

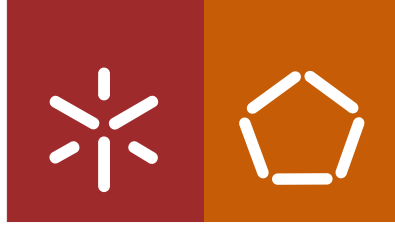


Universidade do Minho
Escola de Engenharia

Bruno Miguel Nogueira Valente

**Expressão e caracterização de uma
lectina recombinante de interesse
biomédico**

outubro de 2015



Universidade do Minho
Escola de Engenharia

Bruno Miguel Nogueira Valente

**Expressão e caracterização de uma
lectina recombinante de interesse
biomédico**

Dissertação de Mestrado
Mestrado Integrado em Engenharia Biomédica
Ramo de Engenharia Clínica

Trabalho efetuado sob a orientação da
Professora Doutora Lucília Domingues
e coorientação da
Doutora Carla Oliveira

DECLARAÇÃO

Nome: Bruno Miguel Nogueira Valente

Título da Dissertação: Expressão e caracterização de uma lectina recombinante de interesse biomédico

Orientador(a): Professora Doutora Lucília Domingues

Coorientador(a): Doutora Carla Oliveira

Ano de conclusão: 2015

Designação de: Mestrado Integrado em Engenharia Biomédica

Ramo: Engenharia Clínica

Escola: de Engenharia

Departamento: de Engenharia Biológica

DE ACORDO COM A LEGISLAÇÃO EM VIGOR, NÃO É PERMITIDA A REPRODUÇÃO DE QUALQUER PARTE DESTA TESE/TRABALHO

Universidade do Minho, ____ / ____ / _____

Assinatura:

Agradecimentos

Em primeiro lugar quero agradecer quero agradecer à minha orientadora, a Prof.^a Dr.^a Lucília Domingues, pela oportunidade que me deu e pelo o meu enquadramento no grupo de trabalho. Agradeço-lhe ainda toda a disponibilidade, o apoio e os conhecimentos que me transmitiu.

Agradeço ainda à minha coorientadora, a Dr.^a Carla Oliveira, por me encaminhar durante, incansavelmente, durante este grande desafio. Agradeço-lhe por todos os momentos de incentivo, pela paciência nas fases mais complicadas e por fim, por toda a experiência partilhada e pelos ensinamentos transmitidos.

Quero também agradecer àqueles que diariamente me acompanharam e ajudaram ao longo deste trabalho, nomeadamente aos meus colegas do grupo de investigação e do laboratório de trabalho. Refiro-me essencialmente ao Rui, à Liliana, à Joana e à Goreti, não só pela entreatajuda no trabalho mas também pelos bons momentos e laços de amizade criados.

Agradeço ainda a todos os meus amigos que, longe ou perto, acompanharam este meu percurso e que tanto ajudaram a quebrar rotinas quando foram necessárias. Em muito ajudou a partilha de momentos de descontração no contorno das adversidades que muito encontrei ao trilhar este percurso académico.

Por último, um especial e sentido agradecimento a toda a minha família, nomeadamente àqueles que de mais perto me acompanharam e tornaram este momento possível. À minha tia Vitória que sempre se preocupou e zelou pelo meu melhor, ao meu irmão que de uma forma ou de outra me acompanhou nesta fase final e claro, aos meus pais que representam os pilares que sustentaram todo o meu percurso. A eles estarei eternamente grato!

Por fim, mais que um agradecimento, uma dedicatória ao meu avô Zé que este ano, por infortúnio do destino, já não se encontra presente entre nós mas que acredito que este seria um momento de orgulho para ele.

Obrigado a todos!

Abstract

Lectins as a protein group establish diversified families with distinct properties. They are commonly distributed in nature and can be found in various living organisms like bacteria, fungi, plants, animals and even humans, acting as mediators of biomolecular processes. The frutalin is a lectin derived from plants and the family of jacalinas. Isolated from breadfruit seeds, *Artocarpus incisa* or *Artocarpus altilis*, it is a homotetrameric glycoprotein with specifically recognition of α -D-galactose residues.

The lectins are distinguished of other proteins for its activity of specific linking to the glycoconjugates receptors what allows to act as a mediator of biological processes with extreme importance for the biomedicine. Previous studies had shown the express frutalina in *P. pastoris* has apoptotic activity in human tumor cells. Still much has to be discovered for their potential however, to the present day, it is possible to evaluate and draw conclusions from the important role that frutalin may have in the biomedical field.

This study was primarily engaged with the cloning and production of a recombinant frutalin in *E. coli* containing, instead of the native linker, a site for cleavage by action of a protease for later separation of β and α chains as a way to evaluate the impact of the cleavage in the activity and properties of the protein isoform cloned in *E. coli*.

This work involved different stages since the genetic construct of the sequences of interest by polymerase chain reaction process and its transformation into a cloning vector suitable for production and expression system in *E. coli* for further purification using immobilized metal on affinity chromatography. To structural characterization of the proteins was used two different technics, dynamic light scattering techniques and fast protein liquid chromatography.

The frutalin is essentially composed of two chains, β and α , linked together by a tetrapeptide linker. This work has been focus on the construction of recombinant forms that differ in the arrangement of the β and α chains relative to the linker region. This region was also subject of changes to be replaced by another recognition site for a *tobacco etch virus* protease (TEV protease) that allows its cleavage. This strategic process acknowledges the influence of cleavage and subsequent refolding in the protein activity. Thus were designed two strategies, one FTL₁ strategy with a β +TEV+ α construction and a parallel strategy, FTL₂ with a α +TEV+ β construction. For each strategy were used different expression vectors, pETM-10 and pETM-20, selected for its properties. Both have a tag (His-tag) very commonly used in purification strategies cited by IMAC technique and nickel affinity to metal ions. For pETM-20 vector showing a

particularity, a Trx fusion partner described as an enhancer of the expression, solubility and protein stability.

Genetic alterations to the sequence of frutalin were observed. For both strategies, we saw the success of the PCR processes, which resulted in the correct inclusion of TEV cleavage section and, for FTL₂ strategy, the change of the arrangement of β and α chains relative to the binding region. It was verified the correct cloning into the expression vectors pETM-10 and pETM-20 that have been properly transformed into *E. coli* strain *Rosetta* observing good levels of soluble expression of His₆-FTL and Trx-His₆-TEV-FTL. Moreover, it was obtained good levels of purification, particularly for His₆-FTL₂ and Trx-His₆-TEV-FTL₂. To remove the associated tags and the separation of the chains was used, to Trx-His₆-TEV-FTL₂, the TEV protease as a digestive agent. Cleavage occurred naturally *in vitro*, verifying good results in particular in the cleavage of Trx-His₆-TEV region. For recovery of the protein of interest was used the technique of reverse purification, a procedure which proved to be effective and which led to the fractional interest in a pure and clean way. Results reinforced the correct cleavage of the region containing Trx by action of TEV protease.

The respective structural analysis proved to be preliminary so, further investigation is required. The observed values suggest that the protein is not analyzed in a tetrameric structure form but into a dimer structure. A more empirical analysis indicates that these values may reflect a wrong protein refolding or even possible errors in the cleavage process.

Resumo

As lectinas como grupo proteico constituem famílias diversificadas com propriedades distintas. Estão amplamente distribuídas na natureza podendo ser encontradas nos mais variados organismos vivos desde bactérias, fungos, plantas, animais e até humanos, atuando como mediadores de processos biomoleculares. A frutalina é uma lectina de origem em plantas e da família das jacalinas. Isolada das sementes de fruta-pão, *Artocarpus incisa* ou *Artocarpus altilis*, uma glicoproteína homotetramérica e que reconhece especificamente resíduos de α -D-galactose.

As lectinas destacam-se de outras proteínas pela sua atividade de ligação específica a recetores glicoconjugados o que permite atuarem como mediadoras de processos biológicos com extrema importância para a biomedicina. Estudos anteriores mostram que a frutalina expressa em *P. pastoris* tem atividade apoptótica em células tumorais humanas. Ainda muito há por descobrir relativamente ao seu potencial no entanto, aos dias de hoje, é já possível fazer uma avaliação e retirar conclusões do importante papel que a frutalina pode ter na biomédica.

Este estudo tem por objetivo a clonagem e expressão de frutalina recombinante em *E. coli* contendo, em vez do *linker* nativo, um local para clivagem por ação de uma protease para posterior separação das cadeias β e α como forma de avaliar o impacto da clivagem nas propriedades e atividade da isoforma clonada em *E. coli*, com a finalidade de caracterização das versões de frutalina recombinante obtidas. A fase de caracterização torna-se importante na medida em que permite avaliar e relacionar a isoforma de frutalina analisada e a influência do processo de clivagem das cadeias na forma e atividade adquirida.

