pressuposto de que a *exu* apenas tinha um papel na localização do mRNA *bcd* na região anterior do oócito. No entanto, esta descoberta é consistente com outras evidências: a *exu* e a *yps* acumulam-se em ambos os pólos do oócito; o mRNA de *osk* acumula-se transientemente no pólo anterior juntamente com o mRNA *bcd*, sendo de seguida transportado para o pólo posterior (Lin *et al.*, 2006; Munro *et al.*, 2006); e em mutantes que não expressem a proteína *exu*, a localização do mRNA *osk* é interrompida (Wilhelm *et al.*, 2000). O gene *yps* interage, ainda, com o gene *orb*, um regulador positivo da localização e tradução do mRNA de *osk*. A natureza desta interacção indica que *yps* actua de modo antagónico a *orb*: as respectivas proteínas competem pela ligação ao mRNA *osk* resultando em efeitos opostos na tradução e localização deste mRNA. A proteína *orb*, codificada por este gene, associa-se às proteínas *yps* e *exu*, nos ovários de *Drosophila*, sendo esta associação mediada por RNA (Mansfield, Wilhelm e Hazelrigg, 2002). A proteína *yps* interage directa e indirectamente com outras proteínas, como se encontra esquematizado na Figura 14.

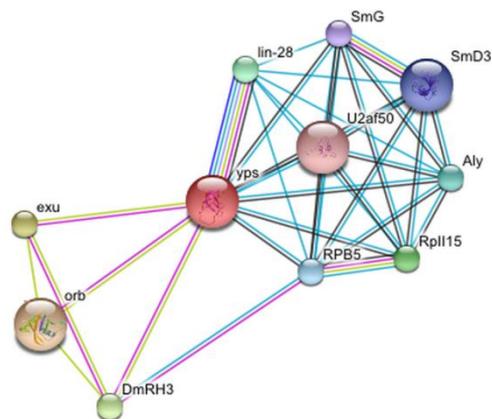


Figura 14 – Interações estabelecidas com a proteína yps. Destacam-se as interações com as proteínas exu e orb. Imagem adaptada da ferramenta online STRING (*Search Tool for the Retrieval of Interacting Genes/proteins*) (Jensen *et al.*, 2008) disponível em http://string-db.org/newstring.cgi/show_input_page.pl.

Considera-se que a yps seja um componente da maquinaria de localização de mRNA, mas o seu envolvimento no transporte e /ou regulação da tradução de mRNA (Mansfield, Wilhelm e Hazelrigg, 2002) no oócito e outros tecidos ainda não se encontra esclarecido. Assim, estudos de clonagem e expressão de yps poderão permitir a determinação da sua estrutura tridimensional, o que ajudará a elucidar acerca do papel desta proteína na localização de mRNA em *Drosophila melanogaster*, bem como compreender melhor o papel biológico das proteínas Y-box.

Depois de compreendida a estrutura e função bioquímica da yps, o próximo passo é a análise da interação desta proteína com os restantes elementos do RNP a que pertence e interação com o mRNA transportado, para melhor compreensão da sua importância neste processo.

Materiais e Métodos

Meios e Tampões

Os meios de crescimento, tampões e soluções utilizados foram preparados conforme consta da tabela que se segue.

Tabela 1 – Composição dos meios de crescimento, tampões e soluções utilizados.

Nome	Composição	Quantidade
Tampão PBS (<i>Phosphate Buffered Saline</i>) 10x (1L em H ₂ O _{mq} , pH=7,4 ajustado com NaOH)	NaCl	80 g
	KCl	2 g
	Na ₂ HPO ₄	14,4 g
	KH ₂ PO ₄	2,4 g
Tampão SDS (<i>Sodium Lauryl Sulfate</i>) 10x (1L em H ₂ O _{mq})	SDS	100 g
Tampão TBE (Tris/Borato/EDTA) 10x (1L em H ₂ O _{mq})	Tris Base	108 g
	Ácido Bórico	55 g
	0.5 M EDTA	40 ml
Solução Corante Gel SDS-PAGE (1L em H ₂ O _{mq})	Coomassie Blue R-250	5 g
	Ácido Acético Glacial	75 ml
	Metanol	450 ml
Solução Descorante Gel SDS-PAGE (1L em H ₂ O _{mq})	Ácido Acético Glacial	75 ml
	Metanol	450 ml
Tampão de Amostra DNA (5 ml em H ₂ O _{mq})	Glicerol	1,5 ml
	Azul de Bromofenol	12,5 mg
Solução II (100 ml, pH=6,8 ajustado com HCl)	Tris Base	6,06 g
Tampão de Amostra Proteína (5 ml em H ₂ O _{mq})	Solução II	5 ml
	10 x SDS	8 ml
	B-mercaptoetanol	1 ml
	Glicerol	2 ml
	Azul de Bromofenol	4 mg

Tabela 1 (continuação)

Nome	Composição	Quantidade
Meio de Crescimento SOB (<i>Super Optimal Broth</i>) (1L em H ₂ O _{mq})	Bactotripton	20 g
	Extracto de Levedura	5 g
	NaCl	0,58 g
	KCl	0,18 g
Meio de Crescimento SOC (<i>SOCatabolite Repression</i>) (1L)	Meio de crescimento SOB	980 ml
	1 M Glucose	20 ml
Meio de Crescimento LB (<i>Lysogeny Broth</i>) (1L em H ₂ O _{mq} , pH=7.5 ajustado com NaOH)	Bactotripton	10 g
	Extracto de Levedura	5 g
	NaCl	10 g
Meio de Crescimento Sólido LB-Agar (1L em H ₂ O _{mq} , pH=7.5 ajustado com NaOH)	Meio de Crescimento LB	1 000 ml
	Agar	15 g
Meio de Crescimento 2xYT (1L em H ₂ O _{mq})	2xYT (Sigma)	31,6 g
Suplemento 50x5052 (1L em H ₂ O _{mq})	Glicerol	250 ml
	Glucose	25 g
	α -Lactose	100 g
Suplemento 20xNPS (1L em H ₂ O _{mq})	(NH ₄) ₂ SO ₄	66 g
	KH ₂ PO ₄	136 g
	Na ₂ HPO ₄	142 g
Solução de Lavagem (1L)	1 x PBS	1 000 ml
	500 mM Imidazol	34,05 g

mq - milli-Q, água destilada que foi purificada num filtro de troca iónica.

EDTA – *EthyleneDiamone Tetraacetic Acid*

Todos os meios de crescimento foram autoclavados antes da adição de antibióticos.

Clone em estudo

O clone correspondente ao gene que codifica para a proteína ypsilon schachtel (*yps*), objecto deste estudo, foi obtido através do *Drosophila Genomics Resource Center* (DGRC). Trata-se de uma instituição que faculta os mais diversos materiais para a investigação em *Drosophila*: clones, linhas celulares, *microarrays*, entre outros. A identificação (ID) do clone utilizado é o código LD37574 da colecção Gold de DNAs complementares (cDNA). Esta colecção está depositada na *Drosophila Gene Collection* (DGC), e foi construída a partir do

Barkeley Drosophila Genome Project (BDGP). A este ID corresponde o ID FBcl0179037 na *Flybase*, uma base de dados de genes e genomas de *Drosophila*. A sequência completa do gene é a abaixo transcrita.

```

5'TCATGTTTCACTTCCGAGCTTGCGAACAAAATAGAATCGCTAAGGCAAAGAAAACCCCTGCGAGTGCGGAGTGTTTAAGA
TCACACTCAAAAAACAACATCCAAAGTGGTTAGTGGCAGTTGAGAGAGTGCCAAGACCCCGAATACCAGCCGGGCAGACAC
GAACAACACCAGCAGCAGCAGCACCACCATCATCAACAGCAGCAGCAGCATTTCCTCGTTTTTCGTTTTCGTCGCGCGAACA
ACAAATAACAAAACCACGACAAACGATGGCTGATGCCGCGGAGAGTAAGCCACTGGCCGCCGAACAGCAGCAGGCGCAGCA
GCAGCCGGAGCAGCAGCAGAATCCGCCGAATCCGCAGGAGCAGGATCACGAGCAGGAGCCGCTGGACGAGCTGCAGGGACA
GCAGGGCCAGCCCGCTCCGCCACCAGGAAGTCAATCCGGCACCAGTCAAGTGGTTCAACGTGAAGAGC
GGCTACGGTTCATCAACCGCAACGACACCAGAGAGGATGTCTTTGTGCACCAGAGCGCCATTGCGCGGAACAACCCCAAAA
AGGCGGTCCGCTCGGTGGGCGACGGTGAGGTCGTTGAGTTCGACGTGGTCATTGGTGAGAAGGCAACGAGGCGGCCAACGT
GACCGGTCCCTCCGGTGAGCCGGTGCGGGGCAATCCAGTTTGCAGCGGACAAGCGCCGTAACCTCCGTCCTGGATGAAGAAG
AATCGCCGCAAGGATGGCGAAGTGAAGGCGAAGACGCCGAATCGTCGGCCAGCAGCAACAGCAGCAGGCGGCACCAATC
GTTGACGGGCAGCCGACGAGCAGGTGCAGTCCGGCCCGCGTCAACCACGACAGAACTTCCGCCGAGGACCACCCGGTGGAC
CGCCCGGAGGACCTCGAGGTGGCCCCGAGGACCGCCTGGTGGAGCTCCCGTGGCCCCGGCGCTACAACAACTATTATCT
GCGTCAGCCGAGACGCGGCCTGGGCGGCGGTGACGGCAGCGCCGAGCCCGGTGTCCACGATCAGAATCCCGAGGGTCTGCA
GCGCGGCGAGGGTTCAGGGACCACGTGCGGTTGGAGGCCACCGGGCGGACCCAGAGGCGCTTCTCCGACGCAACTTCAAC
AATGGTCCACCGCCACCACGCCGACGGCGGAGAATACATTCAGGCCAGGACCGCCACGCCACAACAGCCACGTCCAC
GTCGCCAGAGGAAACCAATGGCCCTGGCGGTGGTTCGGAGCAGCAGCCAGAGAAGAACGGCGCTCAAGAGCTGCAAAATA
CAACCACAGAGAGCACTGCATAGAAAAGGATCCGAAAAGGATCTAAATCGTCAACCGACAAAACAAAACATCGTCAGCAGTA
GCCGAGTATATAACTACAATCGCAGAAAAACAGCAGCAGCAGCCATCCTCAACATCAGCGACCAATGCAACATCATGGTG
CAATAAGATCAGCAAATGTTGCAGCAAATTTTGCTAAATGAGCCGCCGCTTTTTTCGGTTTGTGCTTTTTGTTAGTTTCG
TAGGGCAGAGAATCAATTGTACAACATCATAACAAAACTAACATCCTCAAAATTTCAATCAAATTCAAAAACAAACAA
AAAATTGTATAAAATTTCCATCAAGCCTGGCCGTCGTCAAAAAATTGCATAATTGTGGAATTTTATTGTATGAACCAACAAA
CCAAACACGAAAACATAAATAAATTTCAATTACTGCAAAAAAAAAAAAAAAAAAAAAA3'

```

Figura 15 – Sequência completa do gene *yps*. A azul, sequência de emparelhamento dos *primers* N001 e N001S; a vermelho, sequência de emparelhamento dos *primers* 145C e 145CS; a roxo, local de emparelhamento dos *primers* N55 e N55S.

Esta sequência encontra-se inserida no vector pOT2 (utilizado em 90 % das clonagens dos cDNAs disponibilizados nesta base de dados), entre os locais de restrição 5'EcoRI e 3'XhoI, de acordo com a Figura 16.

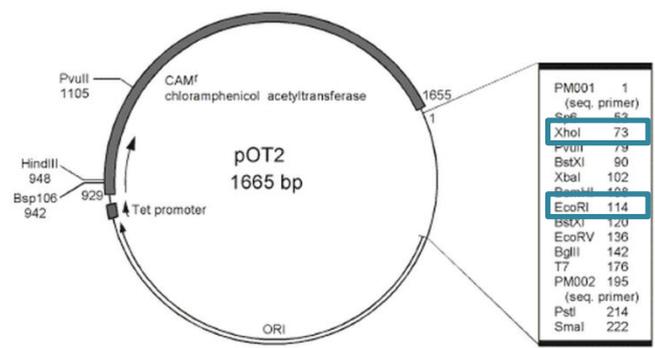


Figura 16 – Esquema do vector pOT2 com os locais de inserção do gene *yps* assinalados.

A proteína yps codificada por este cDNA tem a seguinte sequência:

```
MADAAESKPLAAEQQAQQQPEQQNPPNPQEEDHEQEPLDELQGGQGPAPPTKEVIATKVTGTVKWFNVKSGYGFINRNDTR  
EDVVFVHQSALARNNPKKA VRSVGDGEVVEFDVVIGEKGNAAANVTGPSGEPVVRGSQFAADKRRNFRPVMKKNRRKDGEVEGED  
AESSAQQQQQAAPIVDGQPQQVQSGPRQPRQNFRRGPPGGPPGGPRGGPRGGAPGGPRRYNNYYLRQPRRGLGGGDGSAE  
PGVHDQNPEGLQRGEGQGPRRGGGPPGGPQRRFRFRNFNNGPPPPRRDGGGEYIQGGPPRPQPRRRQRKPNPGGGSEQPEKN  
GAQELQNTTTESTA
```

Figura 17 – Sequência completa da proteína yps

Previsão da Globularidade da Proteína

Pela análise da sequência prevista da proteína yps, observam-se várias zonas altamente repetitivas, em particular na metade C-terminal; o que poderá indicar que esta zona se encontra mais desordenada do que a região N-terminal.

De modo a prever-se qual o fragmento mais estável *in vitro*, ou seja, o que terá maior probabilidade de ser o domínio de ligação ao RNA e outras proteínas, recorreu-se à ferramenta GlobPlot (Linding, 2003) disponível *online* (GlobPlot 2 – *Intrinsic Protein Disorder, Domain & Globularity Prediction* em <http://globplot.embl.de/>). Este programa analisa a tendência da proteína para a ordem/globularidade ou desordem, isto é, prevê regiões não globulares a partir da sequência completa da proteína.

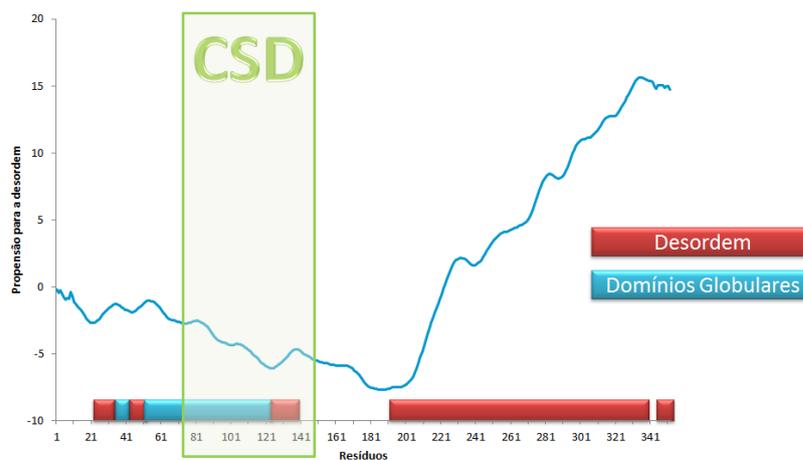


Figura 18 – Representação gráfica das regiões de ordem/desordem da proteína yps obtida através da ferramenta GlobPlot. CSD – *Cold Shock Domain*.

Com base neste gráfico, optou-se pelo estudo de dois fragmentos da proteína: as regiões dos resíduos 1-145 e 55-145, que correspondem às regiões nucleotídicas de 273-705 (fragmento 453 pb) e 432-705 (fragmento 294 pb), pares de base da sequência de cDNA; em detrimento do estudo da sequência na sua totalidade.

De modo a anteciparem-se as propriedades das proteínas resultantes da tradução desses fragmentos de DNA, recorreu-se à ferramenta *ProtParam* (Gasteiger *et al.*, 2005) disponibilizada no servidor ExPASy (*Expert Protein Analysis System*), no endereço <http://www.expasy.org/>. O fragmento de 453 pb que corresponde à região de 1-145 aminoácidos tem um peso molecular aproximado de 15 730,1 Da e um ponto isoelétrico calculado de 4,74; enquanto que o fragmento menor traduz-se numa proteína de 9 664,7 Da com um ponto isoelétrico teórico de 6,81.

Vectores de Clonagem e Expressão

Neste trabalho, foram utilizados quatro vectores de clonagem e expressão, com o intuito de se estudar qual o vector ideal para a expressão e purificação da proteína yps.

Os mapas dos vectores podem ser consultados no Anexo I.

Tabela 2 – Principais características dos vectores utilizados.

Características	pGEX-6P1	pET-14b	pET-15b	pET-SUMO-28a
Tamanho (Kb)	4,900	4,671	5,708	5,369
Promotor	<i>tac</i>	T7	T7 <i>lac</i>	T7 <i>lac</i>
Marca de Seleção	Amp ^r	Amp ^r	Amp ^r	Kan ^r
Local de Clonagem				C-terminal
Caudas de Fusão	GST	His•Tag	His•Tag	His•Tag
	N-terminal	N-terminal	N-terminal	N- e C-terminal
				T7•Ta
				N-terminal
				SUMO
Marca	GE Healthcare	Novagen	Novagen	Novagen

Kb – quilobases; Amp^r – resistência à ampicilina; Kan^r – resistência à kanamicina; GST – Glutathione S-Transferase; His – histidina

Preparação dos Vectores

Os vectores foram obtidos por crescimento em meio 2xYT suplementado com o antibiótico correspondente à marca de selecção de cada plasmídeo a partir de *stocks* em glicerol 50% que se encontravam armazenados a -80°C. Os inóculos foram incubados durante a noite com agitação (150 rpm – rotações por minuto) e a 37°C. Por fim, procedeu-se à purificação dos plasmídeos recorrendo a um *kit* comercial da empresa NZYTech, o NZYMiniprep. Este *kit* permite a preparação rápida em pequena escala de plasmídeos a partir de estirpes recombinantes de *E.coli*: o plasmídeo é selectivamente adsorvido numa coluna de *spin* de sílica e as impurezas, tais como proteínas, sais, nucleótidos e oligonucleótidos são eluídas nos vários passos do protocolo. Os vectores puros foram eluídos em 30 µl do tampão de eluição disponibilizado no *kit*, e armazenados a -20°C até serem utilizados.

Verificação dos Vectores Purificados – Electroforese em Gel de Agarose

A análise dos vectores puros foi feita por electroforese em gel de agarose, uma técnica que tem por base a separação de moléculas de acordo com o seu peso molecular, por aplicação de uma diferença de potencial ao gel onde são depositadas as amostras.

Após terem sido testadas diferentes concentrações de agarose, verificou-se que se conseguia uma visualização óptima das bandas correspondentes aos plasmídeos, a 1% de agarose: 0,9 g de agarose Cloning Grade (NZYTech) em 90 ml de tampão 1 x TBE. De modo a detectar o DNA no gel foram utilizados três agentes corantes ao longo deste trabalho, numa concentração de 0,7x no volume do gel preparado:

- 1) **Brometo de Etídeo (Fluka da Sigma-Aldrich):** intercala a dupla cadeia de DNA e RNA. Forma complexos fluorescentes com os ácidos nucleicos que podem ser observados sobre luz ultravioleta. Contudo, é altamente tóxico, mutagénico e termosensível. Pelo que foi substituído pelo Gel Red (Biotarget);
- 2) **Gel Red (Biotarget):** é também um agente intercalante da dupla cadeia de DNA, mas não mutagénico. Os complexos formados com o DNA, também podem ser

visualizados em transiluminador por excitação com luz ultravioleta. Por fim, utilizou-se, também, o corante SYBR Safe DNA Gel Staining (Invitrogen);

- 3) **SYBR Safe DNA Gel Staining (Invitrogen):** apresenta baixa taxa mutagénica e maior sensibilidade de detecção que os corantes anteriores. Foi o corante que apresentou melhor fluorescência, pelo que foi o corante mais utilizado neste trabalho.

A cada poço do gel, adicionou-se 5 µl de cada vector puro e 3 µl de tampão de amostra de DNA (Tabela 1). O marcador de pesos moleculares utilizado foi o NZYDNA Ladder III (NZYTech) com um padrão de 14 bandas regularmente espaçadas entre os 200 e 10 000 pb. Este marcador, permite ainda a quantificação de DNA pois a cada banda do padrão corresponde uma quantidade precisa de DNA, podendo assim ser determinada a quantidade do DNA em análise comparando-se a intensidade das bandas em estudo com as bandas do padrão. Informação acerca deste padrão pode ser consultada no Anexo II.1.

Correram-se os géis durante 60 min, 90 V e 400 mA. Todos os géis de agarose preparados foram visualizados num transiluminador (Uvitec) de luz ultravioleta.

Amplificação dos Fragmentos

Para se proceder à clonagem de determinado gene são necessárias grandes quantidades do gene individual e puro. A Reacção em Cadeia pela Polimerase (também conhecida por PCR – *Polymerase Chain Reaction*) é uma das técnicas mais poderosas e versáteis para a obtenção dessas quantidades de DNA de uma sequência específica (Lodish, *et al.*, 2008).

Oligonucleótidos Iniciadores – *primers*

Desenharam-se seis *primers* (4 *forward*: N001, N55, N001S e N55S; 2 *reverse*: 145C e 145CS) atendendo aos locais de restrição presentes no local de múltipla clonagem dos plasmídeos e assegurando que estes locais não estavam presentes na sequência dos fragmentos em estudo.

Tabela 3 – Sequências dos *primers* utilizados neste estudo. Locais de restrição assinalados a cor: BamHI, NdeI, BglII, EcoRI, XhoI. Codão STOP assinalado a vermelho.

Vector	Designação	Sequência nucleotídica (5'→3')
pGEX-6P1	N001	GACGGATCCCATATGGCTGATGCCGCGGAG
pET-14b e pET-15b	N55	GACGGATCCCATATGAAGGAAGTCATCGCCACCAAAG
	145C	GTCGAATTCAGATCTTACTTGTCCGCTGCAAACCTGAC
pET-SUMO-28a	N001S	GACCTCGAGGCTGATGCCGCGGAGAG
	N55S	GACCTCGAGAAGGAAGTCATCGCCACCAAAG
	145CS	GTCGGATCCTTACTTGTCCGCTGCAAACCTGAC

Os pares de *primers* N001/145C e N001S/145CS foram desenhados para a amplificação do fragmento de 453 pb; e os pares N55/145C e N55S/145CS foram desenhados para a amplificação do fragmento mais pequeno, de 294 pb. Salienta-se ainda, que os *primers* N001, N55 e 145C possuem dois locais de restrição (sequências não complementares ao cDNA molde) que permitem a utilização dos produtos resultantes da reacção de amplificação, tanto na clonagem em pGEX-6P1 como na clonagem em pET-14b e pET-15b.

Os *primers* foram sintetizados pela empresa Thermo Fisher Scientific.

Simulação da Reacção de Amplificação

Com recurso ao *software* Serial Cloner, foi possível verificar o resultado da reacção de amplificação, tendo em conta os pares de *primers* e a sequência nucleotídica completa do cDNA de *yps*.

Tabela 4 – Tamanho (pb) previsto dos fragmentos de *yps* a amplificar.

Estudo	<i>Primers</i>	Fragmento
pGEX-6P1, pET-14b e pET-15b	N001+ 145C	465
	N55 + 145C	306
pET-SUMO-28a	N001S + 145CS	453
	N55S + 145CS	294

Mistura Reaccional

De modo a otimizar-se a amplificação do DNA desejado, testaram-se várias misturas reacionais até se obter uma amplificação intensa e inequívoca dos fragmentos desejados com o mínimo possível de produtos inespecíficos. A mistura reaccional contém excesso de *primers* e de nucleótidos trifosfatados para que a reacção de polimerização não seja o passo limitante.

Assim, para cada ensaio foram preparadas duas *master mix* (uma para cada fragmento), para um volume final individual de 24,5 µl ao qual se adicionava em último 0,5 µl do cDNA.

Tabela 5 – Volumes (µl) de cada um dos constituintes das *master mix* preparadas.

Composição	N001 ou N001S	N55 ou N55S
H ₂ O _{mq}	14,5	14,5
dNTPs (Fermentas)	1,5	1,5
N001 ou N001S	1,0	
N55 ou N55S		1,0
145C ou 145CS	1,0	1,0
Tampão Taq 10x (NZYTech)	5,0	5,0
MgCl ₂ 50mM (NZYTech)	1,0	1,0
Taq polimerase 5U/µl (NZYTech)	0,5	0,5
TOTAL	24,5	24,5

dNTPs – *deoxyNucleotides TriPhosphates* (dATP, desoxiAdenosina; dCTP, desoxiCitidina; dGTP, desoxiGuanosina e dTTP, desoxiTimidina). U – quantidade de enzima necessária para catalisar a incorporação de 10nmol de dNTPs em 30 min a 37°C.

Programa de Amplificação

A escolha do programa a utilizar para a amplificação dos dois fragmentos em estudo, também foi alvo de vários ensaios até se encontrarem as condições óptimas para o estudo em causa. Foram tidas em conta as temperaturas de fusão (T_m , *melting temperature*) dos *primers* desenhados para a amplificação de ambos os fragmentos, uma vez que estas temperaturas determinam a especificidade de hibridação dos *primers* ao molde. A amplificação foi feita com recurso a um termociclador (*MyCycler Personal Thermal Cycler* da Bio-Rad), tendo-se

optado por um programa *2-step* que consiste em duas temperaturas de fusão diferentes, tal como esquematizado na Figura 19.

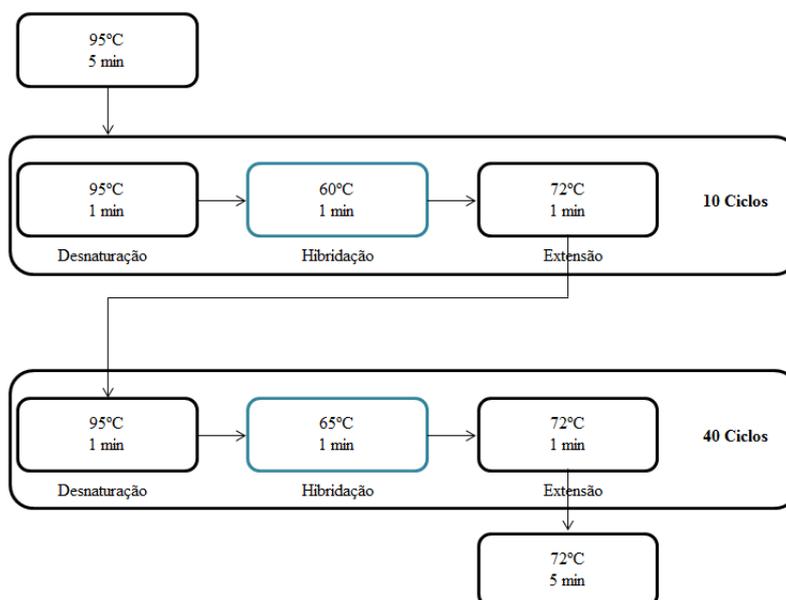


Figura 19 – Esquema do programa utilizado para a amplificação dos fragmentos de *yps*, para os estudos com os vectores pGEX-6P1, pET-14b e pET-15b.

Para o estudo com o vector pET-SUMO-28a utilizou-se um programa idêntico, alterando-se apenas as duas temperaturas de fusão: no primeiro passo foi de 58°C e no segundo passo de 63°C.

Confirmação e Extracção dos Fragmentos Amplificados – Electroforese em Gel de Agarose

O método mais comum de se detectarem os amplicões é por electroforese em gel de agarose. Assim, a confirmação do resultado das reacções de amplificação foi feita por esta técnica, tal como na análise dos vectores após o passo de purificação; tendo-se utilizado as mesmas condições de preparação, corrida e visualização do gel.

A cada poço do gel adicionou-se o volume total de cada reacção de PCR juntamente com 15 µl de tampão de amostra. Sempre que se visualizou um padrão de bandas nos

tamanhos de peso moleculares esperados, as bandas foram excisadas e armazenadas a -20°C para estudos posteriores, ou a 4°C para utilização no próprio dia.

Purificação dos Fragmentos

A purificação dos fragmentos amplificados por PCR foi feita a partir das bandas excisadas do gel de agarose ou directamente em gel, com o auxílio de *kits* comerciais e/ou sistemas automatizados. Calculou-se ainda o rendimento deste passo experimental atendendo ao resultado em gel de agarose antes e depois da purificação. O sinal das bandas foi comparado com a intensidade do sinal das bandas do padrão do marcador de pesos moleculares, e tomou-se como rendimento médio a média dos rendimentos de purificação de cada amostra (Anexo II.1).

Purificação Manual – Em Coluna

Existem vários métodos para a purificação de bandas de DNA extraídas de gel de agarose: tubo de diálise, método da tira de papel e em coluna (Methodbook disponível em <http://www.methodbook.net/dna/gelextrc.html>). Destes três, o método mais célere e eficaz é a purificação da banda de gel em coluna, que tem por base a utilização de uma membrana de sílica que absorve selectivamente fragmentos de DNA na presença dos tampões concebidos para o efeito. Consiste em três passos experimentais típicos: ligação do DNA à membrana imobilizada na coluna, lavagem do DNA e eluição do DNA puro.

Utilizou-se este método sempre que se pretendeu purificar os fragmentos ou vectores a partir de bandas excisadas do gel de agarose, por exemplo, após reacção de hidrólise enzimática.

Neste trabalho, utilizaram-se dois *kits* comerciais para a purificação em gel: QIAquick Gel Extraction Kit (QIAGEN) e NZYGelpure (NZYTech). Em ambos os *kits* o passo final de eluição de DNA foi realizado com 30 µl do tampão de eluição (em vez dos 50 µl propostos nos protocolos), de modo a obterem-se amostras mais concentradas.

Purificação Automatizada – Em Gel

Esta técnica é possível graças ao sistema E-Gel iBase Power System acoplado ao E-Gel Safe Imager Real-Time Transilluminator, ambos comercializados pela Invitrogen.

O primeiro sistema consiste num aparelho eléctrico onde se colocam cassetes de gel de agarose pré-preparadas (com corante SYBR Safe DNA gel da mesma empresa) com duas fileiras de poços: no topo encontram-se os poços de deposição das amostras de DNA a purificar e a meio da cassette, nova fileira de poços onde se recolhe o DNA, depois de separado no gel, puro. Com este método já não é necessário preparar-se o gel de agarose nem purificar DNA excisado do gel através de *kits* comerciais baseados em colunas de *spin*. Apesar de funcionar apenas com géis de agarose comerciais, este sistema torna-se mais económico e célere, ao evitar-se o passo de extracção com outro *kit*.

O segundo sistema foi desenhado para a visualização de géis corridos no sistema iBase, em tempo real e é acoplado ao primeiro sempre que se pretenda visualizar a posição do DNA no gel. Não produz luz ultravioleta, mas sim uma luz azul intensa, que permite uma visualização aceitável das bandas de DNA.

Depois de purificado o DNA desejado, este é quantificado em gel de agarose com um volume mínimo de DNA puro recolhido pelo sistema automatizado, e comparar-se com o padrão de bandas criado pelo marcador. Este sistema pode ser visualizado no Anexo III.

***Kit* de Clonagem**

Ao longo deste trabalho, optou-se por um passo de clonagem adicional prévio à ligação vector/fragmento para garantir que o fragmento que se pretendia integrar no vector de clonagem se encontrava totalmente digerido por ambas as enzimas de restrição. Para tal, utilizou-se um *kit* comercial, o NZYPCR Cloning Kit (NZYTech), que permite a clonagem directa de produtos puros de PCR e baseia-se em três passos experimentais, efectuados conforme protocolo facultado pela empresa:

- 1) **Reacção Phos:** preparação do fragmento para integrar no plasmídeo facultado com o *kit*. A enzima phos, uma fosfatase alcalina, retira o grupo fosfato da extremidade 3' impedindo que o DNA do fragmento recircularize;
- 2) **Reacção de ligação:** entre o fragmento de interesse previamente preparado conforme descrito no passo anterior e o vector de clonagem pNZY28a fornecido com o *kit* (Anexo I.5). As características principais deste vector são a existência de um codão de iniciação *lacZ*, um promotor *T7* e uma região codificante β -lactamase. O local de integração do fragmento no vector de clonagem é na MCR (*Multiple Cloning Region*) que se situa no gene *lacZ* (Anexo I.5);
- 3) **Transformação:** do produto da reacção de ligação em células competentes NZYStar facultadas com o *kit* (*endA1 hsdR17(r_k⁻, m_k⁺) supE44 thi⁻ 1 recA1 gyrA96 relA1 lac[F⁺proA⁺B⁺lacI^qZΔM15 :Tn10(Tc^R)*]). Estas células competentes são uma estirpe de *E.coli* indicada para protocolos de clonagem. Sendo *lacI^q* necessitam de IPTG (*IsoPropyl β-D-ThioGalactopyranoside*) para induzir a expressão a partir do promotor *lac*.

As células transformadas são plaqueadas em meio sólido LB contendo X-Gal (ou BCIG – *Bromo-Chloro-Indolul-Galactopyranoside*, VWR, International), o composto IPTG e o antibiótico correspondente à marca de selecção do vector, neste caso é a ampicilina (100 mg/ml).

A selecção das colónias correspondentes a clones positivos, foi feita com base na técnica de triagem azul/branco: o X-Gal é indicador da expressão da enzima β -galactosidase que é codificada pelo gene *lacZ* do operão *lac*. A enzima cliva o X-Gal, resultando nos produtos galactose e 5-bromo-4-cloro-3-hidroindole. Este último, é oxidado a 5,5'-dibromo-4,4'-dicloro-indigo, um produto insolúvel de cor azul. Assim, se o meio de crescimento sólido for suplementado com X-Gal e um indutor da enzima β -galactosidase (como o IPTG), as colónias que tiverem o gene *lacZ* são facilmente distinguidas das restantes, por apresentarem cor azul.

Ao transformarem-se as bactérias de *E.coli* (que não produzem a enzima β -galactosidase) com os produtos da reacção de ligação entre o vector pNZY28a e os fragmentos de interesse, as bactérias que integraram o vector recombinante deveriam ser

capazes de produzir a enzima β -galactosidase que, ao clivar o X-Gal presente no meio de crescimento sólido, originam colónias azuis. Contudo, a inserção do fragmento é feita na região do gene *lacZ* interrompendo a leitura do gene (inativação por inserção), impedindo-o de produzir uma enzima funcional. Assim, as colónias de bactérias que tenham integrado o vector sem o fragmento, ficam azuis; enquanto que as colónias de bactérias que tiverem assimilado o vector recombinante, ficam brancas.

Confirmação de Clonagem

A partir das colónias brancas seleccionadas, fez-se um PCR de colónia e preparou-se o crescimento das mesmas em meio líquido (2xYT suplementado com ampicilina (100 mg/ml), marca de resistência do plasmídeo pNZY28a). Os inóculos das colónias em meio líquido foram incubados durante a noite com agitação (150 rpm) a 37°C. A purificação dos fragmentos clonados a partir dos inóculos preparados foi feita com recurso ao *kit* NZYMiniprep da (NZYTech) e a análise da sua eficácia e o resultado das reacções de PCR foram analisados em gel de agarose.

As amostras que se entenderam positivas (de ora em diante denominados clones pNZY28a) foram então, sujeitas a uma reacção de hidrólise como abaixo se expõe.

Hidrólise Enzimática dos Fragmentos e Vectores

As hidrólises enzimáticas dos fragmentos, dos vectores e dos clones resultantes do *kit* de clonagem utilizado, foram feitas considerando-se os locais de restrição inseridos nos fragmentos pela amplificação com os respectivos *primers* e o local de múltipla clonagem presente nos quatro vectores estudados.

Utilizou-se o método de dupla hidrólise tanto na “abertura” dos vectores, como na preparação dos fragmentos a serem inseridos e nos clones pNZY28a. Testou-se a eficácia de uma dupla hidrólise em simultâneo e dupla hidrólise sequencial, tendo-se optado como método padrão pelo segundo. A mistura reaccional foi preparada inicialmente com adição de

apenas uma enzima (a enzima com maior actividade) e duas horas mais tarde, adicionava-se a segunda enzima à mistura, deixando que a reacção ocorresse durante a noite (período superior a 12h) a 37°C, temperatura óptima para a actividade das duas enzimas.

Enzimas de Restrição Utilizadas

Como já foi referido, a escolha das enzimas a utilizar contemplou os locais de restrição dos vectores de clonagem e expressão, e dos fragmentos amplificados; sendo imperativo que num par de clonagem vector/fragmento as extremidades criadas na clivagem representassem sequências nucleotídicas complementares. As enzimas utilizadas para a clivagem dos vectores, dos fragmentos e dos clones pNZY28a, são as apresentadas na Tabela 6.

Tabela 6 – Enzimas de restrição utilizadas de acordo com ensaios dos vários vectores.

Vector	Enzimas	Tampão de Hidrólise	Eficácia das Enzimas
pGEX-6P1	BamHI	NEBuffer EcoRI	100 %
	EcoRI		100 %
pET-14 e pET-15b	NdeI	NEBuffer 3	75 %
	BglII		100 %
pET-SUMO-28a	BamHI	NEBuffer 3	100 %
	XhoI		100 %

NEBuffer EcoRI – 50 mM NaCl, 100 mM Tris-HCl, 10 mM MgCl₂, 0,025 % Triton X-100, pH=7,5; NEBuffer 3 – 100 mM NaCl, 50 mM Tris-HCl, 10 mM MgCl₂, 1 mM DTT (*DiThioThreitol*), pH=7,9.

As enzimas e respectivos tampões são da marca New England BioLabs Inc.

Simulação das Reacções de Hidrólise

Fez-se um estudo prévio dos resultados previstos das reacções de hidrólise, de modo a confirmar se a utilização das enzimas escolhidas daria origem aos produtos clivados pretendidos. Este estudo foi possível, graças ao programa Serial Cloner.

Tabela 7 – Tamanho (pb) previsto dos fragmentos e vectores após clivagem com enzimas de restrição.

Vector/Fragmento	Enzimas	Pré-clivagem	Pós-clivagem
pGEX-6P1	BamHI	4 984	4 975
N001x145C	EcoRI	465	453
N55x145C		306	294
pET-14b	NdeI	4 671	4 515
pET-15b	BglII	5 708	5 544
N001x145C		465	455
N55x145C		306	296
pET-SUMO-28a	BamHI	5 369	ND
N001Sx145CS	XhoI	453	441
N55Sx145CS		294	282

N001x145C – indica o fragmento que foi amplificado com este par de *primers*. O mesmo se aplica aos restantes.