Este trabalho envolveu diferentes fases desde a construção genética das sequências de interesse por processos de reação em cadeia da polimerase e a sua transformação em vetores de clonagem apropriados à produção e expressão em sistema *E. coli* para posterior purificação com recurso a cromatografia de afinidade com iões metálicos imobilizados (IMAC) até à caracterização estrutural das proteínas com recurso a técnicas de dispersão de luz dinâmica (*Dynamic Light Scattering*) e de cromatografia rápida líquida de proteínas (*Fast Protein Liquid Chromatography*)

A frutalina é estruturalmente constituída por duas cadeias, β e α , ligadas entre si por um *linker* tetrapeptídico. Este trabalho prendeu-se com a construção de formas recombinantes que diferem na disposição das cadeias β e α em relação à região ligante. Esta região foi também sujeita a alterações, sendo substituída por uma outra zona de reconhecimento para uma

protease *tobacco etch virus* (TEV protease) que possibilite a sua clivagem. Este processo tem por objectivo avaliar a influência da clivagem e posterior *refolding* na atividade proteica da frutalina. Desta forma projetaram-se duas estratégias, uma estratégia FTL₁ com uma construção β +TEV+ α e paralelamente uma estratégia FTL₂ com uma construção α +TEV+ β . Para cada construção foram usados vetores de expressão distintos, pETM-10 e pETM-20, selecionados pelas suas propriedades. Ambos possuem um *tag* de fusão (*His-tag*) muito comumente utilizado em estratégias de purificação pela já referida técnica de IMAC e com afinidade aos íons metálicos de níquel. O vetor pETM-20 apresenta ainda uma outra particularidade, um parceiro de fusão Trx descrito como potenciador da expressão, solubilidade e estabilidade proteica.

Alterações genéticas à sequência da frutalina foram observadas. Para ambas as estratégias verificou-se o sucesso dos processos de PCR e que se traduziram na correta inclusão do local de clivagem TEV bem como, para a estratégia FTL₂, a alteração da disposição das cadeias β e α relativamente à região ligante. Comprovou-se a correta clonagem nos vetores finais de expressão pETM-10 e pETM-20 que foram devidamente transformados na estirpe de *E. coli Rosetta*, observando-se bons níveis de expressão solúvel de His₆-FTL e Trx-His₆-TEV-FTL. Obtiveram-se ainda bons níveis de purificação, nomeadamente para His₆-FTL₂ e Trx-His₆-TEV-FTL₂. Para remoção dos *tags* associados e separação das cadeias utilizou-se, para Trx-His₆-TEV-FTL₂, a TEV protease como agente digestivo. A clivagem ocorreu naturalmente *in vitro*, verificando-se bons resultados nomeadamente na clivagem da região Trx-His₆-TEV. Para recuperação da proteína de interesse recorreu-se à técnica de purificação reversa, procedimento este que mostrou ser eficaz e que permitiu obter as frações de interesse numa forma pura e limpa como se pôde comprovar, resultados que reforçaram a correta clivagem da região que contém Trx por ação da TEV protease.

A respetiva análise estrutural mostrou ser preliminar pelo que existe a necessidade de ser mais explorada. Os valores observados sugerem que a proteína analisada não se encontra na forma estrutural tetramérica mas possivelmente numa estrutura de dímeros. Uma análise mais empírica indica que estes valores podem refletir um mau *refolding* proteico ou até possíveis erros no processo de clivagem.

Índice

Abstract	v – vi
Resumo	vii – viii
Índice de Figuras	xiii – xvi
Índice de Tabelas	xvii
Abreviaturas	xix

Capítulo 1. Introdução 1 - 22

1.1. Enquadramento geral	3
1.2. Biologia Molecular	4 – 5
1.3. Lectinas, proteínas versáteis	6 – 8
1.4. Lectinas relacionadas com a Jacalina (JRL)	8 – 11
1.5. Frutalina, uma lectina de plantas	12 – 13
1.6. <i>Escherichia coli</i> , sistema de expressão de proteínas	14 – 17
1.7. Potencial biomédico de uma frutalina recombinante	18 – 19
1.8. Objetivos do estudo de uma frutalina recombinante	20 – 22

Capítulo 2. Materiais e Métodos 23 – 40

2.1. Esterilização de materiais, soluções e meios de cultura	25
2.2. Estirpes	25
2.3. Vetores	25 – 26
2.4. <i>Primers</i>	26
2.5. Meios e condições de crescimento	27
2.6. Conservação de culturas bacterianas	27
2.7. Estratégias de clonagem de sequências de DNA	28
2.7.1. Reação PCR de construção e amplificação da sequência β +TEV+ α	28
2.7.2. Reação PCR de construção e amplificação da sequência α +TEV+ β	28
2.8. PCR de colónia	29

2.9. Purificação de DNA por kit comercial	29
2.10. Purificação de DNA em gel de agarose por kit comercial	29
2.11. Extração de DNA genómico de plasmídeo por “Miniprep-Caseiro”	30
2.12. Quantificação de DNA	30
2.13. Eletroforese de DNA em gel de agarose	31
2.14. Ligação ao vetor pGEM-T Easy	32
2.15. Transformação de células competentes <i>E. coli</i>	32 – 33
2.16. Digestão com enzimas de restrição <i>NcoI</i> e <i>XhoI</i>	33
2.17. Ligação aos vetores de expressão pETM-10 e pETM-20	33
2.18. Preparação e transformação de células competentes <i>E. coli</i>	33 – 34
2.19. Crescimento e expressão proteica em “ <i>small-scale</i> ”	34
2.20. Produção proteica em “ <i>scale-up</i> ”	35
2.21. Lise celular por sonicação	35
2.22. Purificação por cromatografia de afinidade por iões metálicos	36 – 37
2.23. Diálise	37
2.24. Clivagem por <i>tobacco etch virus</i> protease	37
2.25. Análise proteica por eletroforese de géis SDS-PAGE	37 – 39
2.26. Revelação de géis SDS-PAGE por <i>Comassie Blue</i>	39
2.27. Caracterização por <i>Dinamic Light Scattering</i>	40
2.28. Caracterização por <i>Fast Protein Liquid Chromatography</i>	40

Capítulo 3. Resultados e Discussão 41 – 64

3.1. Construção de uma frutalina recombinante em <i>E. coli</i>	43 – 51
3.1.1. Construção por PCR de FTL1, sequência β +TEV+ α	45 – 49
3.1.2. Construção por PCR de FTL2, sequência α +TEV+ β	49 – 50
3.2. Clonagem em vetor intermédio pGEM-T Easy e transformação em células competentes <i>E. coli</i>	51 – 53
3.3. Clonagem em vetores de expressão pETM-10 e pETM-20 e transformação em células competentes <i>E. coli Rosetta</i>	53 – 57
3.4. Expressão primária de frutalina recombinante	57 – 59

3.5. Produção e purificação de frutalina recombinante expressa em <i>E. coli</i> <i>Rosetta</i>	59 – 62
3.6. Caracterização proteica por <i>Dinamic Light Scattering</i> e <i>Fast Protein Liquid Chromatography</i>	63 – 64
Capitulo 4. Conclusão e Perspectivas Futuras	65 – 70
Capitulo 5. Bibliografia	71 – 76
Capitulo 6. Anexos	77 – 79
Anexo 6.1	79
Anexo 6.2	79

Índice de Figuras

Figura 1.1: Imagem alusiva à ligação entre os elementos de estudo que caracterizam a Biologia Molecular (imagem adaptada da web)	5
Figura 1.2: Representação esquemática das merolectinas, hololectinas, chimerolectinas e superlectinas (<i>adaptado de Liu et al., 2010</i>)	8
Figura 1.3: (a) Representação esquemática da base molecular da apoptose induzida por uma lectina de planta mediada através das mitocôndrias e ou recetores em células tumorais; (b) Representação esquemática da autofagia induzida por uma lectina de planta mediada através da mitocôndria (<i>adaptado de Liu et al., 2010</i>)	10
Figura 1.4: Representação esquemática de um vetor plasmídeo (<i>adaptado de Sorensen and Mortensen et al., 2005a</i>)	15
Figura 3.1: Representação esquemática e estrutural da frutalina <i>template</i>	44
Figura 3.2: Representação esquemática e estrutural das formas recombinantes finais de frutalina. A) Forma recombinante final para a estratégia '1' com alteração da região ligante; b) Forma recombinante final para a estratégia '2' com alteração da região ligante e da disposição das cadeias β e α relativamente à região TEV	45
Figura 3.3: Construção e amplificação do produto final β +TEV. (M) Marcador NZY Ladder; (1) Banda correspondente à sequência β +TEV (tamanho molecular esperado de 92 bp)	46
Figura 3.4: Construção e amplificação do produto final TEV+ α . (M) Marcador NZY Ladder III; (1) Banda correspondente à sequência TEV+ α (tamanho molecular esperado de 429 bp)	48
Figura 3.5: Construção e amplificação do produto final β +TEV+ α . (M) Marcador NZY Ladder III; (1) Banda correspondente à sequência β +TEV+ α (tamanho molecular esperado de 500 bp)	49

Figura 3.6: Construção e amplificação do produto final α +TEV+ β . (M) Marcador NZY Ladder III; (1) Banda correspondente à sequência α +TEV+ β (tamanho molecular esperado de 500 bp) 51

Figura 3.7: Rastreamento de clones positivos da transformação em vetor pGEM-T Easy dos genes de ambas as sequências. (a) Representação esquemática da constituição do vetor pGEM-T Easy; (b) Representação esquemática da construção de ambas as estratégias em vetor pGEM-T Easy; (c) Despistagem das construções em vetor pGEM-T Easy para ambas as estratégias: (M) Marcador NZY Ladder III; (1) Clone selecionado para construção FTL1 (β +TEV+ α) em pGEM-T Easy; (2) Clone selecionado para construção FTL2 (α +TEV+ β) em pGEM-T Easy 52

Figura 3.8: Extração dos *inserts* do vetor pGEM-T Easy por digestão enzimática *Nco*I e *Xho*I. (M) Marcador NZY Ladder III; (1) Confirmação por extração da sequência da estratégia FTL₁ (β +TEV+ α com tamanho molecular esperado de 500 bp); (2) Confirmação da extração da sequência da estratégia FTL₂ (α +TEV+ β com tamanho molecular esperado de 500 bp) 53

Figura 3.9: Representação esquemática da constituição dos vetores de expressão selecionados, pETM-10 e pETM-20. (a) Descrição pETM-10: origem de replicação f1, marcador de seleção Kan, gene promotor T7/lac, *insert* EFH modificado de pET-24d e tag N-His/; (b) Descrição pETM-20: origem de replicação f1, marcador de seleção Amp, gene promotor T7/lac, *insert* RAGE modificado de pET-32, tags N-Trx e N-His e região de clivagem TEV 54

Figura 3.10: Confirmação das construções FTL₁ e FTL₂ nos plasmídeos de expressão pETM-10 e pETM-20, por análise de PCR de colônia de *E. coli* transformada. (a) Esquema da construção das clonagens em vetor pETM-10; (b) Esquema da construção das clonagens em vetor pETM-20; (c) Confirmação por PCR de colônia da clonagem de FTL₁ nos vetores pETM-10 (clones de 1 – 4, com seleção do clone 3) e pETM-20 (clones de 5 – 8, com seleção do clone 6); (d) Confirmação por PCR de colônia da clonagem de FTL₂ nos vetores pETM-10 (clones de 1 – 4, com seleção do clone 2) e pETM-20 (clones de 5 – 8, com seleção do clone 8); (M) Marcador molecular NZY Ladder III; (CN) Controlo negativo; (CP) Controlo positivo 55

Figura 3.11: Extração dos “*inserts*” dos vetores de expressão por digestão enzimática das endonucleases *NcoI* e *XhoI*. (M) Marcador NZY Ladder III; (1) Confirmação da extração da sequência β +TEV+ α do vetor pETM-20; (2) Confirmação da extração da sequência α +TEV+ β do vetor pETM-20; (3) Confirmação da extração da sequência β +TEV+ α do vetor pETM-10; (4) Confirmação da extração da sequência α +TEV+ β do vetor pETM-10. (tamanhos moleculares esperados de 500 bp) 56

Figura 3.12: Análise de His₆-FTL e Trx-His₆-TEV-FTL das amostras obtidas da expressão em *E. coli Rosetta*, por coloração de *Coomassie Blue* e 12% SDS_PAGE. (1) Expressão de His₆-FTL₁; (2) Expressão de His₆-FTL₂; (3) Controlo com vetor de expressão pETM-10; (M) Marcador proteico; (4) Controlo com vetor de expressão pETM-20; (5) Expressão de Trx-His₆-TEV-FTL₁; (6) Expressão de Trx-His₆-TEV-FTL₂ 58

Figura 3.13: Análise de His₆-FTL (tamanho molecular esperado de 18,9 kDa) e Trx-His₆-TEV-FTL (tamanho molecular esperado de 32 kDa) das amostras das fases solúveis e insolúveis obtidas da expressão em *E. coli Rosetta*, por coloração de *Coomassie Blue* e 12% SDS_PAGE. (M) Marcador proteico; (1) Amostra da expressão Trx-His₆-TEV-FTL₁ (fase solúvel); (2) Amostra da expressão Trx-His₆-TEV-FTL₂ (fase solúvel); (3) Amostra da expressão His₆-FTL₁ (fase solúvel); (4) Amostra da expressão His₆-FTL₂ (fase solúvel); (5) Amostra da expressão Trx-His₆-TEV-FTL₁ (fase insolúvel); (6) Amostra da expressão Trx-His₆-TEV-FTL₂ (fase insolúvel); (7) Amostra da expressão His₆-FTL₁ (fase insolúvel); (8) Amostra da expressão His₆-FTL₂ (fase insolúvel) 59

Figura 3.14: Análise de His₆-FTL₂ das amostras obtidas da purificação por afinidade a Níquel (IMAC-*Ni*), por coloração de *Coomassie Blue* e 12% SDS_PAGE. (M) Marcador proteico; (1) Amostra filtrada; (2) Amostra “*flow-through*”; (3) Amostra da passagem de solução “*binding*”; (4) Amostra da passagem de solução “*washing*”; (M) Marcador proteico; (5) Amostra da 1° eluição; (6) Amostra da 2° eluição; (7) Amostra da 3° eluição; (8) Amostra da passagem com solução EDTA 60

Figura 3.15: Análise de Trx-His₆-TEV-FTL₂ das amostras obtidas da purificação por afinidade a Níquel (IMAC-*Ni*), por coloração de *Coomassie Blue* e 12% SDS_PAGE. (M) Marcador proteico; (1) Amostra filtrada; (2) Amostra “*flow-through*”; (3) Amostra da passagem de solução “*binding*”;

(4) Amostra da passagem de solução “washing”; (M) Marcador proteico; (5) Amostra da 1° eluição; (6) Amostra da 2° eluição; (7) Amostra da 3° eluição; (8) Amostra da 4° eluição; (9) Amostra da passagem com solução EDTA 60

Figura 3.16: Análise da clivagem de Trx-His₆-TEV da FTL₂ por TEV protéase (27kDa). (M) Marcador molecular; (1) 1° Amostra clivada Trx-His₆-TEV da FTL₂; (2) 2° amostra clivada Trx-His₆-TEV da FTL₂ 61

Figura 3.17: Análise de His₆-FTL₂ clivada das amostras obtidas da purificação reversa por afinidade a Níquel (IMAC-Ni), por coloração de *Coomassie Blue* e 12% SDS_PAGE. (M) Marcador molecular; (1) Amostra Trx-His₆-TEV-FTL₂ clivada; (2) Amostra da 1° passagem amostra His₆-FTL₂ clivada; (3) Amostra da 2° passagem His₆-FTL₂ clivada; (4) Amostra da 1° passagem de solução “binding”; (5) Amostra da 2° passagem de solução “binding”; (6) Amostra da 1° passagem de solução “washing”; (7) Amostra da 1° eluição; (8) Amostra da passagem com solução EDTA 62

Figura 3.18: Gráfico do ensaio de DLS para uma frutalina recombinante, Trx-His₆-TEV-FTL₂ clivada, Distribuição “Tamanho vs Volume”, raio_{nm} esperado = 3,70 e raio_{nm} obtido = 2,35 ± 0,49 63

Figura 3.19: Cromatograma de exclusão molecular de BSA, frutalina nativa e Trx-His₆-TEV-FTL₂ clivada por sec-FPLC. (1) Amostra padrão de BSA; (2) Amostra padrão de frutalina nativa; (3) Amostra clivada Trx-His₆-TEV-FTL₂ [0,5 mg/ml] 64

Índice de Tabelas

Tabela 1.1: Características do sistema de expressão em <i>E. coli</i> . (adaptado de Demain et al., 2009)	15
Tabela 2.1: Estirpes microbianas usadas durante este trabalho	24
Tabela 2.2: Vetores de clonagem e expressão usados durante este trabalho	24 – 25
Tabela 2.3: <i>Primers</i> usados durante este trabalho. A “sublinhado” estão identificados os sítios de reconhecimento de enzimas de restrição usados nas clonagens. A “duplo sublinhado” está identificado o sítio de reconhecimento usado na clivagem por TEV protéase. A vermelho o correspondente ao codão STOP	26
Tabela 2.4: Tabela resumo da suplementação do meio LB com antibióticos para seleção das culturas das estirpes de <i>E. coli</i> transformadas. (Amp = Ampicilina; Kan = Kanamicina; Cloram = Cloranfenicol)	27
Tabela 2.5: Tamanhos moleculares das bandas do marcador NZYDNA Lader III	31
Tabela 2.6: Tamanhos das bandas moleculares do marcador Page Ruller Unstained Broad Range Protein Ladder	39
Tabela 3.1: Vetores de expressão e proteínas de fusão usados neste trabalho, e respetiva designação	57