ND – Não Disponível.

Liofilização

O processo de liofilização consiste na concentração de uma amostra por desidratação. Neste processo, a amostra é congelada e a água nela contida é retirada por sublimação, ou seja, a água passa do estado sólido ao estado gasoso sem passar pela fase líquida. Pode resumir-se esta técnica em 3 etapas:

- 1) **Congelação:** a temperatura das amostras é reduzida abaixo dos 0°C;
- 2) **Secagem primária:** é eliminada ≈ 90 % da água. Ocorre a sublimação do gelo através do controlo da pressão de vácuo e o aumento gradual da temperatura. As amostras ficam com 15 % de humidade;
- 3) **Secagem secundária:** eliminados os restantes ≈ 10 % de água. A eliminação é feita por evaporação a vácuo com aumento da temperatura (20-60°C) As amostras desidratadas ficam apenas com 2 % de humidade.

Utilizou-se esta técnica para concentração de amostras de DNA após utilização dos métodos de purificação, quer se tratassem dos fragmentos de *yps* quer fossem os vectores de clonagem. Assim, desidrataram-se os 30 μ l de amostra recolhida no tampão de eluição, e ressuspendeu-se a amostra desidratada num volume mínimo de água, nunca superior a 10 μ l.

O liofilizador utilizado foi o Vacuum Concentrator Centrifuge da UniEquip, modelo UNIVAPO 100H com um aspirador de vácuo acoplado, Unijet II.

Reacção de Ligação Vector-Fragmento

As reacções de ligação entre os vectores de clonagem e expressão e os fragmentos foram realizadas com fragmento em excesso relativamente ao vector. Por norma, a mistura reaccional era constituída por: vector e fragmento (em diferentes proporções), T4 DNA Ligase (NZYTech), Tampão de Ligação 10x (NZYTech) e água, para um volume final de 20 μ l ou 50 μ l. Prepararam-se as misturas reaccionais com excesso de fragmento relativamente ao vector de modo a favorecer-se o máximo de formação de moléculas recombinantes. As amostras preparadas foram incubadas durante a noite a 20°C e, posteriormente, transformadas em células competentes de *E.coli*. Os transformantes foram plaqueados em meio de crescimento sólido e incubados durante a noite a 37°C. Observou-se crescimento de colónias brancas e com limites bem definidos.

Confirmação da Reacção de Ligação

Utilizaram-se dois métodos para averiguar o sucesso das reacções de ligação efectuadas: reacção de PCR a partir da colónia (mesmas condições de reacção que as referidas para a reacção de amplificação a partir de cDNA) seleccionada e hidrólise enzimática com as respectivas enzimas de restrição dos plasmídeos purificados a partir de colónias repicadas para meio de crescimento líquido (meio SOC suplementado com o antibiótico respectivo); ambas seguidas de electroforese em gel de agarose.

Esperar-se-ia, através do PCR de colónia o surgimento de bandas do tamanho dos fragmentos em estudo, considerando-se o clone seleccionado como positivo. A partir do gel de agarose com os produtos da reacção de hidrólise enzimática, um clone positivo teria que apresentar um padrão de duas bandas: uma de maior e outra de menor peso molecular correspondendo ao vector e fragmento, respectivamente.

Sequenciação de DNA

Antes de se transformarem os plasmídeos tidos como possíveis clones positivos em células *E.coli* BL21 para expressão proteica, sequenciaram-se as amostras com os *primers* T7 *forward* e T7 *reverse* de modo a confirmar-se se a sequência de DNA integrada no vector de clonagem era de facto a pretendida. A sequenciação foi feita pela empresa STAB VIDA.

Transformação

Os plasmídeos recombinantes, isto é, os clones que se concluíram positivos através dos métodos apresentados no tópico anterior, foram transformados em células competentes de *E.coli*: NZY5 α , NZYStar (descritas anteriormente) e BL21 (DE3), todas comercializadas pela empresa (NZYTech); conforme Protocolo facultado pela empresa.

As células competentes NZY5 α têm propriedades análogas às das células Dh5 α , desenvolvidas para transformações de elevada eficiência ($>10^8$ ufc – unidades formadoras de colónias). O marcador $\phi 80\Delta lacZ\Delta M15$ providencia uma α -complementação (activação) do gene que codifica para a β -galactosidase, e assim podem ser utilizadas para triagens azul/branco de colónias que cresceram em placas com meio sólido suplementado com X-Gal. Estas células, apresentam o seguinte genótipo: *fhuA2* Δ (*argF-lacZ*)U169 *phoA* *glnV44* $\Phi 80$ Δ (*lacZ*)M15 *gyrA96* *recA1* *relA1* *endA1* *thi-1* *hsdR17*.

Por fim, as células competentes BL21 (DE3) contêm uma cópia cromossomal do gene da T7 RNA polimerase. Este hospedeiro é um lisógeno do bacteriófago DE3, um derivado do bacteriófago lambda, que tem a região de imunidade do fago 21 e transporta um fragmento de DNA que contém o gene *lacI*, o promotor *lacUV5* (induzido por IPTG) e o gene da T7 RNA polimerase. Assim, a adição de IPTG ao meio de cultura induz a produção de T7 RNA polimerase, que por sua vez, transcreve o DNA em estudo integrado no plasmídeo. As transformações foram plaqueadas em meio sólido LB suplementado com ampicilina (100 mg/ml) e as placas inoculadas foram incubadas durante a noite a 37°C.

Testes de Expressão

A expressão de genes eucarióticos em bactérias (procariotas) apenas é possível através do cDNA desse gene, pois as bactérias não possuem os mecanismos necessários para o *splicing* dos intrões.

Para a expressão da proteína yps, utilizaram-se duas abordagens distintas: indução com IPTG e auto-indução, ambas com base nas propriedades dos vectores pET e das respectivas células hospedeiras, neste caso, as células competentes BL21 (DE3).

Indução com IPTG

A indução com IPTG é útil quando se utilizam estirpes hospedeiras com o gene que codifica para a enzima T7 RNA polimerase e vectores que contenham o respectivo promotor, T7. Trata-se do caso deste estudo, onde se utilizaram células BL21 (DE3) para transformação do vector recombinante pET-15b. Tal como referido anteriormente, a adição de IPTG ao meio onde ocorre o crescimento das células transformadas vai induzir a activação do promotor contido no vector pET, resultando na tradução do fragmento de DNA integrado neste plasmídeo.

A indução da expressão proteica foi feita a partir de colónia seleccionada do meio sólido que havia sido plaqueado com as células da reacção de transformação. Preparou-se um pré-inóculo a partir da colónia em meio LB suplementado com ampicilina (100 mg/ml) que foi incubado a 37°C durante a noite com agitação. Posteriormente, o pré-inóculo foi adicionado a novo volume (1L) de meio LB com ampicilina (100 mg/ml), e novamente incubado com agitação a 37°C. Mediu-se a densidade óptica (DO) de alíquotas que foram sendo retiradas deste volume de crescimento, até ser atingida uma $DO \approx 0,5-0,6$ (corresponde à fase *mid-log* do crescimento bacteriano, na qual as células se encontram muito activas). Chegado a este valor de DO, dividiu-se o volume de crescimento em 5 (x2) volumes iguais aos quais se adicionaram diferentes concentrações de IPTG no intervalo [0,0;1,0] mM; 5 inóculos foram incubados a 18°C e os restantes 5 a 30°C, todos com agitação. Novamente,

foram retiradas alíquotas de todos os inóculos a diferentes tempos de incubação para medição da DO, até a um tempo final de 16h.

As alíquotas retiradas ao longo deste estudo foram armazenadas a 4°C para posterior análise.

Auto-indução

O protocolo de auto-indução utilizado baseia-se na produção de proteínas recombinantes em bactérias que utilizem o sistema de expressão pET, como as células BL21 (DE3). O método de auto-indução foi originalmente desenvolvido por Studier (Studier, 2005) e a expressão da proteína em estudo ocorre sem que seja necessária a adição de indutores aos meios de crescimento durante a fase de crescimento *mid-log*. Consiste na utilização de um meio de crescimento tamponizado que contém uma mistura de fontes de carbono, incluindo lactose. Utilizou-se um meio de crescimento composto por: 0,1 % de MgSO₄ (1M), 2 % Suplemento 50x5052 (Tabela 1), 5 % do Suplemento 20xNPS (Tabela 1) no meio LB com ampicilina (100 mg/ml).

A auto-indução depende dos mecanismos de que a bactéria dispõe para regular a utilização de carbono e outras fontes de energia presentes no meio. No início do crescimento bacteriano, as células utilizam a glucose como fonte de carbono. Quando esta se esgota no meio, a lactose pode entrar nas células e induzir a expressão da enzima T7 polimerase a partir do lisógeno DE3. A presença de 0,05 % de glucose no meio de crescimento de auto-indução inibe a utilização da lactose no início do crescimento, impedindo assim a indução da T7 polimerase, e é de tal modo eficaz que até mesmo estirpes capazes de expressar proteínas alvo altamente tóxicas para as células hospedeiras, podem crescer e manter plasmídeos funcionais.

Apresenta-se como um método simples no qual ocorre o crescimento bacteriano até à saturação sem a necessidade de se controlar a taxa de crescimento celular.

Preparação do Lisado Bacteriano

Os crescimentos bacterianos foram centrifugados a 6 000 rpm durante 15 minutos, a 4°C em tubos de 500 ml da Beckman. O sobrenadante foi descartado e os pellets ressuspensos cuidadosamente, em 25 ml de 1 x PBS, em gelo. A lise celular foi feita por sonicação com uma sonda UP 200S (Hielscher) a uma amplitude de 70 %: 10 ciclos de 30 segundos com intervalos de 30 segundos, para evitar o sobre-aquecimento das amostras. O lisado foi centrifugado a 20 000 rpm durante 1h a 4°C em tubos de 50 ml da Nalgene. Separou-se o sobrenadante do pellet, retiraram-se as respectivas alíquotas e foram mantidas em gelo.

Análise dos Resultados dos Testes de Expressão

A análise dos resultados de ambos os testes de expressão foi feita através da técnica de electroforese em gel de poliacrilamida desnaturante. Como já foi referido anteriormente, a electroforese baseia-se na separação de macromoléculas por aplicação de uma diferença de potencial gerada por um campo eléctrico. A aplicação deste campo eléctrico a uma mistura de proteínas resulta num padrão de migração distinto de proteína para proteína, dependente da sua dimensão (peso molecular) e carga global: as proteínas de maior massa migram menos que as proteínas de massa menor. A forma das moléculas também influencia a velocidade de migração: moléculas grandes e assimétricas migram mais lentamente que moléculas esféricas com a mesma massa.

A separação de proteínas e a consequente análise do seu padrão de migração é possível através da técnica de electroforese em gel de poliacrilamida e vulgarmente conhecida como, PAGE (*PolyAcrylamide Gel Electrophoresis*). Às amostras é adicionado o detergente SDS (*Sodium Dodecyl Sulfate*) que tem como função desnaturar proteínas e a neutralizar as moléculas com carga negativa, mantendo constante a razão carga/massa. Este detergente liga-se às cadeias laterais hidrofóbicas, destabilizando as interacções hidrofóbicas no centro da proteína que contribuem para uma conformação estável. O tratamento com SDS é combinado com o aquecimento na presença de agentes redutores (Tampão de Amostra de Proteína, Tabela 1), quebrando-se as ligações dissulfureto. Deste modo, a separação é feita tendo

apenas por base o tamanho das proteínas tal como acontece com moléculas de DNA na electroforese em gel de agarose.

A preparação dos géis de poliacrilamida consiste em dois passos: preparação do gel de resolução (12 %) sobre o qual, depois de polimerizado, se preparava o gel de concentração (*stacking*). Os géis foram preparados conforme consta na Tabela 8.

Tabela 8 – Composição (em μl) dos géis de acrilamida de resolução e de *stacking*.

Reagentes	Gel de Resolução	Gel de <i>Stacking</i>
Acrilamida	2 000	325
Tris 8,8	1 250	
H₂O_{mq}	1584	1 525
APS	17	12,5
TEMED	3,5	2,5
Tris 6,8		625

APS – Ammonium PerSulfate; TEMED – *N,N,N',N'*-TEtraMethylEthyleneDiamine

A 10 μl das alíquotas retiradas nos dois estudos de expressão, adicionaram-se 10 ou 20 μl de tampão de amostra de proteína, consoante se tratassem de amostras em solução ou amostras de *pellet*, respectivamente. As amostras foram fervidas a 100°C durante 5 minutos. As amostras fervidas foram carregadas nos vários poços, tendo como referência de padrão de migração o marcador de pesos moleculares NZYColour Protein Marker (NZYTech), podendo ser consultado no Anexo II.2. A electroforese foi feita em tampão SDS 0.1% nas condições: 150 V, 50 mM durante 60 minutos; num sistema vertical Mini-Protean Tetra Cell da Bio-Rad.

Os géis foram colocados na Solução Corante (Tabela 1) e posteriormente diferenciados com a Solução Descorante (Tabela 1) para visualização das bandas correspondentes às proteínas expressas.

Purificação da Proteína – Cromatografia por Afinidade

Apenas a proteína expressa através dos testes de auto-indução, foi sujeita ao passo de purificação. Esta proteína recombinante apresenta na sua extremidade N-terminal uma cauda

de histidinas, facultada pelo vector pET onde o respectivo DNA se encontrava inserido. A técnica utilizada para este efeito foi a técnica de cromatografia de afinidade com metal imobilizado ou IMAC (*Immobilized Metal Affinity Chromatography*) que se baseia na afinidade das proteínas de fusão com caudas de histidinas, pelos metais de transição, tais como: Cu^{2+} , Ni^{2+} , Zn^{2+} ou Co^{2+} .

O metal seleccionado para o efeito, encontra-se imobilizado numa resina empacotada numa coluna. Ao introduzir-se na coluna a solução de proteínas, estas interagem com os iões metálicos imobilizados no suporte sólido. A separação das várias proteínas presentes na solução e consequente purificação da proteína em estudo, é possível atendendo ao conteúdo relativo de resíduos de histidina, cisteína e cadeias laterais aromáticas e de grupos N-terminal acessíveis, que constituem a proteína. Como a proteína expressa contém uma sequência de 6 resíduos de histidina expostos no N-terminal, estes podem interagir com o metal imobilizado na coluna, ficando assim a proteína retida. As restantes proteínas do extracto bruto interagem de forma mais fraca com o metal, pelo que atravessam a coluna. Conseguem-se deste modo uma purificação de elevado rendimento, num único passo de purificação.

Neste processo, utilizou-se a resina Ni-NTA Agarose da empresa Qiagen. A resina foi lavada com 3-5 volumes de coluna com $\text{H}_2\text{O}_{\text{mq}}$ e 5-10 volumes de coluna com 1 x PBS. Seguidamente, foi incubada com o lisado bacteriano (sobrenadante do passo de preparação) a $\approx 4^\circ\text{C}$ durante 1h, com agitação constante (120 rpm). Terminado o tempo de incubação, colocou-se esta suspensão numa coluna de vidro ou polipropileno (PD-10 da GE Healthcare, filtro com poro de 20-85 μm) e recolheu-se o extracto bruto. De seguida, a purificação foi efectuada com um gradiente de imidazol em tampão 300 mM NaCl. O imidazol competiu com a cauda de poli-histidinas pela ligação ao níquel imobilizado na coluna. O gradiente utilizado foi bastante amplo, uma vez que se desconhecía a que concentração deste composto a proteína yps seria eluída, utilizou-se uma gama de concentrações desde os 10 aos 500 mM (diluições preparadas a partir da solução de lavagem concentrada de 500 mM). A purificação foi feita a 4°C .

Recolheram-se as várias fracções e retiraram-se alíquotas para posterior análise em gel SDS-PAGE e análise da reacção com o reagente de Bradford. A cada fracção, adicionou-se 1

ml de DTT 1 mM (forte agente redutor) e 1 ml de EDTA 1 mM (agente quelante), e foram todas armazenadas a 4°C.

Método de Bradford

Trata-se de um método colorimétrico para determinação da concentração de proteínas totais numa solução. Baseia-se na interação entre o corante presente no reagente de Bradford (*Comassie Brilliant Blue*, BG-250) e as cadeias laterais básicas e aromáticas dos resíduos de aminoácidos das proteínas presentes na solução. A ligação covalente do reagente de Bradford às proteínas é acompanhada de alteração da cor do reagente (castanho), em meio ácido a solução fica azul. A concentração da proteína total na amostra é estimada através da leitura espectrofotométrica deste produto corado.

Contudo, neste trabalho não se utilizou este método para quantificar a proteína presente nas várias amostras. Utilizou-se este método somente como forma de comparar a quantidade de proteína (yps ou proteínas de *E. coli*) eluída em cada fracção, permitindo acompanhar visualmente este processo.

O reagente de Bradford utilizado foi o Quick Start Bradford Dye Reagent 1x, comercializado pela Bio-Rad Laboratories, Inc.

Resultados

Nota: a maioria das imagens apresentadas são digitalizações de impressões feitas a partir da imagem recolhida com a máquina fotográfica, pelo que em algumas não é nítido o sinal presente no gel, visível de forma clara aquando da recolha da imagem no transiluminador.

A apresentação dos resultados contemplará os ensaios efectuados com os quatro vectores seleccionados para este estudo: pGEX-6P1, pET-14b, pET-15b e pET-SUMO-28a. Para maior clareza da exposição dos resultados, apresenta-se abaixo um esquema com os passos experimentais realizados para os ensaios com os vectores referidos.

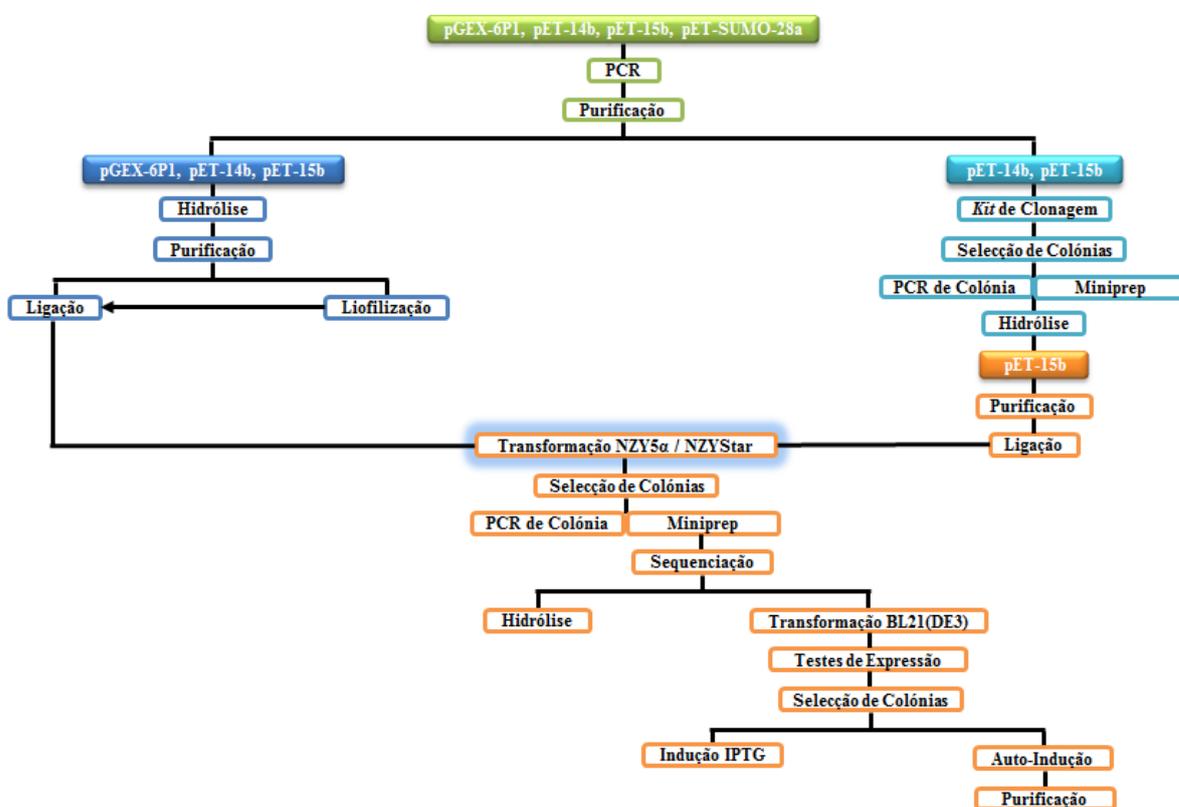


Figura 20 – Esquema do Procedimento Experimental adoptado para os ensaios com os vários vectores. As caixas a cheio indicam os fragmentos preparados para clonar nos plasmídeos indicados; as caixas a vazio indicam as técnicas experimentais utilizadas.

Hidrólise e Purificação dos Vectores

Após hidrólise dos plasmídeos com as respectivas enzimas, estes foram purificados e a análise da integridade dos mesmos foi feita em gel de agarose, onde se colocaram as amostras digeridas e puras, e o gel foi visualizado no transiluminador. Em todas as purificações doravante mencionadas, do volume final de eluição, foram retirados 5 μ l aos quais se adicionaram 3 μ l de tampão de amostra, e carregou-se este volume em gel de agarose preparado conforme descrito anteriormente. As condições de corrida utilizadas e a visualização do gel foram as já referidas.

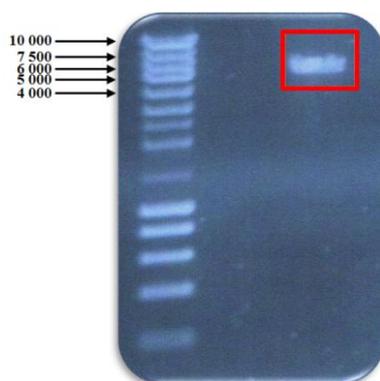


Figura 21 – Visualização do vector pGEX-6P1 digerido e puro. A purificação foi feita com o *kit* NZYGelpure (NZYTech), com um rendimento médio de \approx 60 %.

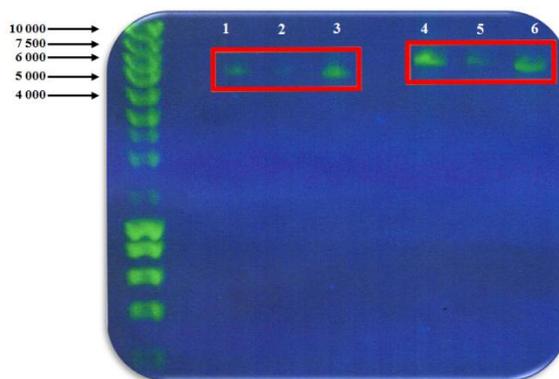


Figura 22 – Análise dos vectores pET-14b (poços 1-3) e pET-15b (poços 4-6) digeridos e puros. A purificação foi feita com o *kit* QIAquick Gel Extraction Kit (QIAGEN), com um rendimento médio de \approx 50 %.

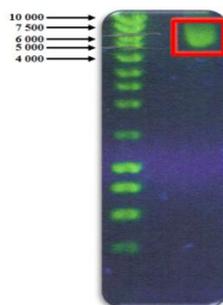


Figura 23 – Análise do vector pET-SUMO-28a digerido e puro. A purificação foi feita com o sistema E-Gel iBase™ Power System (Invitrogen), com um rendimento médio de $\approx 40\%$.

Ensaio com os Vectores pGEX-6P1, pET-14b, pET-15b e pET-SUMO-28a

Reacções de PCR e Purificação dos Fragmentos Amplificados

A partir do cDNA correspondente ao gene *yps*, amplificaram-se os dois fragmentos desejados para os ensaios com estes vectores. A amplificação de cada fragmento foi feita em quadruplicado (A,B,C,D) utilizando-se em índice o número 1 com correspondência ao fragmento de maior peso molecular, e o número 2 para o fragmento menor. Salienta-se ainda, que os resultados aqui apresentados foram sujeitos a diversos ensaios para otimizar tanto a composição da mistura reaccional, como o programa de amplificação a utilizar.

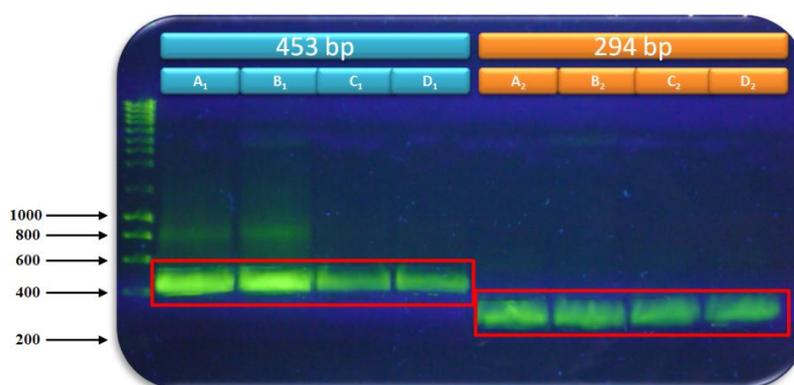


Figura 24 – Fragmentos de amplificação do gene *yps* para ensaios com pET-SUMO-28a.

Após a extracção das bandas a partir do gel de agarose onde se carregaram os produtos de PCR relativos aos ensaios com pGEX-6P1, pET-14b e pET-15b, os fragmentos foram purificados, e o resultado da purificação analisado em gel de agarose.



Figura 25 – Análise dos fragmentos puros de *yps* para os ensaios com os vectores pGEX-6P1, pET-14b e pET-15b. A purificação foi feita com o *kit* QIAquick Gel Extraction Kit (QIAGEN), com um rendimento médio de \approx 90 %.

Ensaio com os Vectores pGEX-6P1, pET-14b e pET-15b

Hidrólises Enzimáticas e Purificação dos Produtos Hidrolisados

As enzimas de restrição utilizadas estiveram de acordo com os locais de corte dos vectores e dos fragmentos: nos ensaios com pGEX-6P1 os fragmentos de *yps* foram clivados com as enzimas EcoRI e BamHI; e nos ensaios com os vectores pET-14b, pET-15b os respectivos fragmentos de *yps*, foram clivados com as enzimas NdeI e BglII. As impurezas existentes nas misturas reaccionais das hidrólises enzimáticas, foram eliminadas por purificação através de uma das técnicas referidas.



Figura 26 – Análise dos fragmentos digeridos e puros para os ensaios com pGEX-6P1. A purificação foi feita com *kit* NZYGelpure (NZYTech) com um rendimento médio de $\approx 33\%$. Amostras A₁, B₁ e A₂, B₂ são as que apresentam maior concentração.

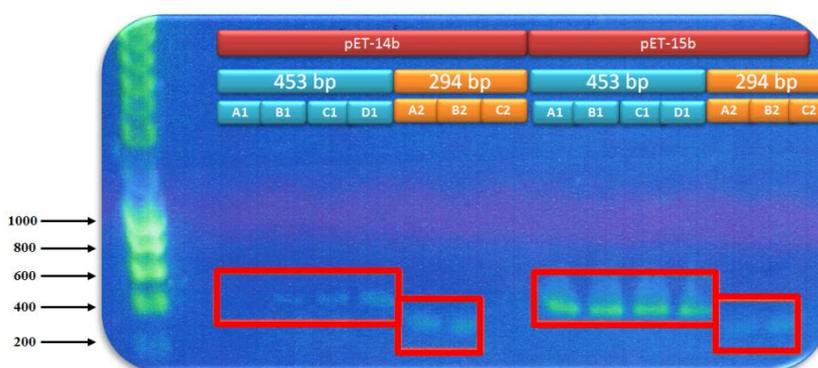


Figura 27 – Análise dos fragmentos digeridos e puros para os ensaios com pET-14b e pET-15b. As purificações foram feitas com *kit* NZYGelpure (NZYTech) para pET-14b e o *kit* QIAquick Gel Extraction Kit (QIAGEN) para pET-15b, com rendimentos médios de $\approx 33\%$ e de $\approx 50\%$. Amostras D₁, A₂ e B₂ referentes ao pET-14b e as amostras A₁, B₁ e A₂ relativas ao pET-15b, são as que apresentam maior concentração.

Reacção de Ligação

Depois de digeridos e purificados, os vectores e os respectivos fragmentos, preparou-se a reacção de ligação para um volume final de 50 μ l, com proporção molar de vector/fragmento de 1:3. Apenas foram utilizados os fragmentos de *yps* que apresentavam em gel uma concentração mais elevada. Depois de incubadas as misturas reaccionais, o resultado das reacções de ligação foi analisado em gel de agarose, para averiguar se ocorreria de facto ligação entre o vector e o fragmento a inserir, comparando-se para tal as amostras das misturas reaccionais com os respectivos vectores de clonagem, não hidrolisados.



Figura 28 – Análise da eficácia das reacções de ligação nos ensaios com pGEX-6P1, pET-14b e pET-15b.

Na tentativa de aumentar a concentração das amostras dos fragmentos em estudo para se conseguir um maior concentração de fragmento relativamente ao vector na reacção de ligação, fizeram-se novos ensaios nos quais as amostras de fragmento hidrolisadas e puras foram liofilizadas até desidratação completa. Por fim, as amostras foram ressuspensas num volume mínimo (10 µl) de água estéril e através do gel de agarose, analisou-se se ocorrera ou não um aumento da concentração do DNA dos fragmentos de *yps*.



Figura 29 – Análise da eficácia do processo de liofilização nos ensaios com pGEX-6P1, pET-14b e pET-15b.

Os rectângulos vermelhos indicam os fragmentos que eram visíveis.

Da liofilização resultaram amostras menos concentradas que as amostras de partida, o oposto do desejado. Contudo, fizeram-se novas reacções de ligação a partir das amostras liofilizadas que apresentaram sinal no gel de agarose, utilizando-se para tal as condições mencionadas anteriormente, mas a proporção vector/fragmento diminuiu de 1:3 para 1:2, respectivamente. O resultado das reacções de ligação, foi carregado em novo gel de agarose, juntamente com alíquotas dos três vectores puros e não digeridos.

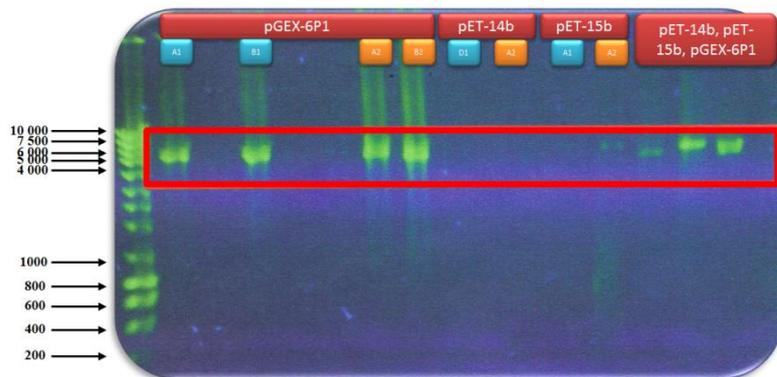


Figura 30 – Análise da eficácia das reacções de ligação após liofilização, nos ensaios com pGEX-6P1, pET-14b e pET-15b.

O único plasmídeo recombinante que apresentou sinal no gel de agarose, foi o pGEX-6P1 ligado aos fragmentos A₁, B₁, A₂ e B₂. As bandas destas amostras aparentavam o tamanho esperado, pelo que se prepararam as respectivas transformações em *E.coli*.

Transformação dos Clones em Células NZY5 α

A transformação das reacções de ligação destas células competentes foi feita de acordo com o protocolo descrito nos Materiais e Métodos. As placas de Petri inoculadas com as células transformadas, não apresentaram crescimento.

Ensaio com os Vectores pET-14b e pET-15b

Kit de Clonagem

Na tentativa de se perceber se os maus resultados anteriores se deviam a problemas com o fragmento a integrar, com o vector, ou até mesmo com outro constituinte da mistura reaccional de ligação, testou-se a utilidade do *kit* de clonagem NZYPCR Cloning Kit (NZYTech) na exclusão da hipótese de eventuais problemas com os fragmentos. Os produtos de PCR puros utilizados, foram os que constam na Figura 25: amostras B_{1,2} e D_{1,2} preparadas para clonar em pET-14b, e as amostras A_{1,2} e C_{1,2}, preparadas para clonagem em pET-15b. Estas amostras foram tratadas com a enzima *phos*, ligadas ao vector pNZY28a, transformadas

nas células competentes de *E.coli* (NZYStar ou NZY5 α) e por fim, os transformantes plaqueados em meio sólido LB com ampicilina (100 mg/ml) e X-Gal (8 placas, uma para cada amostra). Depois de incubadas, as placas de Petri inoculadas com os transformantes apresentavam uma maioria de colónias azuis (referentes a clones negativos) e algumas colónias brancas individualizadas e de margens bem definidas, que representariam possíveis clones positivos. De cada placa, seleccionaram-se duas colónias brancas e a designação de cada colónia foi atribuída da seguinte forma: às duas colónias seleccionadas da placa onde se inocularam os transformantes da reacção de ligação do fragmento A₁ com o vector de clonagem, denominaram-se de colónias A_{1.1} e A_{1.2}. Procedendo-se de modo análogo com as restantes colónias. As placas foram armazenadas a 4°C.

A confirmação de que as colónias seleccionadas corresponderiam de facto a clones positivos foi feita através de reacção de PCR e, preparação de inóculos em meio líquido (posteriormente purificados), a partir de cada colónia.

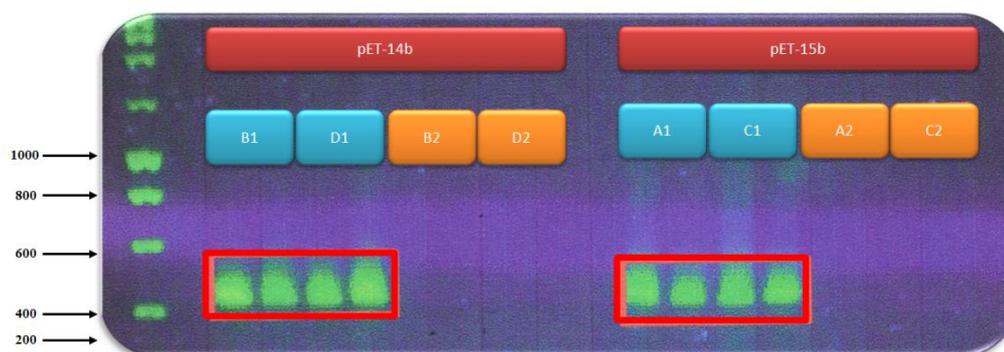


Figura 31 – Confirmação da eficácia da clonagem por PCR de colónia.

Das amostras puras obtidas do crescimento das colónias em meio líquido, confirmou-se o resultado obtido com PCR de colónia (não foi tirada fotografia do respectivo gel de agarose).

Procedeu-se à hidrólise enzimática de alguns dos clones tidos como positivos (B_{1.2}, B_{2.2} e A_{1.1}, A_{2.1}), com as enzimas respeitantes aos vectores pET-14b e pET-15b: NdeI e BglII. Caso representassem de facto clones positivos, esperar-se-ia o aparecimento de duas bandas: uma referente ao vector pNZY28a e outra referente ao fragmento de *yps* que foi integrado.

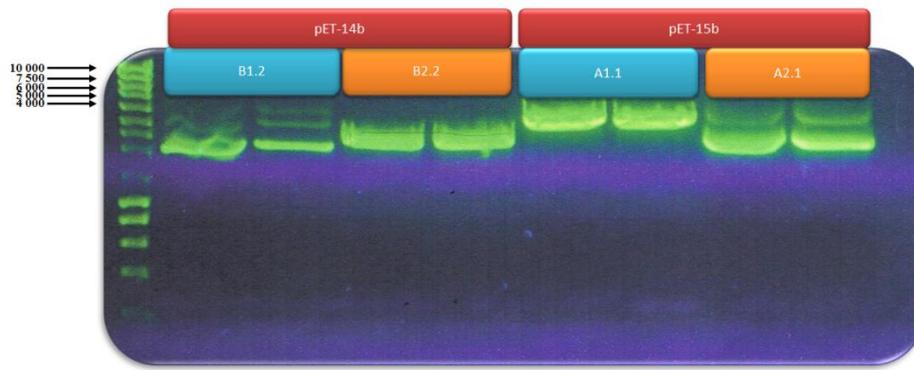


Figura 32 – Análise do resultado da hidrólise enzimática dos clones B_{1,2}, B_{2,2} e A_{1,1}, A_{2,1}.

Perante os resultados obtidos, repetiu-se a reacção de hidrólise enzimática, desta vez com a adição em simultâneo das duas enzimas. Obteve-se o mesmo resultado, tomando-se como falsos positivos os clones seleccionados. Assim, seleccionaram-se novas colónias a partir das mesmas placas, utilizando-se a mesma nomenclatura para denominar os clones. Fizeram-se novos PCRs a partir das colónias seleccionadas.

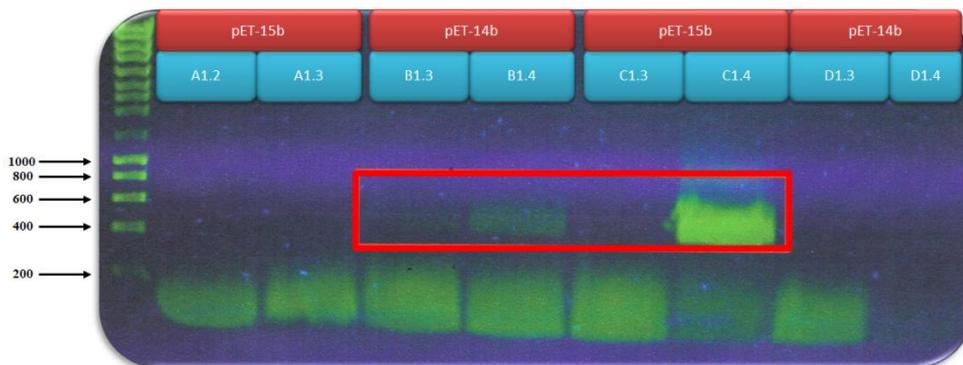


Figura 33 – Confirmação da eficácia da clonagem por PCR de colónia, fragmentos de 465 pb.