Abreviaturas

a.a.: Aminoácido

bp: Par de bases

BSA: Bovin Serum Albumin

cDNA: Ácido desoxirribonucleico complementar

ConcA: Concanavalin A

DNA: Ácido desoxirribonucleico

DLS: Dynamic Light Scattering

DO: Densidade ótica

EDTA: Ethylenediamine Tetraacetic Acid

FPLC: Fast Protein Liquid Chromatography

FTL: Frutalina

FW: Forward

HeLa: Human Cervical Cancer Cells

JRL: Jacalin Related Protein

IMAC: Immobilized Metal Ion Affinity Chromatography

IMAC-Ni: Cromatografia de afinidade ao níquel

IPTG: Isopropyl β -D-1-thiogalactopyranoside

kDa: Kilodalton

LB: Meio Luria Bertani

mRNA: RNA mensageiro

Ni: Níquel

PCR: Polimerase Chain Reaction

RIP: Ribosome Inactivating Protein

RNA: Ácido ribonucleico

RV: Reverse

SDS-PAGE: Sodium dodecyl sulfate polyacrylamide gel electrophoresis

SEC: Cromatografia de exclusão molecular

SOC: Super Optimal Broth with Catabolite Repression

TAE: Tris-acetate EDTA

TEV: *Tobacco Etch Virus*

Trx: Tioredoxina

Capítulo 1

Introdução

1.1. Enquadramento geral

A Engenharia Biomédica é um ramo em constante crescimento onde a evolução de novos domínios para a resolução de problemas permitiram um desenvolvimento evolutivo à escala mundial que atualmente reconhecemos. Segundo Bronzino, em *“The Biomedical Engineering Handbook”*, esta aplica princípios de engenharia, das mais diversas áreas como a electrónica, física, informática, química, biologia entre outras, para assim conseguir descodificar e compreender os vários sistemas com o objetivo de poder modificar, controlar e monitorizar as suas funcionalidades”.

O sistema “corpo humano” apresenta uma complexidade e sofisticação de funcionamento que o torna um fascínio e desperta a atenção e a curiosidade por parte de toda a comunidade científica. Compartilhando esta ideia todos se focam em torno de um objetivo principal, desvendar os mistérios das áreas da ciência. Uma dessas áreas é a Biologia Molecular, que conseguiu criar um impacto revolucionário na investigação científica e no conhecimento das diversas áreas da biologia, como a engenharia genética ou a tecnologia do DNA recombinante que, seguramente, continuaram a ter um desenvolvimento e especial importância em aplicações para a sociedade futura. Os primeiros passos desta ciência e o seu progresso são sumarizados na história da década de 1940, época em que novos desenvolvimentos desta área começaram por tentar resolver problemas na tentativa de descrever a extensão dos temas associados às diversas disciplinas, nomeadamente a descrição de teorias ligadas à autorregulação das células e dos microrganismos. (*Ladik et al., 2004*) Como uma ciência, emerge ligada a uma geração mais analítica da área onde muitas e importantes moléculas ativas e as suas reações foram descobertas, e que lhe valeu o verdadeiro consenso de disciplina por volta de 1960. (*Rheinberger et al., 2009*)

Na altura, não era considerada uma disciplina em si contrariamente aos dias de hoje. Ganhou forma e encontrou um rumo ligado a outras ciências da vida, estando intimamente ligada a disciplinas como a biologia celular, a microbiologia, a genética e a bioquímica (*Festa et al., 2007*). Assim, pelo seu promissor potencial era vista como uma ciência que encarnava um novo sentido de vida e que hoje abre a porta a uma série de soluções, como por exemplo ao nível de tratamentos e meios de diagnóstico, de farmacologia e de biotecnologia.

1.2. Biologia Molecular

A história da Biologia Molecular é vasta enfrentado várias crises de dogmatismos, emergindo e adaptando novas abordagens e conexões. A evolução da Biologia Molecular mostra que as disciplinas são construções sociais historicamente e culturalmente situadas. (*Strasser et al., 1999*). Emerge em meados do século vinte, e pode ser definida como a mera descrição da biologia em termos moleculares, ou seja, o estudo da vida à escala molecular. A biologia molecular pode não ser vista como uma nova ciência mas como resultado cruzado entre áreas distintas como a bioquímica e a genética e com objetos de investigação bem definidos: os genes estão para a genética como a produção e expressão proteica/enzimática estão para a bioquímica. Este conceito fica definido com a definição da relação entre os objetos de estudo o que permitiu avançar e abarcar uma vasta gama de aplicações publicamente fascinantes e relevantes na história científica. (*Morange et al., 1998*)

Nas últimas décadas o estudo da estrutura genética e da função molecular tornou-se um dos campos mais promissores na área da biologia molecular pela existência de inúmeras proteínas de interesse biológico e importância biomédica com utilidade e benefícios na interação com a biociência e em aplicações clínicas que vão desde os diagnósticos às terapias. Simulações moleculares dinâmicas são usadas para caracterizar e descrever as propriedades e o comportamento das proteínas como a solubilidade molecular, alterações conformacionais e interações entre moléculas. (*Festa et al., 2007*). Estes avanços são o reflexo de investigações científicas com base em estudos teóricos e resultados práticos alcançados pelos diversos grupos e comunidades de investigadores, contudo continuaram a existir inúmeros problemas por resolver nesta área e que merecem toda a atenção necessária à sua reformulação, otimização e melhoria das soluções.

Por outras palavras a Biologia Molecular pode ser caracterizada pelos princípios repartidos das várias áreas da física e química para compreenderem os conceitos centrais do universo biológico, desde a replicação, expressão e evolução genética do DNA às limitações conformacionais, à dinâmica estrutural e rearranjos tridimensionais que definem as funções biológicas das proteínas. (*Echols et al., 2001*) O DNA e as proteínas são os protagonistas centrais da biologia molecular e o foco da genética clássica: genótipo, o rearranjo como unidade hereditária, e o fenótipo, expressão genética observada no organismo. Desta forma podemos resumir todo o princípio básico desta ciência ao raciocínio de que os genes, estabelecidos em

1.3. Lectinas, proteínas versáteis

Em 1888 Peter Hermann Stillmark deu início ao estudo das proteínas, promovendo a hemaglutinação de eritrócitos pela procura de fatores proteicos tóxicos em *Ricinus communis*. (Kennedy et al., 1995) Stillmark consegue desta forma descrever a presença de um fator proteico tóxico em extratos de *Ricinus communis* e que tinham a propriedade de aglutinar as células vermelhas do sangue. Esta foi a primeira lectina extraída e caracterizada e que foi denominada de ricina, mais tarde usada como modelo antigénico em estudos imunológicos. (Teixeira et al., 2012)

Mais tarde novos estudos e descobertas foram surgindo. James B. Summer, trinta anos depois isolou pela primeira vez uma hemaglutinina pura, a concanavalina A (ConcA), com a qual conseguiu descobrir que a sua atividade hemaglutinante era inibida pela sacarose, demonstrando em primeira mão a especificidade das lectinas pelos açúcares. Este resultado sugere que a hemaglutinação resulta da reação com os carboidratos expressos na superfície dos eritrócitos. (Teixeira et al., 2012)

Foi no ano de 1954 que o termo “lectina” foi proposto por Boyd e Shapleigh, termo baseado no fato de esta família de proteínas terem a capacidade de distinguir os eritrócitos nos diferentes tipos de sangue. O termo lectina inicialmente refere-se à capacidade de proteínas ligantes de carboidratos aglutinarem eritrócitos, particularmente do sangue humano. (Vasconcelos et al., 2004) Entretanto este termo evolui pelo seu comportamento de aglutinação, surgindo um novo termo sinónimo de lectina, a aglutinina, que refere a capacidade das proteínas ligantes aglutinarem eritrócitos e outras células. (Peumans and Van Damme et al., 1995) Por definição, são glicoproteínas e/ou proteínas oligoméricas de origem não imune que contêm domínios não catalíticos que permitem reconhecer e criar ligações a mono- ou oligossacáridos específicos sem que haja uma alteração estrutural dos carboidratos ou dos glicolípidos, no entanto, esta definição foi evoluindo no sentido de que as lectinas são proteínas de fusão com grupos de ligação a carboidratos e com domínios catalíticos, como é o exemplo da ricina, uma proteína tipo 2 RIP (do inglês Ribosome Inactivating Protein). (Teixiera et al., 2012)

As lectinas como grupo proteico constituem famílias diversificadas com propriedades distintas. Estão amplamente distribuídas na natureza podendo ser encontradas nos mais variados organismos vivos desde bactérias, fungos, plantas, animais e até humanos, atuando como mediadores de processos biomoleculares. (Teixeira et a., 2012) Isoladas de plantas, são comumente mais abundantes nas sementes vegetais mas podem ainda ser encontradas

noutros tecidos vegetais como as raízes, folhas, flores. Por observação, estas lectinas não mostram uma ação importante no crescimento e desenvolvimento da planta, contudo mostram um envolvimento marcante no processo de reconhecimento específico de glicanos externos ao sistema da planta, indicando um papel essencial na sua defesa. (Astoul *et al.*, 2002)