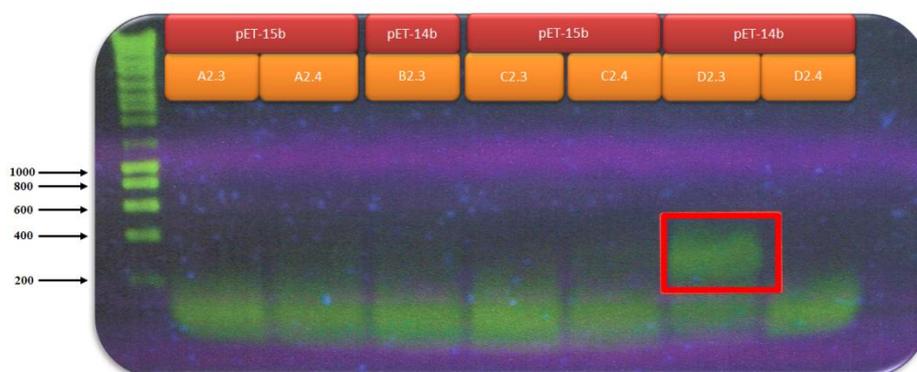


Figura 34 – Confirmação da eficácia da clonagem por PCR de colónia, fragmentos de 306 pb.

Apenas os clones B_{1.3}, B_{1.4}, C_{1.4} e D_{2.3}, apresentaram sinal no tamanho esperado, confirmando-se a presença do fragmento nas células seleccionadas. Para confirmar que de facto os fragmentos estavam integrados no vector utilizado no *kit*, prepararam-se inóculos destes clones em meio líquido.

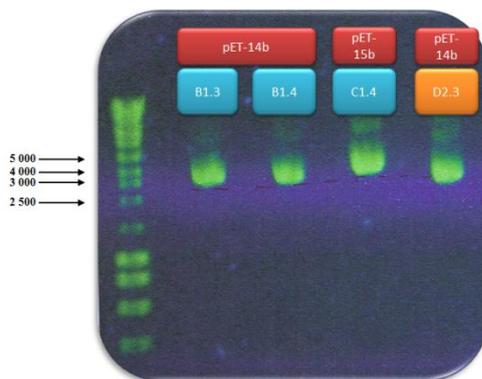


Figura 35 – Confirmação da eficácia da clonagem por crescimento em meio líquido, plasmídeos com os fragmentos de 465 e 306 pb, inseridos.

Pelos resultados anteriores esperava-se que estes clones representassem verdadeiros positivos: nos 4 clones visualizam-se bandas no tamanho aproximado do vector pNZY28a verificando-se que a banda correspondente à amostra D_{2.3} (fragmento de menor peso molecular) migrou ligeiramente mais que as restantes. A confirmação destes clones, foi feita tal como anteriormente, por hidrólise enzimática.

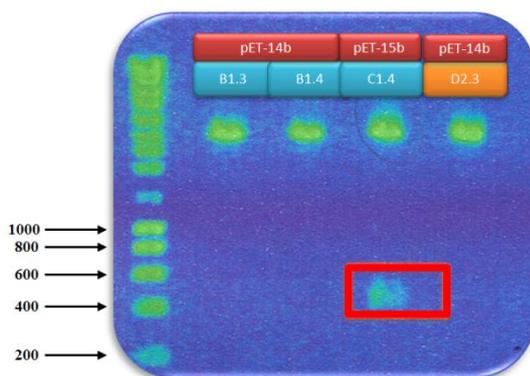


Figura 36 – Análise dos produtos da hidrólise enzimática, fragmentos de 465 e 306 pb.

Ensaio com o Vector pET-15b

Dos 4 clones tidos como positivos, apenas o clone C_{1.4} (preparado para integrar em pET-15b) apresentou um padrão de duas bandas nos tamanhos correspondentes ao vector utilizado e ao fragmento que nele havia sido integrado. Considerou-se o clone C_{1.4} como verdadeiro positivo.

Purificação do Fragmento Extraído do Clone C_{1.4}

O restante volume da reacção de hidrólise foi purificado através do sistema E-Gel iBase Power System (Invitrogen). O fragmento extraído do clone, de ora em diante será designado por fragmento C_{1.4}.

Reacção de Ligação

Com o fragmento C_{1.4} e o vector pET-15b puros, prepararam-se duas reacções de ligação, A e B.

Tabela 9 – Volumes utilizados na preparação das reacções de ligação A e B

Reacção	pET-15b	C _{1.4}	H ₂ O _{mq}	T4 Ligase	10xTampão	Volume final
A	1	9	7	1	2	20
B	3	14	0	1	2	

Sequenciação

Antes de se transformarem as duas reacções de ligação em células de *E.coli* BL21 (DE3), enviaram-se amostras dos vectores recombinantes puros para sequenciação com os *primers* T7 *forward/reverse* para amplificação completa do vector pET. Depois de recebidos os resultados, alinharam-se as sequências destes com a sequência de DNA do fragmento clonado, o fragmento de 465 pb. Para tal, utilizaram-se os *softwares* BioEdit Sequence Alignment Editor e Serial Cloner. Do resultado da sequenciação verifica-se que foram

sequenciados 1 100 nucleótidos do vector recombinante, confirmando-se que o fragmento do gene *yps* se encontra clonado na região de 111-559 nucleótidos, de acordo com o alinhamento das sequências.

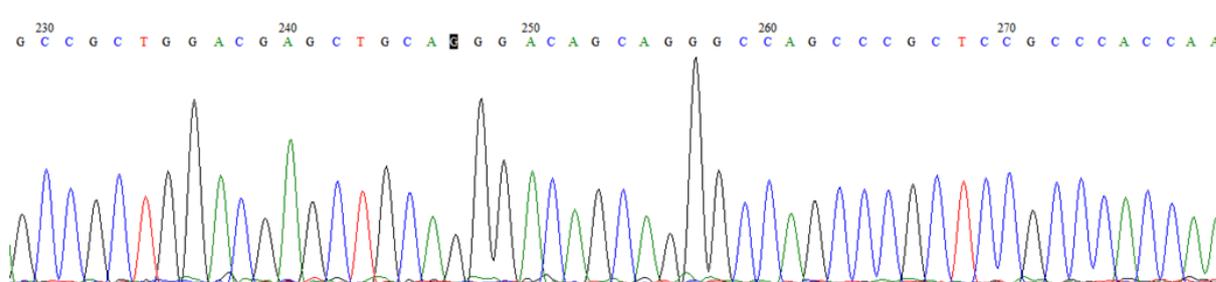


Figura 37 – Parte do gráfico referente à sequenciação do plasmídeo recombinante da reacção de ligação A, 229-279 nucleótidos. O código de cores dos picos corresponde aos respectivos nucleótidos, identificados acima.

Pelo alinhamento entre a sequência de DNA do fragmento de 465 pb e os nucleótidos sequenciados, não se verificou a inserção de qualquer mutação.

Reacção de Ligação A		Reacção de Ligação B	
9	-----ccatatggct 18	9	-----ccatatggc 17
4741	AGCAG CATCATCATCATCA TACAGCAGCGGCCCTGGTGCCGCCGGCAGCCATATGGCT 4800	61	CAGCAGC CATCATCATCATCA TACAGCAGCGGCCCTGGTGCCGCCGGCAGCCATATGGC 120
19	gatgccgaggagAGTAAGCCACTGGCCGCCGAACAGCAGCAGGCCGACAGCAGCAGCCGGAG 78	18	tgatgccgaggagAGTAAGCCACTGGCCGCCGAACAGCAGCAGGCCGACAGCAGCAGCCGGA 77
4801	GATGCCGCCGAGAGTAAGCCACTGGCCGCCGAACAGCAGCAGGCCGACAGCAGCAGCCGGAG 4860	121	TGATGCCGCCGAGAGTAAGCCACTGGCCGCCGAACAGCAGCAGGCCGACAGCAGCAGCCGGA 180
79	CAGCAGCAGAATCCGCCGAATCCGAGGAGCAGGATCACGAGCAGGAGCCGCTGGACGAG 138	78	GCAGCAGCAGAATCCGCCGAATCCGAGGAGCAGGATCACGAGCAGGAGCCGCTGGACGA 137
4861	CAGCAGCAGAATCCGCCGAATCCGAGGAGCAGGATCACGAGCAGGAGCCGCTGGACGAG 4920	181	GCAGCAGCAGAATCCGCCGAATCCGAGGAGCAGGATCACGAGCAGGAGCCGCTGGACGA 240
139	CTGCAGGGACAGCAGGGCCAGCCCCTCCGCCACCAGGAAGTTCATCGCCACC AAGTTC 198	138	GCTGCAGGGACAGCAGGGCCAGCCCCTCCGCCACC AAGGAAGTTCATCGCCACC AAGTTC 197
4921	CTGCAGGGACAGCAGGGCCAGCCCCTCCGCCACCAGGAAGTTCATCGCCACC AAGTTC 4980	241	GCTGCAGGGACAGCAGGGCCAGCCCCTCCGCCACC AAGGAAGTTCATCGCCACC AAGTTC 300
199	ACCGGCACCCTCAAGTGGTTCAACGTGAAGAGCGGCTACGGCTTCATCAACCGCAACGAC 258	198	CACCGGCACCCTCAAGTGGTTCAACGTGAAGAGCGGCTACGGCTTCATCAACCGCAACGA 257
4981	ACCGGCACCCTCAAGTGGTTCAACGTGAAGAGCGGCTACGGCTTCATCAACCGCAACGAC 5040	301	CACCGGCACCCTCAAGTGGTTCAACGTGAAGAGCGGCTACGGCTTCATCAACCGCAACGA 360
259	ACCAGAGAGGATGCTTTTGTGCCACAGAGCGCCATTGCGCGGAACAACCCAAAAAGGC 318	258	CACCAGAGAGGATGCTTTTGTGCCACAGAGCGCCATTGCGCGGAACAACCCAAAAAGGC 317
5041	ACCAGAGAGGATGCTTTTGTGCCACAGAGCGCCATTGCGCGGAACAACCCAAAAAGGC 5100	361	CACCAGAGAGGATGCTTTTGTGCCACAGAGCGCCATTGCGCGGAACAACCCAAAAAGGC 420
319	GTCCGC TC GGTGGGC GACGGTGAGGTCGTTGAGTTCGACGTGGTCATTGGTGAGAAGGGC 378	318	GGTCCGCTCGGTGGGC GACGGTGAGGTCGTTGAGTTCGACGTGGTCATTGGTGAGAAGGG 377
5101	GTCCGC TC GGTGGGC GACGGTGAGGTCGTTGAGTTCGACGTGGTCATTGGTGAGAAGGGC 5160	421	GGTCCGCTCGGTGGGC GACGGTGAGGTCGTTGAGTTCGACGTGGTCATTGGTGAGAAGGG 480
379	AACGAGGCGGCCAACGTGACCGGTCCTCCGGTGAGCCGGTGGGGCAGtcagtttgc 438	378	CAACGAGGCGGCCAACGTGACCGGTCCTCCGGTGAGCCGGTGGGGCAGtcagtttgc 437
5161	AACGAGGCGGCCAACGTGACCGGTCCTCCGGTGAGCCGGTGGGGCAGtcagtttgc 5220	481	CAACGAGGCGGCCAACGTGACCGGTCCTCCGGTGAGCCGGTGGGGCAGtcagtttgc 540
439	gctgacaagttaaatgac----- 455	438	agcggacaagttaaatgac----- 455
5221	GC GGACAAGTAAAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTG 5280	541	AGCGGACAAGTAAAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTG 600

Figura 38 – Alinhamento entre a sequência do fragmento de 465 pb de *yps* (a rosa) e os produtos das reacções de ligação A (a verde) e B (a azul). Os rectângulos vermelhos assinalam os resíduos de histidina em cada um dos clones.

Transformação do Clone em NZYStar (NZYTech)

A transformação foi feita de acordo com o protocolo descrito em Materiais e Métodos, mas utilizaram-se as células competente NZYStar (NZYTech) em alternativa às células NZY5 α , da mesma empresa. Os transformantes de cada reacção de ligação (C_{1.4}A e C_{1.4}B) foram inoculados em duas placas de Petri com meio sólido LB suplementado com ampicilina (100 mg/ml). Depois de incubadas durante a noite, a 37°C, ambas as placas apresentaram crescimento, tendo-se seleccionado de cada placa 5 colónias para estudo. A confirmação do sucesso das clonagens pET-15b x C_{1.4}A e pET-15b x C_{1.4}B, foi feita como referido anteriormente: PCR de colónia e crescimento em meio líquido.



Figura 39 – Confirmação da eficácia do passo de clonagem por PCR de colónia.



Figura 40 – Confirmação da eficácia do passo de clonagem por crescimento dos clones em meio líquido.

Todos os clones apresentaram-se como potenciais positivos. Assim, hidrolizaram-se as amostras com as enzimas NdeI e HindIII, para verificar quais representavam efectivamente positivos verdadeiros.



Figura 41 – Análise dos produtos da hidrólise enzimática, clones $C_{1.4}A$ (1-5) e $C_{1.4}B$ (1-5).

Os clones $C_{1.4}A$ e $C_{1.4}B$ apresentaram um padrão de bandas nos tamanhos esperados, considerando-se assim como clones positivos verdadeiros.

Transformação do Clone em BL21 (DE3) (NZYTech)

Os clones que se revelaram positivos foram transformados em células competentes BL21 (DE3), optimizadas para expressão de proteína em *E.coli*. O protocolo utilizado para a transformação dos clones foi análogo ao já referido. Cada clone transformado foi inoculado numa placa de Petri com meio sólido e antibiótico. Depois de incubadas, durante a noite e a $37^{\circ}C$, todas as placas apresentavam crescimento.

Expressão da Proteína yps por Indução com IPTG

Seguiu-se o protocolo facultado em Materiais e Métodos para sobre-expressão em *E.coli*. Para este estudo seleccionou-se uma colónia da placa inoculada com o clone $C_{1.4}A2$. Os valores de densidade óptica medida a 600 nm nos ensaios a $18^{\circ}C$ e $30^{\circ}C$ são os que constam na Tabela 10.

Tabela 10 – Valores de DO_{600} medidos nos ensaios a 18°C e a 30°C, para as várias concentrações de IPTG.

		Incubação a 18°C					Incubação a 30°C				
[IPTG] (mM)		0,0	0,2	0,4	0,8	1,0	0,0	0,2	0,4	0,8	1,0
Tempo (h)	0	0,601	0,581	0,566	0,566	0,565	0,610	0,647		0,577	0,583
	2	1,058	0,951	0,955	0,935	0,915	1,169	1,139	1,106	1,052	1,069
	5	1,575	1,313	1,286	1,260	1,324	1,661	1,704	1,623	1,580	1,584
	18	2,129	2,029	1,891	1,918	1,816	2,133	2,108	2,072	2,036	2,026

Ao longo deste protocolo retiraram-se várias alíquotas das fracções solúvel e insolúvel, às quais se adicionou tampão de amostra para análise em gel SDS-PAGE, e verificou-se se ocorrera ou não expressão da proteína yps e em que condição é que a expressão foi mais significativa. Prepararam-se géis SDS-PAGE para cada tempo das duas temperaturas de incubação, carregando-se em cada um as amostras respectivas. Os géis foram corridos nas condições referidas em Materiais e Métodos.

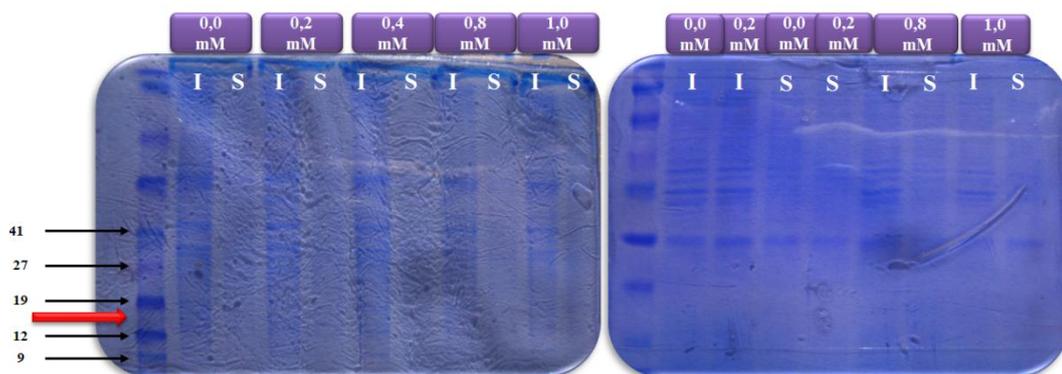


Figura 42 – Gel SDS-PAGE para $t = 0h$ a 18°C (esquerda) e 30°C (direita). I – fracção insolúvel; S – fracção solúvel.

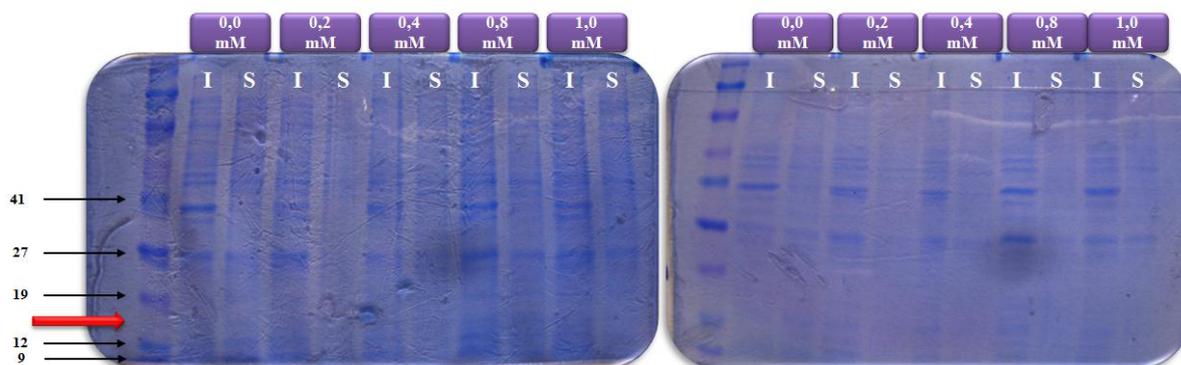


Figura 43 – Gel SDS-PAGE para $t = 2h$ a 18°C (esquerda) e 30°C (direita). I – fracção insolúvel; S – fracção solúvel.

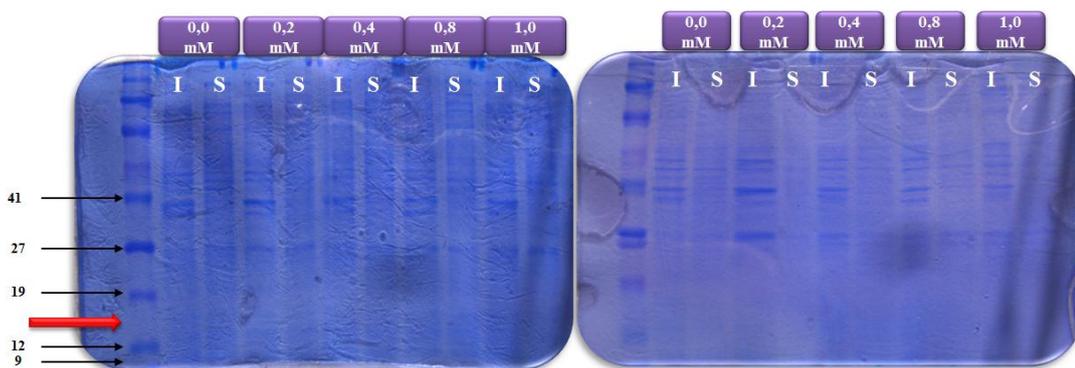


Figura 44 – Gel SDS-PAGE para t = 5h a 18°C (esquerda) e 30°C (direita). I – fracção insolúvel; S – fracção solúvel.

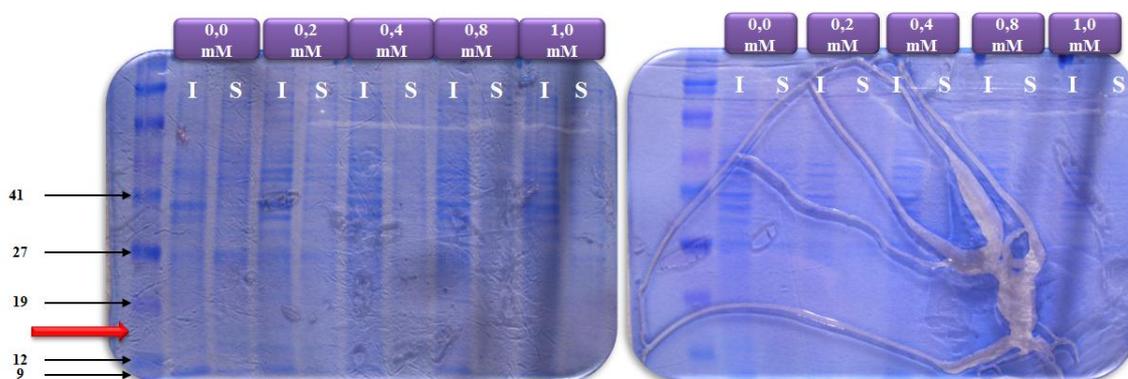


Figura 45 – Gel SDS-PAGE para t = 18h a 18°C (esquerda) e 30°C (direita). I – fracção insolúvel; S – fracção solúvel.

O fragmento clonado (C_1) corresponde ao fragmento de 465 pb que se traduz numa proteína de 15 730.1 Da, pelo que esperar-se-ia o aparecimento de bandas no gel em valores aproximados a este peso molecular; o que não se verificou nas duas temperaturas de incubação utilizadas.

Assim, optou-se pela expressão desta proteína através de um procedimento experimental de auto-indução.

Expressão da Proteína yps por Auto-Indução

Para posterior comparação entre os dois métodos de expressão, a selecção de colónias foi feita a partir da mesma placa, inoculada com o clone $C_{1,4}A2$. De modo a tornar-se mais célere e claro o processo de avaliação da eficácia da auto-indução na expressão da proteína

yps, procedeu-se à purificação da proteína expressa e as alíquotas retiradas das fracções solúveis e insolúveis antes e depois do passo de purificação, foram analisadas em gel SDS-PAGE. Desconhecendo-se o comportamento da proteína quando sujeita a purificação por cromatografia de afinidade com níquel imobilizado, optou-se por se utilizar um gradiente de imidazol para avaliar qual a concentração óptima deste composto, para obtenção de maior quantidade da proteína em estudo e com maior grau de pureza.

Prepararam-se géis SDS-PAGE com as alíquotas das fracções insolúvel e solúveis a diferentes concentrações de Imidazol.

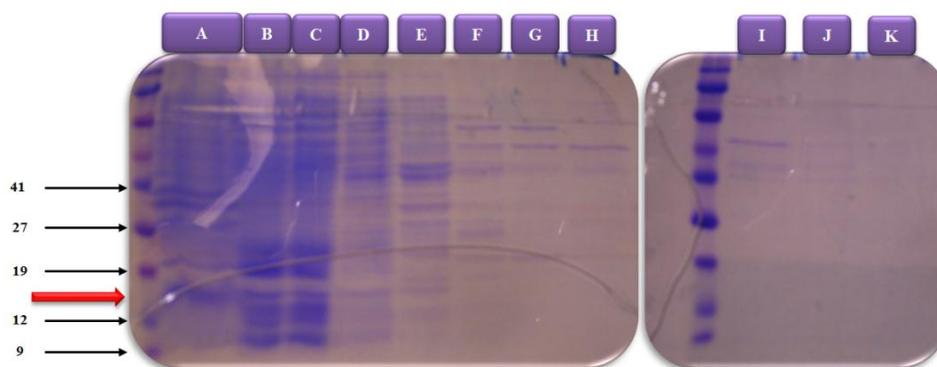


Figura 46 – Gel SDS-PAGE para o ensaio de auto-indução. A-Pellet, B-Sobrenadante, C-Extracto bruto, D-10 mM Imidazol, E-25 mM Imidazol, F-50 mM Imidazol, G-100 mM imidazol, H-150 mM Imidazol, I-300 mM Imidazol, J-400 mM Imidazol, K-500 mM Imidazol.

De igual modo, esperar-se-ia a presença de uma banda próxima do peso molecular de 15 kDa; o que se verifica até uma concentração de Imidazol de 25 mM.

Discussão

Pretendeu-se com este trabalho, a sobre-expressão e purificação da proteína Ypsilon Schachtel, para posterior análise de qual ou quais regiões da sua sequência proteica interagem com as outras proteínas do complexo RNP, nomeadamente com a proteína exu; e com o mRNA transportado por este complexo.

Um dos maiores desafios deste trabalho foi a optimização de todos os passos experimentais que culminaram com a expressão de uma parte da região da proteína ypsilon schachtel. Este processo pela sua minúcia e carácter de experimentação “tentativa-erro” foi bastante moroso, mas os resultados obtidos representam as bases de sustentação para posteriores estudos desta proteína, nomeadamente a determinação da sua estrutura tridimensional.

Uma experiência de clonagem típica envolve, de um modo geral, 5 componentes: gene ou fragmento do gene a clonar; endonucleases de restrição; vector de clonagem e/ou expressão; enzima DNA ligase e células hospedeiras apropriadas. E o principal objectivo de uma experiência deste género é a produção em larga escala de uma proteína de interesse.

O ponto de partida para o estudo da proteína yps foi a análise da sua sequência proteica: a extremidade C-terminal apresenta sucessivas repetições de aminoácidos, podendo indicar alguma desordem nesta região. Posto isto, tentou-se perceber qual a melhor abordagem para o seu estudo: clonar-se a totalidade do gene que codifica a proteína ou clonar-se apenas a região que se pensa estar envolvida nas possíveis interacções yps-exu, yps-RNP e yps-mRNA. Com o auxílio da ferramenta bioinformática GlobPlot confirmou-se que a região C-terminal se encontra altamente desordenada, comparativamente à região N-terminal, devido essencialmente às sucessivas repetições existentes na extremidade C-terminal (Figura 18). Assim sendo, supôs-se que a sequência nucleotídica envolvida nas interacções que se pretendiam estudar se encontra na região N-terminal e tomaram-se como objecto de estudo dois fragmentos desta região: um fragmento composto pelos primeiros 145 aminoácidos, e um fragmento menor no intervalo de aminoácidos 55-145, ambos contemplando o domínio CSD nas suas regiões N-terminal. Este domínio encontra-se presente em todas as proteínas Y-box e

a sua estrutura tridimensional já se encontra resolvida (Figura 11), pelas técnicas de cristalografia de raios-X e RMN. A existência de estudos de cristalografia indica que esta região é suficientemente estável para ser cristalizada, o que poderá ajudar no processo de cristalização da proteína yps.

Depois de escolhida a região da proteína mais passível de um estudo *in vitro*, procurou-se definir quais os vectores de clonagem e expressão mais apropriados para o estudo pretendido. Na escolha dos vectores foi necessário atenderem-se a vários factores: os vectores teriam que ser estáveis depois de transformados nas células hospedeiras, permitindo a sua replicação; deveriam ter a capacidade de serem replicados nas células hospedeiras; serem de pequena dimensão, o que facilita o processo de transformação; possuírem marcas de selecção que permitam a distinção entre células transformadas e células não transformadas; possuírem sequências de ligação ao ribossoma de modo a que a proteína de interesse possa ser expressa; e por fim, devem ser de fácil isolamento a partir dos lisados bacterianos. Com base nestes factores, escolheram-se três vectores para a clonagem do DNA de *yps*, posterior expressão da proteína yps em *E.coli* (que reconhece a origem de replicação dos vectores) e detecção da mesma: pGEX-6P1, pET-14b e pET-15b. Estes vectores têm em comum a marca de resistência a antibióticos (ampicilina) e o facto da expressão da proteína codificada no gene neles inserido, poder ser induzida pela adição do composto IPTG ao meio de crescimento bacteriano.

A clonagem de um fragmento de DNA no local de clonagem múltipla do vector pGEX-6P1 resulta na produção de uma proteína de fusão GST. O sistema GST tem por base a expressão indutível de genes (ou fragmentos destes) enquanto fusões com *Schistosoma japonicum* GST. A expressão da proteína ocorre sob o controlo do promotor *tac*, o qual é induzido pelo análogo da lactose, o IPTG. O vector pGEX-6P1 contém também o gene *lacI*^q. O produto deste gene é uma proteína repressora que se liga à região do operador do promotor *tac*, impedindo a expressão até que esta seja induzida pelo IPTG, controlando assim de forma restricta a expressão do fragmento inserido no vector. As proteínas de fusão GST são purificadas a partir dos lisados bacterianos através da técnica de cromatografia de afinidade com glutathiona imobilizada: as proteínas com a fusão GST ligam-se à coluna e as impurezas são removidas pelas sucessivas lavagens. A proteína de interesse pode ser eluída pela adição de glutathiona reduzida à coluna e a cauda GST pode ser clivada através de uma protease que

reconhece a sequência proteica expressa entre a GST e a proteína de interesse. Estirpes de *E.coli* que não produzam proteases citoplasmáticas (Lon, OmpT, DegP ou HtpR), ajudam na expressão das proteínas de fusão ao minimizarem os efeitos de degradação proteolítica pelo hospedeiro. A estirpe de *E.coli* BL21 é das mais indicadas para a expressão das proteínas de fusão GST pois não produz as enzimas OmpT e Lon, permitindo elevados níveis de expressão da proteína alvo. Relativamente aos vectores pET-14b e pET-15b, a expressão de proteínas clonadas ocorre sob o controlo do promotor T7 no caso do pET-14b ou sob o controlo do promotor T7lac no caso do pET-15b. Em ambos os casos a estirpe bacteriana mais indicada para a expressão da proteína alvo é, tal como no sistema pGEX, a estirpe BL21 (DE3) de *E.coli*. Nesta estirpe, a expressão da T7 RNA polimerase, que controla o promotor T7, é induzida pelo IPTG (adicionado ao meio de crescimento bacteriano na fase *mid-log* de crescimento). No entanto, níveis basais da T7 RNA polimerase promovem a tradução, apesar de ténue, do gene alvo, o que poderá resultar num abrandamento do crescimento celular se a proteína a produzir for tóxica para as células. Assim, no caso de se suspeitar que a proteína a ser expressa poderá ser tóxica (no caso da proteína yps, esta informação era desconhecida) para as células hospedeiras, a clonagem no vector pET-15b é a melhor opção. Este plasmídeo possui um operador *lac* a jusante do promotor T7 e tem também uma cópia do gene *lacI*. O repressor adicional do gene *lac* codificado pelo gene *lacI* reprime o promotor T7lac e também o gene cromossomal que codifica a T7 RNA polimerase, inibindo quase totalmente a expressão da T7 RNA polimerase. Com a adição do indutor IPTG deixa de haver repressão e a expressão da T7 RNA polimerase leva à síntese da proteína. Ambos os vectores têm uma cauda de fusão com seis histidinas que permite a detecção e/ou purificação da proteína alvo através da técnica de cromatografia de afinidade com metal imobilizado. No estudo exposto nesta tese, o metal escolhido foi o níquel.

De acordo com a escolha da sequência da proteína que se pretendia estudar e com os vectores de clonagem/expressão seleccionados para o efeito, foi possível desenharem-se *primers* específicos para a amplificação dos dois fragmentos de cDNA que se pretendia estudar. No desenho dos *primers* (Tabela 3), consideraram-se alguns factores, tais como: a região 3' de ambos os *primers* tinha de ser complementar com a sequência de cDNA a amplificar; as regiões 3' dos *primers* não deviam ser complementares entre si, de modo a evitar-se a formação de dímeros de *primers*; e nas extremidades 5' colocaram-se algumas bases não complementares ao cDNA para permitir a introdução de locais de restrição em cada

uma das extremidades dos produtos de PCR. Nos *primers* N001, N55 e 145C introduziram-se duas sequências de restrição para permitir a clonagem dos fragmentos amplificados em diferentes plasmídeos a partir de um único par de *primers*. Salienta-se que estas sequências de restrição têm que existir no local de múltipla clonagem dos vectores seleccionados, e não podem estar presentes dentro da sequência nucleotídica dos fragmentos em estudo, o que levaria à produção de proteína truncada. Outro pormenor importante no desenho dos *primers* utilizados é a existência de um codão STOP no *primer reverse*, o que permite, aquando da expressão da respectiva proteína, restringir a tradução proteica até este ponto, evitando desta forma a produção indefinida de proteína até o ribossoma encontrar um codão STOP. O programa de amplificação utilizado consistiu numa amplificação em dois passos (Figura 19): no primeiro passo, composto por 10 ciclos, amplificam-se os fragmentos individuais de *yps* a partir do cDNA integrado no plasmídeo pOT2; no segundo passo, com 40 ciclos, amplificaram-se os fragmentos de DNA obtidos no primeiro passo. Nos primeiros 10 ciclos a sequência do cDNA emparelha apenas com parte do *primer*. Passados estes 10 ciclos, existe quantidade suficiente de DNA para hibridar com a totalidade do *primer*. As temperaturas de fusão de partida foram calculadas através do *software* Serial Cloner (com base na composição dos *primers* e da sua complementaridade com o DNA alvo), tendo sido ajustadas à medida que se optimizava a reacção. Outros factores preponderantes para esta reacção, e que foram sujeitos a optimização, foram as concentrações dos diferentes constituintes da mistura reaccional. Verificou-se, por exemplo, que volumes de cDNA superiores a 0,5 µl inibiam a acção da Taq polimerase, havendo uma fraca amplificação dos fragmentos. A detecção dos amplicões foi feita através da técnica de electroforese em gel de agarose. Os produtos puros das reacções de PCR efectuadas são os apresentados nas Figuras 24 (fragmentos para integrar em pET-SUMO-28a) e 25 (fragmentos para integrar em pGEX-6P1, pET-14b e pET-15b). Nestes géis de agarose é possível visualizarem-se bandas fluorescentes nos pesos moleculares esperados para os dois fragmentos amplificados, mas também algum arrastamento destas bandas ao longo do gel. Este arrastamento poderá dever-se a produtos inespecíficos resultantes da amplificação. De salientar que, contrariamente à maioria das DNA polimerases, a Taq polimerase não tem a capacidade de corrigir erros à medida que sintetiza DNA (acção exonuclease 3'→5'), podendo introduzir alguns erros embora ocorra com baixa frequência, 1 nucleótido em 9 000 (Tindall e Kunkel, 1988).

Tanto os plasmídeos (Figuras 21, 22 e 23) como os fragmentos de *yps* amplificados (Figuras 26 e 27), foram hidrolisados com as respectivas endonucleases de restrição, gerando-se em cada par vector/fragmento extremidades coesivas complementares. Os plasmídeos linearizados foram sujeitos a um tratamento prévio à reacção de ligação, que consistiu na remoção dos grupos 5'-fosfato por acção de uma enzima fosfatase alcalina, impedindo-se deste modo que os vectores recircularizassem sem que o fragmento de DNA tivesse sido inserido. Aquando da reacção de ligação, a fosfatase foi inactivada pelo aumento da temperatura da mistura reaccional, e os plasmídeos linearizados que apresentam o local de corte das duas enzimas ligam os fragmentos de DNA. A partir dos fragmentos de *yps* hidrolisados e puros, mais concentrados, ensaiaram-se as reacções de ligação com os três vectores, também hidrolisados e puros (Figura 28). A reacção de ligação entre o DNA do fragmento e o DNA do vector foi mediada pela enzima DNA ligase (neste caso, a enzima T4 DNA ligase) que catalisou a formação de uma ligação fosfodiéster entre o grupo 3'-hidroxilo de um segmento de DNA do fragmento e o grupo 5'-fosfato do DNA do vector. No entanto, na Figura 28 apenas são visíveis as bandas correspondentes aos vectores puros não hidrolisados; não se observando qualquer sinal referente às amostras das reacções de ligação efectuadas. A primeira conclusão é de que não ocorreu reacção de ligação com nenhum dos três vectores utilizados. Mas não havendo ligação, esperar-se-ia um padrão de duas bandas para cada amostra: uma banda referente ao vector linearizado (a alto peso molecular) e outra banda referente aos fragmentos de *yps*. A ausência deste padrão poderá indicar que o DNA (vectorial e de *yps*) não se encontrava numa concentração mínima para ser detectado e possivelmente, para se ligar. Assim, optou-se por se concentrarem os hidrolisados de *yps* e vectores por liofilização (Figura 29). O resultado obtido foi oposto ao esperado, as amostras liofilizadas apresentavam-se menos concentradas do que as amostras hidrolisadas anteriormente testadas (Figuras 26 e 27), podendo dever-se à degradação de DNA durante este processo. Mesmo assim, procedeu-se a nova reacção de ligação (com as amostras mais concentradas) apresentando-se o resultado na Figura 30. Surpreendentemente, visualiza-se um sinal correspondente ao vector pGEX-6P1. Comparando a posição destas bandas com o marcador de pesos moleculares e com a banda relativa a este vector puro e não digerido não foi possível concluir se o passo experimental anterior foi ou não bem-sucedido, pois a diferença de pesos moleculares entre o vector nativo puro e o vector recombinante é mínimo. Contudo, permitiu verificar a presença do vector nas amostras, podendo este vector ser ou não recombinante. Com base nestes resultados, procedeu-se à transformação dos clones

supostamente positivos em células competentes bacterianas, mas não se registou o crescimento de colónias nas placas inoculadas com a transformação, concluindo-se deste modo que os clones representavam falsos positivos.

Perante os resultados indesejados obtidos pelo método de clonagem directa, optou-se por se inserir um passo de clonagem extra que permitiria averiguar se o insucesso na reacção de clonagem se devia a problemas com os fragmentos de *yps* ou com os vectores de clonagem. Nesta fase do trabalho, entendeu-se dar prioridade à clonagem nos vectores pET uma vez que estudos anteriores demonstravam que a expressão de proteína a partir de vectores pGEX poderia resultar na formação de corpos de inclusão (agregados de proteína insolúveis) e porque se pretendia testar a utilidade de um protocolo de expressão de proteína por auto-indução apenas aplicável aos vectores pET. Apesar de existirem protocolos para isolamento de proteína a partir de corpos de inclusão, seria necessário proceder ao “*refolding*” da proteína, que é um processo pouco eficiente, moroso e com uma taxa de sucesso muito reduzida. Caso a clonagem nos vectores pET não resultasse, proceder-se-ia à optimização da clonagem em pGEX com posterior expressão e sequestro da proteína dos corpos de inclusão. Utilizou-se o *kit* comercial referido em Materiais e Métodos; e que tem por base a clonagem dos fragmentos em estudo no vector pNZY28a, que, após transformação em células competentes apropriadas, permite a selecção de clones positivos pelo método de triagem azul/branco. Os primeiros clones seleccionados, apesar de apresentarem uma banda referente a um dos fragmentos de *yps* no tamanho esperado (Figura 31), a hidrólise dos mesmos permitiu concluir trataram-se de falsos positivos (Figura 32), pois não se visualizou o padrão de duas bandas esperado (banda do vector e banda do fragmento). Seleccionaram-se novos clones a partir das mesmas placas de crescimento. Novamente, confirmou-se a presença dos fragmentos de *yps* através de PCR de colónia (Figuras 33 e 34) e do vector recombinante (Figura 35). Neste último, visualizam-se bandas no tamanho aproximado dos vectores pET-14b e pET-15b, podendo estas amostras representarem clones positivos. Para confirmar, hidrolizaram-se estas amostras com as endonucleases de restrição respectivas e verificou-se que apenas o clone identificado por C_{1.4} era de facto um positivo verdadeiro (Figura 36). Este clone corresponde ao fragmento de maior tamanho de *yps* (465 pb) preparado para clonar no vector pET-15b.