Vários avanços foram feitos na tentativa de organizar este grupo heterogêneo das lectinas de plantas sendo que inicialmente foram classificadas de acordo com o número de grupos específicos com base na sua habilidade de reconhecer e criar ligações específicas a açúcares. (Vandenborre *et al.*, 2011) Pneumans e Van Damme classificaram as lectinas de plantas de acordo com a sua estrutura e baseado na capacidade de interação e com outros glicoconjugados em: merolectinas (um domínio de ligação a carboidratos); hololectinas (possuem no mínimo duas regiões ligantes muito idênticas); superlectinas (possuem no mínimo duas regiões ligantes mas não idênticas) e chimerlectinas (são proteínas de fusão constituídas por duas cadeias diferentes, uma delas com atividade catalítica ou outra atividade biológica), representadas na imagem da figura 1.2. (Pneumans and Van Damme *et al.*, 1995)

A maioria das lectinas de plantas são hololectinas e têm a capacidade, pelos seus domínios ligantes de aglutinar células ou precipitados glicoconjugados. (Teixeira *et al.*, 2012) Por outro lado, as lectinas de plantas podem ainda ser subdivididas em cinco grupos de acordo com o monossacárido pelo qual mostram uma maior afinidade: D-manose/D-glucose, D-galactose/N-acetil-D-galactosamina, N-acetil-D-glucosamina, L-fucose e N-ácido acetilneuramínico. (Vasconcelos *et al.*, 2004)

Numa perspectiva evolucionária, sistemas de classificação mais elaborados aparecem tendo em consideração as sequências das lectinas disponíveis nas últimas décadas. Por análise do genoma/transcriptoma as lectinas de plantas estão divididas em doze famílias diferentes de acordo com a estrutura do domínio de ligação aos carboidratos. Uma dessas famílias pertence ao grupo de lectinas relacionadas com a jacalina (Jacalin-related-lectin ou JRL) que apresentam características específicas de *fold*ing com um ou mais domínios de ligação a carboidratos. (Vandenborre *et al.*, 2011)

Os últimos anos mostram avanços no estudo molecular, estrutural e bioquímico e revelam que, desta família de lectinas – JRL – fazem parte proteínas ligantes de carboidratos que partilham da mesma sequência da proteína de origem em *Artocarpus integrifolia* (jaca) e de nome corrente jacalina. (Pneumans *et al.*, 2000)

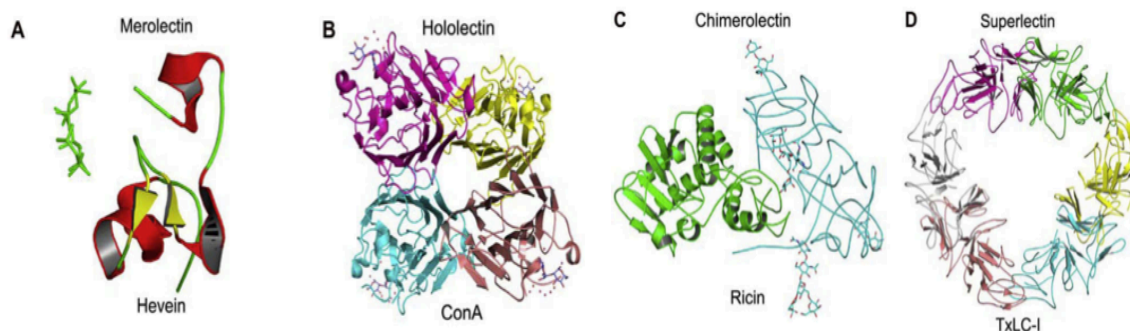


Figura 1.2: Representação esquemática das merolectinas, hololectinas, chimerolectinas e superlectinas (*adaptado de Liu et al., 2010*)

1.4. Lectinas relacionadas com a Jacalina (JRL)

A jacalina é uma lectina glicosilada, homotetramérica de quatro promotores idênticos. Consiste em duas cadeias polipeptídicas, uma mais leve, cadeia β de 20 resíduos de aminoácidos e uma cadeia mais pesada, cadeia α de 133 resíduos de aminoácidos associadas não covalentemente numa estrutura monomérica β -prisma $\alpha\beta$. A sua estrutura primária (preprolectina) consiste em 21 resíduos peptídicos sinal, 39 resíduos propéptidos e duas cadeias β e α ligados por um *linker* tetra peptídico (Thr-Ser- Ser-Asn). A conformação final resulta de processos de co- e pós-tradução em que os péptidos sinal são removidos pela co-tradução e a N-glicosilação resulta numa prolectina parcialmente glicosilada maturando a proteína pela excisão do *linker* de ligação e alteração do N-glicano. No final do processo, os promotores formam uma estrutura tetramérica. (*Bourne et al., 2002*)

A particularidade de ligação aos carboidratos da jacalina foram estudados pelo que, a existência de uma especificidade de interações diferentes subdivide este grupo em dois outros subgrupos: man-JRL, lectinas de específica ligação a resíduos de manose ou maltose e de domínios não clivados e formadas por um simples promotor; e T-antigen/gal- JRL, lectinas com facilidade de ligação à galactose e de domínios clivados por proteólise que resultam na separação das cadeias β e α . Aparentemente, man-JRL e gal-JRL não diferem apenas na estrutura sequencial como também na sua localização intracelular, pelo que explica as diferenças no seu processamento e na sua biossíntese. Foi referenciado que as gal-JRL são uma forma evolutiva das man-JRL, resultado da má repartição nos vacúolos onde a clivagem proteolítica permitiu um reconhecimento de glicanos presentes em proteínas. (*Fernandez-del-Carmen et al., 2012*)

As lectinas distinguem-se de outras proteínas de plantas pela sua específica atividade de ligação a carboidratos. As suas ligações específicas a 9írus99res glicoconjugados (carboidratos com estrutura complementar do sítio de ligação à lectina, podem ser: glicoproteínas, glicolípidos, polissacarídeos) atuam como mediadoras de processos biológicos no reconhecimento proteína/carboidrato, definindo interações moleculares que desencadeiam respostas particularizadas na comunicação e desenvolvimento celular, defesas imunes, infeções e inflamações e outras respostas específicas a nível tumoral, sendo muitas vezes usadas como marcadores tumorais e diferenciadores de tumores malignos e inclusive como reconhedores do nível/grau de glicosilação que está associado às metástases tumorais. *(Fu et al., 2011)*

Os últimos anos mostraram um avanço no conhecimento dos mecanismos específicos ligados às vias de elevada complexidade, pelo que os estudos mostram que as lectinas, como proteínas com propriedades antitumorais, têm atividade específica de induzir a morte celular programada como supressão de células malignas – apoptose (*self-killing*) e a autofagia (*self-eating*). *(Oliveira et al., 2011)* Estes são dois diferentes tipos de mecanismos moleculares de morte celular programada, induzidos pela própria célula como resposta a condições de stress celular, como a limitação de nutrientes, danos genéticos, excesso de componentes celulares, por exemplo.

Estudos recentes demonstram que pode existir uma co-regulação entre a apoptose e a autofagia como indutores da morte celular, estando interconectados e até regulados pelo mesmo “trigger” em simultâneo. *(Hanningan et al., 2009)* Mas como atuam? Existem três processos moleculares que podem ser a resposta a esta questão: (1) por inativação do ribossoma da célula tumoral; (2) por endocitose localizada em certos organelos da célula cancerígena, mitocôndria por exemplo; (3) por ligação a açucares conjugados na superfície da célula tumoral. Baseado nestas três possíveis vias, as lectinas de plantas podem induzir a morte a diversos tipos de células tumorais. *(Fu et al., 2011)*

Alguns destes fenómenos já foram descritos em certos estudos focados na morte induzida por lectinas. Baseado nesses estudos mostrou-se que estas proteínas em específico atuam via mitocondrial ao nível dos processos mediados e que originam a morte celular programada por estímulo, criando um balanço entre a via apoptótica e autofágica. Esta última é regulada por fatores e componentes comuns que se sobrepõem sobre as funções biológicas funcionais da célula, e que por sua vez vai desencadear e regular a atividade apoptótica. *(Liu et al., 2010)*