Este método permitiu a selecção de um fragmento de *yps* que se sabe agora ser passível de clonagem. Assim, purificou-se o fragmento C_{1.4} e prepararam-se duas reacções de ligação com o vector pET-15b, A e B, variando-se a razão vector/fragmento. Os produtos das reacções de ligação foram transformados em *E.coli*. Os vectores recombinantes purificados a partir dos clones positivos foram enviados para sequenciação (Figura 37), confirmando-se a inserção do fragmento de *yps* através das duas reacções de ligação, pET-15b x C_{1.4}A e pET-15b x C_{1.4}B (Figura 38). Dos resultados da sequenciação constata-se que a cauda de fusão de histidinas apresenta apenas 5 dos 6 resíduos iniciais. De notar, que a sequenciação automática em capilar (utilizada pela empresa contratada) não lê de modo eficiente as primeiras 40-60 bases das sequências nucleotídicas. Dada a posição do local de clonagem dos fragmentos no vector, a omissão do primeiro resíduo de histidina poderá dever-se a este motivo, e não à ausência deste resíduo na amostra sequenciada. Paralelamente à sequenciação, confirmou-se também a eficácia das reacções de ligação por PCR de colónia (Figura 39), por purificação do vector recombinante a partir de crescimento em meio líquido (Figura 40) e ainda pela hidrólise enzimática do mesmo (Figura 41). De notar que na hidrólise enzimática dos produtos da reacção de ligação se utilizaram as enzimas NdeI e HindIII, em vez do par enzimático NdeI-BglII. Esta alteração deve-se à perda da sequência de restrição BglII quando se ligou o fragmento de *yps* ao vector cortado com BamHI (tanto a BglII como a BamHI deixam sequências coesivas idênticas, apesar da sequência de corte diferir nas bases dos extremos do local de corte), tendo sido necessário escolher-se nova enzima de acordo com o local de múltipla clonagem do vector e garantindo que a mesma sequência não estava presente no fragmento integrado. Todos estes testes confirmaram tratar-se de clones positivos, ou seja, foi possível clonar o fragmento de 465 pb do gene *yps* no vector de clonagem e expressão pET-15b.

Por fim, procedeu-se à transformação do clone positivo em células BL21 (DE3) de *E.coli*, tendo-se obtido crescimento em todas as placas inoculadas com o produto da reacção de transformação. Perante este resultado positivo procedeu-se à expressão da proteína *yps* segundo dois métodos distintos, mas ambos com base nas propriedades do sistema pET: expressão induzida pela adição de IPTG e expressão por auto-indução. Tal como referido no sistema pGEX, a expressão de proteínas recombinantes em *E.coli* leva frequentemente à produção de corpos de inclusão. Contudo, apesar de se formarem estes agregados insolúveis, alguma proteína encontra-se no extracto solúvel. O sistema de expressão pET apresenta-se

como uma vantagem pois aumenta a razão de solúvel/insolúvel de proteína. Por outro lado, qualquer factor que diminua a velocidade de síntese da proteína resulta geralmente no aumento da quantidade de proteína na forma solúvel. Entre os factores que afectam a solubilidade da proteína alvo, aqueles que podem ser facilmente manipulados em laboratório são: a temperatura a que ocorre a indução do crescimento bacteriano, a composição do meio de cultura e a concentração. Verificou-se que o crescimento dos meios de cultura induzidos a 37°C aumenta a probabilidade da proteína se acumular em corpos de inclusão, enquanto que crescimentos a 30°C leva à produção de maior quantidade de proteína solúvel e activa (Schein, 1989). Outra estratégia eficaz na produção de maior quantidade de proteína solúvel, é a indução a baixas temperaturas (15-20°C) com crescimento prolongado por várias horas. De acordo com estes dados, escolheram-se duas temperaturas para o ensaio de indução com IPTG, 18°C e 30°C. A cada temperatura incubaram-se os crescimentos preparados com as células BL21 (DE3) transformadas com o vector recombinante, que foram induzidos a diferentes concentrações de IPTG, para análise do efeito da concentração de IPTG na expressão da proteína yps. Ao longo das 18h de crescimento bacteriano às temperaturas indicadas, retiraram-se várias alíquotas dos crescimentos a tempos diferentes e mediram-se as respectivas DO (Tabela 10). Por comparação dos valores de DO às duas temperaturas, verifica-se que houve maior taxa de crescimento de bactérias a 30°C, como esperado. No entanto, apesar de se verificar um crescimento óptimo nestas condições, não significa que tenha havido expressão de proteína solúvel. Da análise em gel SDS-PAGE das fracções solúvel/insolúvel das alíquotas recolhidas, verifica-se a ausência de bandas no tamanho previsto do fragmento de proteína, ≈ 16 kDa (Figuras 42, 43, 44 e 45). Este facto poderá indicar que não houve expressão da proteína. Com o intuito de se excluir um possível problema com a expressão da proteína, experimentou-se novo protocolo de expressão, desta feita utilizando-se o protocolo de auto-indução. Este método é muito menos moroso e trabalhoso que o protocolo anterior: as células transformadas são inoculadas num meio de crescimento rico para auto-indução e as células bacterianas crescem até à saturação, sem ser necessário medirem-se os valores de DO. Além disso, a densidade da cultura e a concentração da proteína alvo por volume de cultura são geralmente mais altas que as obtidas com a indução por IPTG. De facto, através deste método foi possível expressar-se a proteína yps (Figura 46), que se encontra presente nas fracções insolúvel (A) e solúvel (B) do crescimento bacteriano, numa concentração de proteína de $\approx 1,4$ mg/L.

A purificação da proteína foi feita através da cromatografia por afinidade com níquel imobilizado. Ao carregar-se a coluna com o sobrenadante do meio de crescimento (fracção solúvel) a proteína nele existente deveria ligar-se ao níquel imobilizado nas partículas da coluna através das histidinas que constituem a sua cauda de fusão. De seguida, lavagens sucessivas com concentrações crescentes de imidazol, removem a maior parte das impurezas presentes nesta fracção e que não interagem com o níquel imobilizado. A eluição da proteína deveria ocorrer a concentração elevada de imidazol que compete com a proteína pela ligação ao metal imobilizado, permitindo que a proteína seja eluída da coluna. Ao longo deste protocolo experimental retiraram-se várias alíquotas para posterior análise da expressão da proteína yps em gel SDS-PAGE e para avaliar a sua quantidade com o reagente de Bradford. As amostras A, B e C apresentavam a maior concentração de proteína. Este resultado é concordante com o esperado, uma vez que estas fracções contêm a maioria das proteínas extraídas das células: pellet, sobrenadante e extracto bruto, respectivamente. Da amostra C para a amostra D observou-se uma diminuição significativa da quantidade de proteína eluída; resultado confirmado no respectivo gel SDS-PAGE (Figura 46). Contrariamente ao ensaio de expressão induzida com IPTG, neste ensaio de auto-indução detectou-se a presença da proteína yps em todas as fracções até uma concentração de 25 mM de imidazol (Figura 46, A-E). No entanto, este resultado indica que a cauda de histidinas poderá não se ter ligado covalentemente ao metal da coluna, pois foi eluída na sua totalidade a baixas concentrações de imidazol. Nas restantes fracções recolhidas (Figura 46, E-K) não se torna a visualizar uma banda no peso molecular pretendido. Qualquer problema com a cauda de histidinas é excluído à partida pois a sequenciação demonstrou a existência de 5 das 6 histidinas na extremidade N-terminal dos fragmentos (Figura 38). Uma possível explicação para este fenómeno é o tamanho relativamente pequeno da cauda de histidinas: se esta não estiver totalmente exposta ao solvente (enrolando-se sobre a restante proteína, por exemplo) poderá não conseguir interagir com o metal presente na coluna.

De futuro, seria interessante tentar a purificação deste vector recombinante pelo mesmo método, mas com outros metais imobilizados na coluna. Colunas de Cu^{2+} , Co^{2+} e Zn^{2+} permitem a separação do produto de fusão das proteínas bacterianas com rendimentos de purificação até 95 %.

Outra alternativa para aumentar a expressão de yps, seria testar a utilidade da fusão da proteína yps com a proteína SUMO (SMT3), um pequeno modificador relacionado com a ubiquitina; através da clonagem dos fragmentos de DNA no plasmídeo pET-SUMO-28a. Este sistema de expressão consiste na clonagem do gene de interesse no vector de expressão pET-SUMO e a consequente expressão da proteína alvo com a proteína SUMO traduzida na região N-terminal. Ao contrário da ubiquitina, a SUMO está envolvida na estabilização e localização de proteínas *in vivo*, apresentando-se como uma mais-valia para a expressão proteica quando traduzida como cauda de fusão na proteína alvo. A SUMO sendo uma proteína monomérica e altamente solúvel, vai aumentar a solubilidade das proteínas recombinantes que lhe estejam associadas. Depois de expressa a proteína de fusão, a região da SUMO pode ser clivada por uma protease altamente específica e activa (ULP-1), resultando na obtenção da proteína nativa. No entanto, a proteína de fusão tem que ter também disponível uma cauda de histidinas (por exemplo) para que seja possível detectar e purificar a proteína depois de clivada a região SUMO. Estas condições de expressão e detecção/purificação encontram-se reunidas no vector pET-SUMO-28a, apresentando-se este vector como uma excelente alternativa para obtenção de grande quantidade da proteína yps pura. Não só aumentará a probabilidade de obter maiores quantidades de proteína solúvel, como a cauda de histidinas estará mais exposta ao solvente (como comprovado em estudo onde foi utilizado o mesmo vector de expressão). Por fim, o facto de a SUMO ser reconhecida por uma protease altamente específica, aumentará o rendimento da clivagem da proteína de fusão, permitindo obter maiores quantidade de yps pura.

Depois de se obter quantidade suficiente da proteína yps pura, o próximo passo será a preparação da proteína para determinação da sua estrutura tridimensional. A análise da estrutura da proteína yps por cristalografia de raios-X pressupõe, de modo resumido, os seguintes passos: concentração da proteína, cristalização da proteína, irradiação do cristal da proteína com feixe de raios-X, recolha de dados de difracção e respectivo tratamento informático. Assim que se determine a estrutura desta região da proteína yps (1-145 aminoácidos). Em paralelo, proceder-se-á a estudos de ligação da yps ao mRNA e à exu (também se encontra a ser estudada no mesmo laboratório). Com estes estudos será possível perceber o modo de interacção entre a yps e o complexo ribonucleoproteico, respondendo assim se de facto a interacção deste complexo com o mRNA ocorre através da yps ou de outra proteína do mesmo complexo.

Bibliografia

- Ainger, K. *et al.* 1993. Transport and localization of exogenous myelin basic protein mRNA microinjected into oligodendrocytes. *The Journal of Cell Biology* 123 (2): 431-441.
- Alberts, B. *et al.* 2002. *Molecular Biology of the Cell*, 4^a ed., Garland Science, New York.
- Anthony, J.F., *et al.* 2000. *An Introduction to Genetic Analysis*, 7^a ed., W. H. Freeman, New York.
- Arn, E.A. e Macdonald, P.M. 2001. RNA localization goes direct. *Developmental Cell* 1: 155-156.
- Bashirullah, A., Cooperstock, R.L. e Lipshitz, H.D. 1998. RNA localization in development. *Annual Review of Biochemistry* 67: 335-394.
- Bastock, R. e St. Johnston, D. 2008. *Drosophila* oogenesis. *Current Biology* 18 (23): R1082-R1087.
- Bullock, S.L. 2007. Translocation of mRNAs by molecular motors: think complex? *Seminars in Cell & Developmental Biology* 18: 194-201.
- Bullock, S.L. e Ish-Horowicz, D. 2001. Conserved signals and machinery for RNA transport in *Drosophila* oogenesis and embryogenesis. *Nature* 414: 611-616.
- Cha, B.J., Koppetsch, B.S. e Theurkauf, W.E. 2001. In vivo analysis of *Drosophila bicoid* mRNA localization reveals a novel microtubule-dependent axis specification pathway. *Cell* 106: 35-46.
- Cooley, L. e Theurkauf, W.E. 1994. Cytoskeletal functions during *Drosophila* oogenesis. *Science* 266: 590-596.
- Deshler, J.O. *et al.* 1998. A highly conserved RNA-binding protein for cytoplasmic mRNA localization in vertebrates. *Current Biology* 8: 489-496.
- Dollar, G. *et al.* 2002. Rab11 polarization of the *Drosophila* oocyte: a novel link between membrane trafficking, microtubule organization, and *oskar* mRNA localization and translation. *Development* 129: 517-526.
- Driever, W. e Nüsslein-Volhard, C. 1988. A gradient of *bicoid* protein in *Drosophila* embryos. *Cell* 54: 83-93.
- Emery, P. 2000. RNase protection assay. *Methods in Molecular Biology* 362: 343-348.
- Fonseca, P.M.S. 1997. Caracterização da γ -Tubulina23C em *Drosophila melanogaster*. Tese de Doutorado. Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto.
- Gasteiger, E. *et al.* 2005. Protein identification and analysis tools on the ExPASy Server. *In The Proteomics Protocols Handbook* (Walker, J.M. ed.), pp 571-607, Human Press Inc., Totowa.
- Gigliotti, S. *et al.* 2004. Molecular genetics of oogenesis in *Drosophila melanogaster*. *Invertebrate Survival Journal* 1: 72-81.
- Gilbert, S.E. 2003. *Developmental Biology*, 7^a ed., pp 307-355 e pp 711-748, Sinauer Associates Inc., Sunderland.
- Goldstein, L.S.B. e Fyrberg, E.A. 1994. *Methods in Cell Biology*. Volume 44 – *Drosophila melanogaster*: practical uses in cell and molecular biology. Academic Press Inc., California.
- Griffiths *et al.* 2000. Sex chromosomes and sex-linked inheritance. *In An Introduction to Genetic Analysis*, 7^a ed., W. H. Freeman, New York.
- Gunkel, N. *et al.* 1998. Localization-dependent translation requires a functional interaction between the 5' and 3' ends of *oskar* mRNA. *Genes & Development* 12: 1652-1664.

- Hudson, A. e Cooley, L. 2010. *Drosophila* kelch functions with cullin-3 to organize the ring canal actin cytoskeleton. *The Journal of Cell Biology* 188 (1): 29-37.
- Huynh, J.R. e St. Johnston, D. 2004. The origin of asymmetry: early polarisation of the *Drosophila* germline cyst and oocyte. *Current Biology* 14: R438-R449.
- Irion, U. e St. Johnston, D. 2007. *Bicoid* RNA localization requires specific binding of na endosomal sorting complex. *Nature* 445: 554-558.
- Irion U. *et al.* 2006. Miranda couples *oskar* mRNA/Staufen complexes to the *bicoid* mRNA localization pathway. *Developmental Biology* 297: 522-533.
- Jansen, R.P. 2001. mRNA localization: message on the move. *Nature Reviews Molecular Cell Biology* 2: 247-256.
- Januschke, J. *et al.* 2002. Polar transport in the *Drosophila* oocyte requires dynein and kinesin I cooperation. *Current Biology* 12: 1971-1981.
- Jeffrey, W. B., Tomlinson, C. R. e Brodeur, R. D. 1983. Localization of actin messenger RNA during early ascidian development. *Developmental Biology* 99: 408-417.
- Jensen, L.J. *et al.* 2008. STRING 8 – A global view on proteins and their functional interaction in 630 organisms. *Nucleic Acids Research* 37: D412-D416.
- Johnstone, O. E Lasko, P. 2001. Translational regulation and RNA localization in *Drosophila* oocytes and embryos. *Annual Review of Genetics* 35: 365-406.
- Kindler, S. *et al.* 2005. RNA transport and local control of translation. *Annual Review of Cell and Developmental Biology* 21: 223-245.
- Kloc, M. e Etkin, L.D. 2005. RNA localization mechanisms in oocytes. *Journal of Cell Science* 118: 269-282.
- Kuersten, S. e Goodwin, E.B. 2003. The power of the 3'UTR: translational control and development. *Nature Reviews Genetics* 4: 626-637.
- Kugler, J.M. e Lasko, P. 2009. Localization, anchoring and translational control of *oskar*, *gurken*, *bicoid* and *nanos* mRNA during *Drosophila* oogenesis. *Fly* 3 (1): 15-28.
- Landsman, D. 1992. RNP-1, na RNA-binding motif is conserved in the DNA-binding cold shock domain. *Nucleic Acids Research* 20 (11): 2861-2864.
- Lasko, P. 1999. RNA sorting in *Drosophila* oocytes and embryos. *The FASEB Journal* 13: 421-433.
- Lasko, P. 2003. Cup-link *oskar* RNA localization and translational control. *The Journal of Cell Biology* 163 (6): 1189-1191.
- Lécuyer, E. *et al.* 2007. Global analysis of mRNA localization reveals a prominent role in organizing celular architecture and functions. *Cell* 131: 174-187.
- Linding, R. *et al.* 2003. GlobPlot: exploring portein sequences for globularity and disorder. *Nucleic Acids Research* 31: 3701-3708.
- Leptin, M. 1999. Gastrulation in *Drosophila*: the logic and the celular mechanisms. *The EMBO Journal* 18 (12): 3187-3192.
- Leptin, M. e Affolter, M. 2004. *Drosophila* gastrulation: identification of a missing link. *Current Biology* 14: R480-R482.

- Lin, H., Yue, L. e Spradling, A.C. 1994. The *Drosophila* fusome, a germlike-specific organelle, contains membrane skeletal proteins and functions in cyst formation. *Development* 120: 947-956.
- Lin, M.D. *et al.* 2006. *Drosophila* decapping protein 1, dDcp1, is a component of the *oskar* mRNP complex and directs its posterior localization in the oocyte. *Developmental Cell* 10: 601-613.
- Lipshitz, H.D. e Smibert, C.A. 2000. Mechanisms of RNA localization and translational regulation. *Current Opinion in Genetics & Development* 10: 476-488.
- Lodish, H. *et al.* 2008. *Molecular Cell Biology*, 6^a ed., pp 165-214, W. H. Freeman and Company, New York.
- Macdonald, P.M. 1990. *Bicoid* mRNA localization signal: phylogenetic conservation of function and RNA secondary structure. *Development* 110: 161-171.
- Macdonald, P.M., Luk, S.K.S. e Kilpatrick, M. 1991. Protein encoded by the *exuperantia* gene is concentrated at sites of *bicoid* mRNA accumulation in *Drosophila* nurse cells but not in oocytes or embryos. *Genes & Development* 5: 2455:2466.
- Malhotra, A. 2009. Tagging for protein expression. *Methods in Enzymology* 463: 239-258.
- Mansfield, J.H., Wilhelm, J.E. e Hazeirigg, T. 2002. Ypsilon Schachtel, a *Drosophila* Y-box protein, acts antagonistically to Orb in the *oskar* mRNA localization and translation pathway. *Development* 129: 197-209.
- Matsumoto, K. e Bay, B.H. 2005. Significance of the y-box proteins in human cancers. *Journal of Molecular and Genetic Medicine* 1 (1): 11-17.
- Matsumoto, K. e Wolffe, A.P. 1998. Gene regulation by y-box proteins: coupling control of transcription and translation. *trends in Cell Biology* 8: 318-323.
- Micklem, D.R. 1995. mRNA localisation during development. *Developmental Biology* 172: 377-395.
- Mische, S. *et al.* 2007. Direct observations of regulated ribonucleoprotein transporta cross the nurse cell/oocyte boundary. *Molecular Biology of the Cell* 18: 2254-2263.
- Mohr, E. e Richter, D. 2001. Messenger RNA on the move: implications for cell polarity. *The International Journal of Biochemistry & Cell Biology* 33: 669-679.
- Munro, T.P. *et al.* 2006. A repeated IMP-binding motif controls *oskar* mRNA translation and anchoring independently of *Drosophila melanogaster* IMP. *The Journal of Cell Biology* 172 (4): 577-588.
- Nakamura, A. *et al.* 2001. Me31B silences translation of oocyte-localizing RNAs through the formation of cytoplasmic RNP complex during *Drosophila* oogenesis. *Development* 128: 3233-3242.
- Nasmyth, K. e Jansen, R.P. 1997. The cytoskeleton in mRNA localization and cell differentiation. *Current Opinion in Cell Biology* 9: 396-400.
- Oleynikov, Y. e Singer, R.H. 1998. RNA localization: diferente zipcodes, same postman?. *trends in Cell Biology* 8: 381-383.
- Palacios, I.M. 2007. How does na mRNA find its way? Intracellular localisation of transcripts. *Seminars in Cell & Developmental Biology* 18: 163-170.
- Palacios, I.M. e St. Johnson, D. 2001. Getting the message across: the intracelular localization of mRNAs in higher eukaryotes. *Annual Review of Cell and Developmental Biology* 17: 569-614.
- Petterson, E.F. *et al.* 2004. UCSF – Chimera – A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 25 (13): 1605-1612.

- Riechmann, V. e Ephrussi, A. 2001. Axis formation during *Drosophila* oogenesis. *Current Opinion in Genetics & Development* 11: 374-383.
- Riechmann, V. e Ephrussi, A. 2004. Par-1 regulates *bicoid* mRNA localisation by phosphorylating Exuperantia. *Development* 131: 5897-5907.
- Rubin, G.M. e Lewis, E.B. 2000. A brief history of *Drosophila*'s contributions to Genome research. *Science* 287: 2216-2218.
- Sander, K. 1976. Specification of the basic body pattern in insect embryogenesis. *Advances in Insect Physiology* (J.E. Treherne), pp 125-238, Academic Press Inc., London.
- Schein, C.H. 1989. Production of soluble recombinant proteins in bacteria. *Nature* 7: 1141-1149.
- Schnorrer, F., Bohmann, K. e Nüsslein-Volhard, C. 2000. The molecular motor dynein is involved in targeting Swallow and *bicoid* RNA to the anterior pole of *Drosophila* oocytes. *Nature Cell Biology* 2: 185-190.
- Schnorrer, F. *et al.* 2002. γ -Tubulin37C and γ -tubulin ring complex protein 75 are essential for *bicoid* RNA localization during *Drosophila* oogenesis. *Developmental Cell* 3: 685-696.
- Schroeder, M.D. *et al.* 2004. Transcriptional control in the segmentation gene network of *Drosophila*. *PLoS Biol* 2 (9): 1396-1410.
- Slack, J.M.W. 1991. From egg to embryo – Regional specification in early development, 2^a ed., pp 53, Cambridge University Press, New York.
- St. Johnston, D. 1995. The intracellular localization of Messenger RNAs. *Cell* 81: 161-170.
- St. Johnston, D. 2005. Moving messages: the intracellular localization of mRNAs. *Nature Reviews Molecular Cell Biology* 6: 363-710.
- St. Johnston, D. e Nüsslein-Volhard, C. 1992. The origin of pattern and polarity in the *Drosophila* embryo. *Cell* 68: 201-219.
- Studier, F. W., 2005. Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification* 41 (1): 207-234.
- Tindall, K.R. e Kunkel, T.A. 1988. Fidelity of DNA synthesis by the *Thermus aquaticus* DNA Polymerase. *American Chemical Society* 27 (16): 6008-6013.
- Theurkauf, W.E. e Hazelrigg, T.I. 1998. In vivo analyses of cytoplasmic transport and cytoskeletal organization during *Drosophila* oogenesis: characterization of a multi-step anterior localization pathway. *Development* 125: 3655-3666.
- Thieringer, H.A. *et al.* 1997. Identification and developmental characterization of a novel Y-box protein from *Drosophila melanogaster*. *Nucleic Acids Research* 25 (23): 4764-4770.
- Wang, H. *et al.* 2002. Dendritic BC1 RNA: functional role in regulation of translation initiation. *The Journal of Neuroscience* 22 (23): 10232-10241.
- Wang, S. e Hazelrigg, T. 1994. Implications for *bcd* mRNA localization from spatial distribution of *exu* protein in *Drosophila* oogenesis. *Nature* 369: 400-403.
- Wilhelm, J.E. e Smibert, C.A. 2005. Mechanisms of translational regulation in *Drosophila*. *Biology of the Cell* 97: 235-252.
- Wilhelm, J.E. e Vale, R.D. 1993. RNA on the move: the mRNA localization pathway. *The Journal of Cell Biology* 123 (2): 269-274.

- Wilhelm, J.E. *et al.* 2000. Isolation of a ribonucleoprotein complex involved in mRNA localization in *Drosophila* oocytes. *The Journal of Cell Biology* 148 (3): 427-439.
- Wilkie, G.S., Dickson, K.S. e Gray, N.K. 2003. Regulation of mRNA translation by 5'- and 3'-UTR-binding factors. *TRENDS in Biochemical Sciences* 28 (4): 182-188.
- Wilsch-Bräuninger, M., Schwarz, H. e Nüsslein-Volhard, C. 1997. A sponge-like structure involved in the association and transport of maternal products during *Drosophila* oogenesis. *The Journal of Cell Biology* 139 (3): 817-829.
- Wolpert L. 1969. Positional information and the spatial pattern of cellular differentiation. *Journal of Theoretical Biology* 25: 1-47.
- Zimyanin, V.L. *et al.* 2008. In vivo imaging of *oskar* mRNA transport reveals the mechanism of posterior localization. *Cell* 134: 843-853.

Anexos

Anexo I - Vectores de Clonagem e Expressão

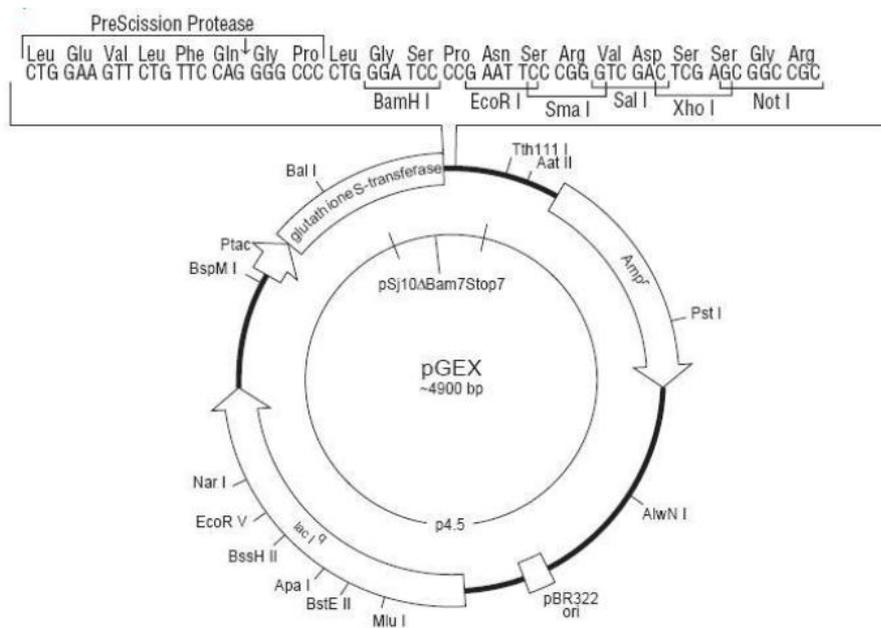


Figura I.1 – Esquema do vector pGEX-6P1

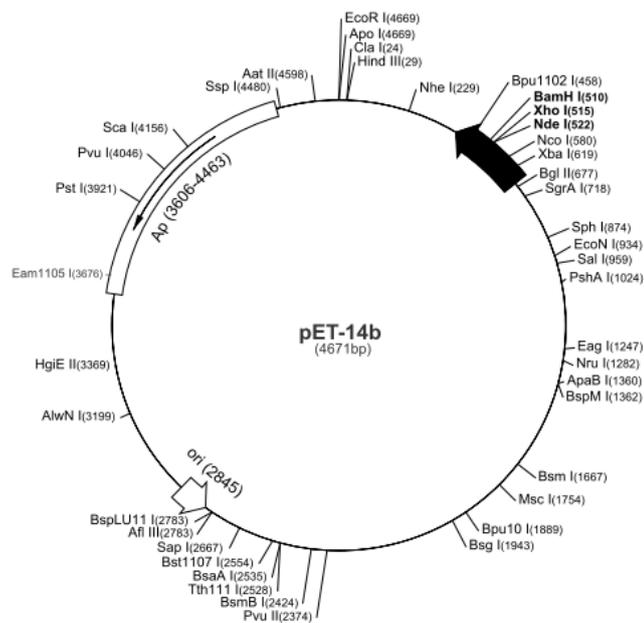


Figura I.2 – Esquema do vector pET-14b

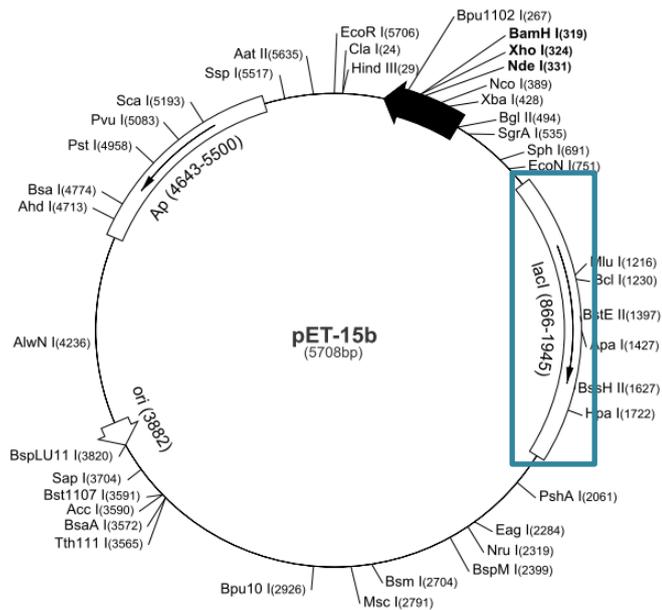


Figura I.3 – Esquema do vector pET-15b

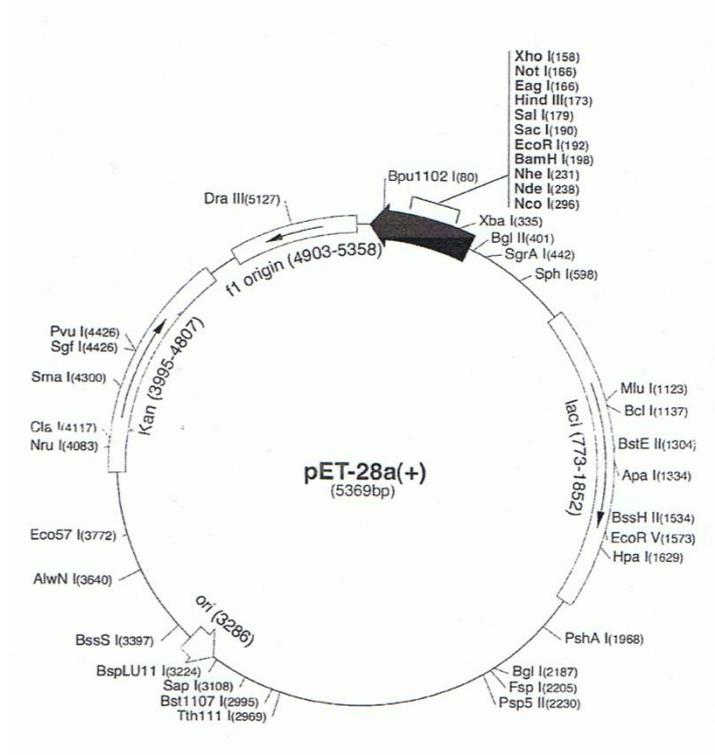
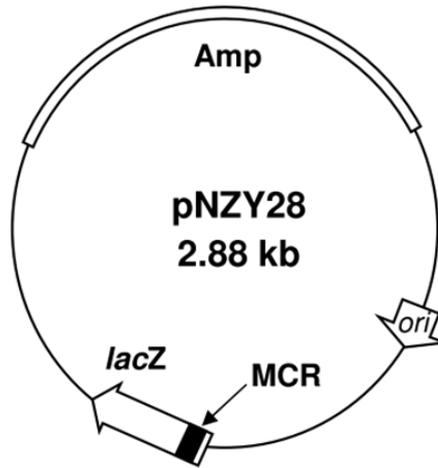


Figura I.4 – Esquema do vector pET-SUMO-28a. A seqüência SUMO (SMT3) encontra-se inserida entre os locais de restrição NdeI e BamHI



Múltipla Região de Clonagem (MCR)

KpnI

HindIII *EcoRI* *PstI* *SalI* *XbaI* *BamHI* *SpeI* *NdeI* ***EcoRV*** *BamHI* *SmaI* *SacI* *EcoRI*

AAGCTTGAATTCTCTGCAGGTCGACTCTAGAGGATCCACTAGTCATATGG**GATATC**GGATCCCCGGGTACCGAGCTCGAATTC

Figura I.5 – Esquema do vector pNZY28a

Anexo II – Marcadores de DNA e Proteína

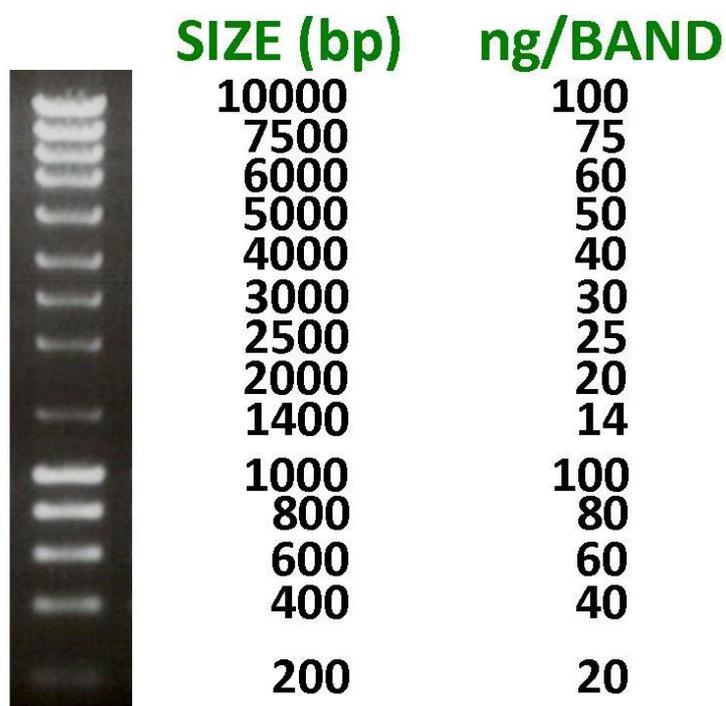


Figura II.1 – Padrão de Bandas do marcador de DNA *NZYDNA Ladder III* (NZYTech®)

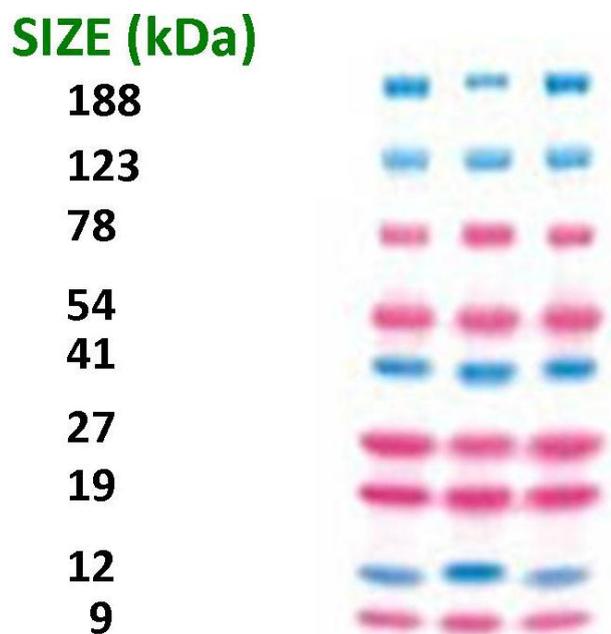


Figura II.2 – Padrão de Bandas do marcador de proteínas *NZYColour Protein Marker* (NZYTech®)

Anexo III – E-Gel iBase Power System e E-Gel Safe Imager Real-Time Transilluminator

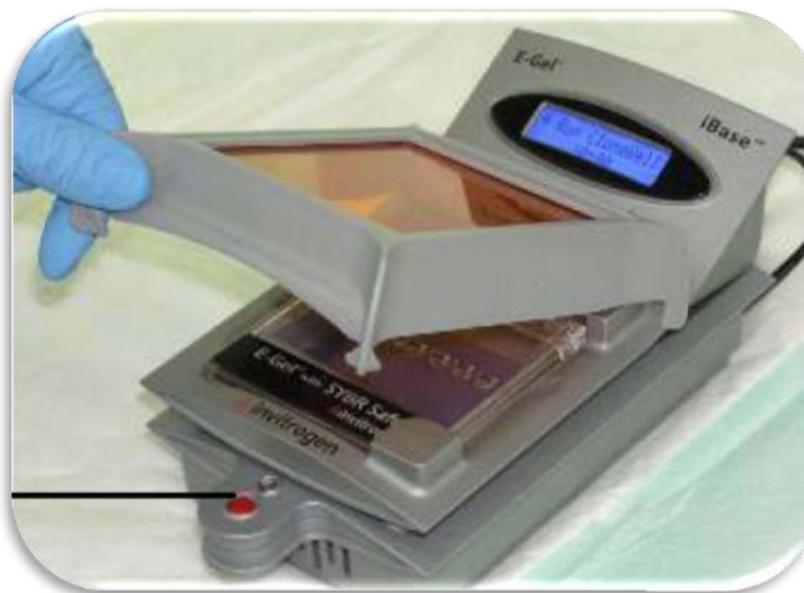


Figura III.1 – Imagem dos dois sistemas acoplados.