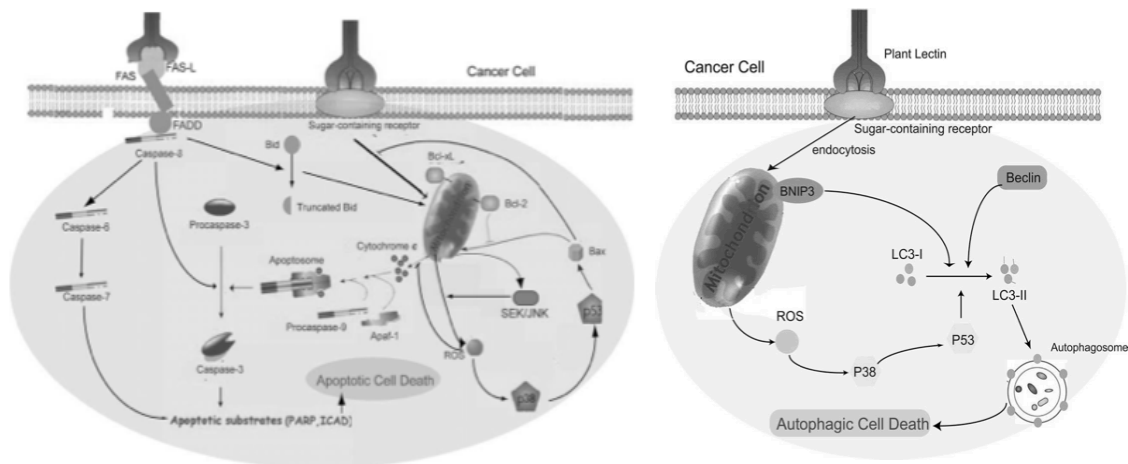


Figura 1.3: (a) Representação esquemática da base molecular da apoptose induzida por uma lectina de planta mediada através das mitocôndrias e ou recetores em células tumorais; (b) Representação esquemática da autofagia induzida por uma lectina de planta mediada através da mitocôndria (*adaptado de Liu et al., 2010*)

Estas possibilidades permitem o combate a vários tipos de tumores pelo ataque alvo às vias moleculares da morte celular programada, esquematizada na imagem da figura 1.3, potenciando novos agentes antitumorais na descoberta de novos fármacos. Estas moléculas apresentam um grau de toxicidade natural pelo que o futuro passa por criar sistemas recombinantes de proteínas modificadas com propriedades antineoplásicas aprimoradas, melhorando a eficiência e o rendimento dos processos minimizando os níveis de toxicidade para a terapêutica em causa.

Para além do seu potencial antitumoral as lectinas mostram ainda um importante papel na atividade antiviral e antimicrobiana: foi descrito que certas lectinas de plantas (tipo 2 RIP) apresentam atividade antiviral inibindo infeções, a replicação e a sua propagação sistémica, enquanto que outras podem ter um papel antiviral indireto. As lectinas podem também ter um papel de defesa contra bactérias por mecanismos indiretos como a interação com carboidratos da parede celular ou glicanos extracelulares. Podem ainda ter um papel antifúngico em que estudos já demonstraram resultados inibitórios de germinação de esporos e o crescimento de hifas. (*Peumans and Van Damme et al., 1995*)

A frutalina, como uma lectina de origem de plantas, apresenta propriedades notáveis imunomoduladoras, capacidade como biomarcador tumoral e atividade antitumoral por indução das vias da morte celular programada por apoptose ou inibição da proliferação de células

cancerígenas. Apresenta ainda uma forte atividade aglutinante (três vezes superior à jacalina) de eritrócitos de origem humana e de coelhos. *(Oliveira et al., 2014)*

Resumidamente, o campo científico tem vindo ao longo das últimas décadas a descrever as lectinas marcando uma história cronológica que dá a conhecer uma série de eventos promissores na investigação e que deixa um futuro de perspectivas que ainda não permanecem claras. Estes estudos mostram um progresso significativo e crucial na compreensão do papel que estas proteínas podem ter na sociedade científica ao nível de processos biológicos, aplicações biomédicas e como ferramenta da biotecnologia, por exemplo.

Durante muitas décadas as lectinas eram consideradas substâncias tóxicas para as células e animais, no entanto estudos demonstraram que este grupo de proteínas tem propriedades de interesse. Caracterizam-se pela sua atividade hemaglutinante, capacidade de estimulação linfócita, propriedades imuno-histoquímicas e imunomoduladoras como biomarcador no reconhecimento tumoral e atividade antitumoral. Estudos abrem novas perspectivas para as lectinas, sobretudo de origem em plantas, como potências de diagnósticos, terapêuticas futuras e de drogas antineoplásicas. *(Mislovicova et al., 2009)*

Pelas suas propriedades físico-químicas e biomoleculares as lectinas tem diversas aplicações com emergente importância na área biomédica. Estudos clínicos mostram as vantagens significativas naqueles que são os mecanismos antitumorais. À medida que as complexidades bioquímicas e moleculares das vias apoptóticas e autofágicas começam a ser melhor descritas e compreendidas, novas estratégias terapêuticas são desenvolvidas com potencial utilização clínica.

1.5. Frutalina, lectina de plantas

A frutalina é uma lectina de origem em plantas e da família das jacalinas. Isolada das sementes de fruta-pão, *Artocarpus incisa* ou *Artocarpus altilis*, é uma glicoproteína homotetramérica e que reconhece especificamente resíduos de α -D-galactose. (Oliveira et al., 2009) O nome “Frutalina” (a seguir o designado FTL) deriva do português “fruta” que dá nome à planta fonte de origem “fruta-pão” seguido do sufixo “lina”. (Oliveira et al., 2014)

Pertence à família de lectinas relacionadas com a jacalina (JRL) e à subfamília das lectinas específicas de galactose (gal-JRL). Mostra uma evolução estrutural semelhante pelo que apresenta homologas características que se refletem nas suas propriedades e atividade proteica, como a especificidade por açúcares e a identidade sequencial com a jacalina. A tradução primária do produto gal-JRL mRNA em proteína inclui modificações *co* e *pós*-tradução que resultam na remoção do péptido sinal e do propéptido *N*-terminal, numa glicosilação parcial e na excisão do *linker* tetra peptídico, separando as duas cadeias polipeptídicas (α e β), seguido do *folding* proteico. Diferentes isoformas da frutalina ou isolectinas, como diferentes sequências maduras podem apresentar diferentes atividades biológicas. (Oliveira et al., 2014)

O isolamento de frutalina das suas fontes naturais apresenta desvantagens, resulta numa mistura heterogénea de diferentes isoformas proteicas. Uma forma de ultrapassar este tipo de limitações, como fonte de proteína mais homogénea, passa por processos de biologia molecular de expressão recombinante e respetiva purificação. Com processos otimizados é possível obter elevados níveis de pureza e controlo nas propriedades finais. (Oliveira et al., 2009^a)

Vários estudos mostraram com sucesso que os sistemas heterólogos de expressão de frutalina recombinante podem ser vantajosos com uma forma de produção e purificação, fácil e viável, de frutalina funcional. Demonstraram que formas recombinantes expressas em sistemas diferentes (por exemplo *P. pastoris* e *E. coli*) têm atividade específica à imagem da forma nativa, como atividade hemaglutinante indicando o seu potencial interesse para o ramo da biomédica. (Oliveira et al., 2008)

Análises sequenciais de ácidos nucleicos mostram que a frutalina contém 471 pares de bases e que podem ocorrer diferentes sintetizações de sequências de DNA, o que indica que esta proteína em particular pode ser codificada por uma ampla família de genes. Este fenómeno não é de todo desconhecido pelo que existem estudos que revelaram que a frutalina pode ser codificada em múltiplo-*loci* em genoma *A. Integrifolia*. (Oliveira et al., 2009^b)

É uma proteína que apresenta uma conformidade muito peculiar, pois é uma proteína de ligação a carboidratos e de estrutura homotetramérica, sendo o seu modelo conformacional proveniente de estudos teoricamente computacionais.

Existem alguns estudos que caracterizam a frutalina, nomeadamente a sua forma recombinante, expressa em diferentes microrganismos como o caso de *E. coli*. A sua sequência correspondem a um total de 157 aminoácidos que equivalem a um peso molecular de 17,1 kDa e um respetivo ponto isoelétrico de 8,75. A sequência peptídica da frutalina é representada em baixo, e comparada com a sequência correspondente da jacalina. A análise mostra uma similaridade de cerca de 95% na dedução das sequências de aminoácidos entre a frutalina e a jacalina. Os diferentes aminoácidos entre a cadeia de frutalina e a cadeia de jacalina estão representados a vermelho. (*Oliveira et al., 2009b*)

Sequência da frutalina:

NEQSGK**S**QTVIVGPWGAKV**STSSN**GKAFDDGAFTGIREINLSYNKETAIGDFQVIYDLNG**R**PFVQ**NH**TS**FIK**GFTPV
KISLDFPSEYI**V**EVSGHTG**K**VSGYVWRSLTFKTNK**K**TYGPY**G**VTSGTPFNLPIENGL**V**VGFKGSIGYW**M**DYFSMYLSL
(157 aminoácidos)

Sequência da jacalina:

NEQSG**I**SQTVIVGPWGAKV**STSSN**GKAFDDGAFTGIREINLSYNKETAIGDFQ**V**YDLNG**S**PFVQ**NH**K**S**FI**T**GFTPV
KISLDFPSEYI**M**EVSGYT**G**NVSGYVWRSLTFKTNK**K**TYGPY**G**ITSGTPFNLPIENGL**I**VGFKGSIGYW**L**DYFSMYLSL
(157 aminoácidos)

Do total da sequência, 20 aminoácidos correspondem à cadeia β e 133 aminoácidos à cadeia α (de *N*-terminal para *C*-terminal) e que se encontram ligados entre si por um “*linker*” tetra peptídico de (T-S-S-N). (*Oliveira et al., 2009*) No entanto, é de referir que a frutalina nativa e a frutalina recombinante apresentam diferentes propriedades moleculares e consequente atividade biológica. A grande afinidade da frutalina por resíduos de galactose depende da correta excisão do *linker* tetra peptídico entre as cadeias β e α . (*Oliveira et al., 2011*)

1.6. *Escherichia coli*, sistema de expressão de proteínas

Os microrganismos têm vindo a mostrar que cada vez mais são peças importantes na comunidade nas mais diversas indústrias, como a indústria alimentar, farmacêutica, biotecnológica, biomédica. O seu estudo permite obter novos conhecimentos e perceber novas aplicações que possam ter um impacto vantajoso na nossa sociedade.

O campo de produção de proteínas recombinantes é muito vasto pelo que continuamos a assistir a uma expansão e um progresso impressionante na área da tecnologia de proteínas recombinantes. Nas passadas décadas presenciámos uma evolução na área terapêutica ligada à atividade proteica com uma quantidade vasta de novas aplicações clínicas. (Chen *et al.*, 2011)

De entre muitos sistemas de produção e expressão heteróloga de proteínas, o mais utilizado, que continua a dominar a área científica no que toca aos sistemas de expressão bacterianos, é a bactéria Gram-negativa, *E. coli* sendo um dos sistemas mais antigos e que continua a ser pelo conhecimento genético envolvente, pela sua capacidade de crescimento rápido, alta densidade, elevada expressão, cultura fácil ou até pela inexistência de substratos inespecíficos, a escolha mais atrativa para investigação laboratorial. (François *et al.*, 1999)

Contudo, apresenta certas limitações, a elevada densidade celular pode ser prejudicial pela toxicidade por formação de acetatos. Este fenómeno pode ser ultrapassado pelo ajuste dos níveis de oxigénio e glucose, controlando desta forma a taxa específica de crescimento nas culturas. Por outro lado, as proteínas são muito vulgarmente expressas em corpos de inclusão o que afeta a sua solubilidade e *folding* tornando-a praticamente inativa. Este último pode ainda ser afetado pela existência de pontes dissulfureto. (Demain *et al.*, 2009)

O sistema de produção em *E. coli* torna-se excelente para a expressão de proteínas não glicosilada que é a razão pelo que muitos anticorpos produzidos não reconheçam proteínas de origem mamífera. A sua genética é agora melhor compreendida do que qualquer outro microrganismo. Recentes progressos na compreensão fundamental da transcrição, tradução, e *folding* de proteínas em *E. coli*, juntamente com a disponibilidade de ferramentas genéticas melhoradas, tornam esta bactéria num sistema valioso de expressão de proteica. (Demain *et al.*, 2009)

Tabela 1.1: Características do sistema de expressão em *E.coli*. (adaptado de Demain et al., 2009)

Vantagens	Desvantagens
Expressão rápida	Proteínas com ligações dissulfido dificultam a expressão
Rendimento alto	Produção de proteínas não glicosiladas
Culturas e modificações genômicas mais fáceis	Proteínas produzidas com endotoxinas
Menor custo	Formação de acetatos resultantes da toxicidade da célula
Produção em massa mais fácil com custo efetivo	Proteínas produzidas como os corpos de inclusão são inativos e requerem o <i>refolding</i>

Este sistema tem conseguido aliar tecnologias de biologia molecular com as propriedades e características específicas da estirpe. Um número central de elementos são essenciais no design do sistema de expressão, a combinação de tecnologias de DNA recombinante é um ponto forte e que assenta na expressão do cDNA plasmídeos (figura 1.4), onde a proteína é amplificada pelo sistema escolhido. A expressão genética por plasmídeos inclui uma origem de replicação, uma marca de resistência a antibiótico, um promotor e terminador de transcrição, uma região iniciadora da tradução e um respetivo terminador. (Sorensen and Mortensen et al., 2005^a)

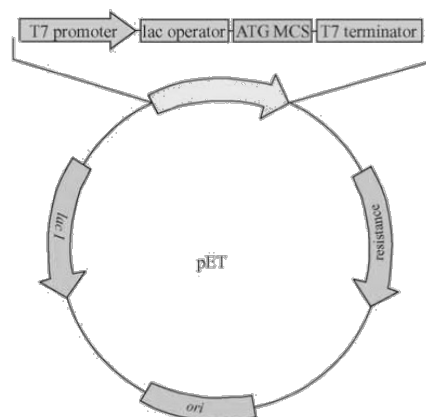


Figura 1.4: Representação esquemática de um vetor plasmídeo (adaptado de Sorensen and Mortensen et al., 2005a)

Na expressão pode ocorrer, quando sob certas condições não seletivas, uma multiplicação aleatória de plasmídeos durante o processo de replicação e na qual verificamos uma perda alvo como resultado do fenômeno de multimerização. Uma forma de contornar este problema é criar uma pressão seletiva, aproveitando as vantagens de genes plasmídeos como marcas de resistência a antibióticos, suplementar o meio de crescimento com antibiótico correspondente e assim eliminar outros plasmídeos livres nas células. (François et al., 1999)

As marcas de resistência a antibióticos como a ampicilina, kanamicina e cloranfenicol são das mais comuns na maioria dos sistemas de expressão recombinante. Esta pressão seletiva resulta na degradação e inativação das moléculas por enzimas periplasmáticas permitindo o crescimento no meio da biomassa de produção. Por exemplo, a resistência à ampicilina é mediada pela expressão de β -lactamase pelo gene *bla* e secretado no periplasma ao hidrolisar degrada o anel β -lactâmico degradando a ampicilina. (Sorensen and Mortensen et al., 2005^a)

Um sistema de expressão de proteínas recombinantes em *E. coli* requer uma combinação do vetor plasmídeo na estirpe de expressão. Neste caso o foco é a expressão citoplasmática de frutalina solúvel, em que os mais indesejáveis obstáculos são a formação de corpos de inclusão e o conseqüente *refolding* proteico. (Sorensen and Mortensen et al., 2005^b)

A purificação de proteínas solúveis a funcionais é um marco dos últimos tempos das ciências biotecnológicas e a otimização dos procedimentos pode ser a solução para a integridade proteica. Estirpes mutantes de *E. coli* contribuem na solubilidade de expressão de proteínas, a cisteínas, aminoácido com um importante papel na manutenção da estrutura conformacional das proteínas pela formação de ligações dissulfureto que aumentam a estabilidade e a resistência à proteólise, são ativamente mantidas por vias que envolvem a tioredoxina redutase e a glutaredoxina, sendo que o *fold*ing e a formação de ligações dissulfureto na proteína alvo aumentam com a fusão da tioredoxina (trx). Por outro lado existem ainda as interações com os parceiros de fusão potenciadores da solubilidade da proteína alvo e que envolvem estratégias mais complexas. (Sorensen and Mortensen et al., 2005^b)

A atividade proteica exige um rearranjo conformacional e tridimensional preciso das estruturas. Situações de condições mais adversas, como temperaturas mais elevadas, são limitativas das reações de ligação *in vivo*. Conformações intermedias tendem ainda a associar-se em grânulos amorfos de proteína aos quais denominamos de corpos de inclusão, resultado de

um equilíbrio instável entre a agregação de proteínas *in vivo* e a solubilização. (Sorensen and Mortensen et al., 2005a)

A desnaturação proteica e a sua carga metabólica podem ser controladas por fatores ambientais minimizadores deste fenómeno pelo controlo de parâmetros que permitem o aumento da solubilidade proteica *in vivo*. Dos diversos fatores podemos referenciar técnicas como a redução da temperatura da cultura, técnica em que estudos já provaram ser efetiva no aumento da solubilidade de certas proteínas. A consequência direta da redução da temperatura é a parcial eliminação de proteases induzidas pelo stress das condições de expressão. As estirpes, já referenciadas anteriormente, também têm um contributo importante na expressão de proteínas recombinantes solúveis bem como a possibilidade de interação com parceiros de fusão. Outro fator igualmente importante é o uso de *tags* de afinidade na proteína recombinante. Esta técnica tem longa tradição nos processos de purificação proteica mas não só, podem ter também influência no rendimento proteico e no aumento da solubilidade *in vivo*. (Sorensen and Mortense et al., 2005b)

Em (Oliveira et al., 2009b) comprovou-se com sucesso a expressão de frutalina recombinante em *E. coli*. Partindo da clonagem de cDNA em vector de expressão para a produção de frutalina, este estudo pretendia ainda maximizar a solubilidade da lectina. Resultados indicaram que a temperatura, o tempo de indução, a concentração de Isopropil- β -D-1-tiogalactopiranosido (IPTG) e a interação entre a concentração de IPTG e o tempo de indução tinha os efeitos mais significativos sobre o nível de expressão de frutalina solúvel recombinante.

Resumidamente, a variedade de aplicações de proteínas recombinantes e seus sistemas de expressão representam um vasto campo de estudo ainda que em certos casos com algumas limitações e necessidades o que prova uma capacidade de melhoria. O sistema de produção e expressão em *E. coli* apresenta-se como a primeira escolha laboratorial em investigação e as suas estratégias comuns de expressão resultam de uma combinação de protocolos genéticos, focados na melhoria da estabilidade estrutural e atividade proteica. O desafio é atender à procura em termos de quantidade, qualidade e custo-efetividade. (Chen et al., 2011)

1.7. Potencial biomédico de uma frutalina recombinante

Vários são os estudos de investigação que mostram com clareza a relação entre as propriedades da frutalina com o potencial interesse biomédico que advém desta proteína, especificamente para a área clínica como drogas antineoplásicas e terapêuticas anticancerígenas futuras.

Estudos indicam que esta proteína mostra uma ativa capacidade de aglutinação de eritrócitos à imagem da atividade proteica da frutalina nativa. O processo e as propriedades funcionais das formas recombinantes estão relacionados com o sistema microbiano e estratégia de produção. Ensaio revelaram uma especificidade idêntica de frutalina recombinante expressa em *E. coli* para formar ligações a carboidratos à imagem da estrutura nativa no entanto, esses mesmos resultados sugerem uma menor afinidade de ligação mas com funcionalidade preservada. (Oliveira et al., 2009b)

Outros estudos avaliaram a expressão dos glicoconjugados específicos de frutalina em tecidos do carcinoma humano da próstata e da hiperplasia benigna da próstata por imuno-histoquímica. Foi comparado o padrão de ligação de frutalina nativa com frutalina recombinante expressa em *P. pastoris* nos tecidos da próstata e avaliadas pelo potencial uso como biomarcador tumoral. Pela primeira vez, uma lectina recombinante foi utilizada e avaliada em estudos de imuno-histoquímica em tecidos da próstata humanos. Este estudo envolveu um total de 20 casos do cancro da próstata e 25 casos de hiperplasia benigna da próstata e pelos quais foi possível demonstrar que ambas as formas de frutalina estudadas têm um valor potencial a considerar na diferenciação das alterações neoplásicas de lesões benignas da próstata humana. (Oliveira et al., 2009a)

Por sua vez, a atividade citotóxica de ambas as formas (nativa e recombinante) foi provada pela sua ação em células HeLa. Ambas as frutalinas, apesar das diferenças biomoleculares, apresentam um efeito citotóxico irreversível por indução da apoptose e da inibição da proliferação celular, mostrando ainda uma afinidade de migração para este tipo específico de células cancerígenas, reforçando o seu potencial como agente anticancerígeno. A interação com as células HeLa ao nível intracelular permite por ações biomoleculares afetar o crescimento, morfologia, viabilidade de uma forma irreversível das células alvo. Foi descoberto que a frutalina nativa migra para uma região localizada perinuclear enquanto a frutalina recombinante localiza-se no próprio núcleo. (Oliveira et al., 2011)

A importância das JRL, especialmente a própria jacalina, em diagnósticos cancerígenos e terapêuticas anti tumorais tem vindo a ser uma constante presença em estudos relacionados. O potencial de frutalina recombinante expressa em *P. pastoris* foi ainda avaliado em termos de propriedades anti tumorais e como biomarcador tumoral. O estudo incidiu em tecidos da próstata e por comparação com a frutalina nativa mostrou que, esta última reconhece qualquer tipo de células dos tecidos da próstata, neoplasias (células malignas) ou hiperplasias (células não malignas) mas com ligação mais forte pelas células de carácter maligno, enquanto a forma recombinante mostrou uma atividade mais alvo e de reconhecimento específico para células neoplásicas. (*Oliveira et al., 2014*)

A frutalina é comumente caracterizada pelas suas propriedades imunomoduladoras e como marcadora tumoral responsável pelo desencadeamento de um conjunto de respostas biológicas pela ligação a receptores da superfície celular. Automaticamente é ativada uma resposta imunológica com a indução da migração de neutrófilos e reorganização do seu citoesqueleto, a ativação mitogénica de linfócitos e a indução da quimiotaxia. Tal não ocorre na presença de galactose, pelo que este resultado sugere que haja um reconhecimento específico de carboidratos. (*Brando-Lima et al., 2006*)

Desta forma, estes estudos e os resultados a eles anexados permitem elucidar sobre a linha diretiva a seguir para o futuro, ao mesmo tempo que cria expectativas sobre o potencial da frutalina, na sua forma recombinante, como uma ferramenta útil de aplicação e desenvolvimento biomédico.

1.8. Objetivos do estudo de uma frutalina recombinante

Estudos anteriores demonstraram o potencial da construção e expressão biológica de lectinas recombinantes, potencialmente ativas para efeito de aplicações biomédicas. Relativamente à frutalina a sua forma recombinante apresenta vantagens quando comparada com a forma nativa, até porque a extração de frutalina das sementes de plantas é um processo que consome tempo e custos com resultados não homogêneos mas sim uma mistura heterogênea de diferentes isoformas (isoelectinas) com atividades distintas, citotóxicas por exemplo.

Por outro lado, a sua aplicação está limitada pela capacidade de obtenção a partir das suas fontes naturais, havendo limitações a nível de demora e rendimento de extração, probabilidades de contaminação. Assim sendo, uma dos modos de se conseguir contornar estas limitações passa pela expressão recombinante de uma forma de elevado grau de pureza.

O objetivo principal deste trabalho foi a clonagem e produção de frutalina recombinante em *E. coli* contendo, em vez do *linker* nativo, um local *tobacco etch virus* (TEV) de reconhecimento para a TEV protease atuar na separação, por clivagem, das cadeias β e α e como forma de avaliar o impacto desta clivagem nas propriedades da isoforma de frutalina clonada em *E. coli*.

Um segundo objetivo foi a caracterização das versões de frutalina recombinante obtidas. Estudos anteriores que revelaram versões de frutalina produzidas em *E. coli* não apresentaram atividade citotóxica em células cancerígenas contrariamente a isoformas glicosiladas produzidas em *P. pastoris*. Desta forma levantam-se questões sobre a influência da glicosilação na atividade. Assim, a fase de caracterização torna-se influente na medida em que permitirá avaliar e relacionar a isoforma de frutalina estudada e a influência da clivagem das cadeias na atividade adquirida.

Especificando um pouco mais os principais objetivos deste trabalho, de uma forma mais prática, consistem na clonagem e construção de formas recombinante que diferem na disposição das cadeias β e α em relação à região ligante por PCR. Esta região também será sujeita a alterações, nomeadamente a substituição por uma outra região que possibilite a clivagem por uma protease (*tobacco etch 20irus* protease), processo estratégico que faz parte do estudo na medida em que a clivagem e um *refolding* posterior poderá influenciar a atividade da proteína. A alteração da disposição das cadeias está de certa forma relacionada com o resultado da clivagem. A zona de clivagem TEV é constituída por 7 aminoácidos (E-N-L-Y-F-Q-G) sendo que

a TEV protease tem ação de clivagem entre os aminoácidos Q e G (E-N-L-Y-F-Q || G) deixando seis resíduos de aminoácidos ligados a uma cadeia e o restante à outra cadeia. Tendo duas cadeias com tamanhos distintos (β com 20 a.a. e α com 133 a.a.) a razão relativamente aos resíduos deixados pela clivagem difere e que por sua vez poderá ter influência na atividade específica das suas cadeias. Esta alteração da disposição é desenhada no sentido de mitigar a influência que estes resíduos de a.a. possam ter na atividade final da frutalina recombinante.

No fim, teremos uma estratégia FTL₁ com uma construção β +TEV+ α e paralelamente uma estratégia FTL₂ com uma construção α +TEV+ β . Para cada construção, foram usados vetores de expressão distintos, selecionados pelas suas propriedades, o que possibilitam a purificação e recuperação das proteínas produzidas. Esses vetores correspondem aos plasmídeos pETM-10 e pETM-20. Como propriedades falamos nomeadamente das suas regiões constituintes. Destacando os pontos fortes de cada um dos vetores, ambos possuem um *tag* de fusão (*His-tag*) muito comumente utilizado em estratégias de purificação pela já referida técnica de IMAC. Para o pETM-20, este apresenta uma particularidade, um parceiro de fusão Trx descrito como potenciador da expressão, solubilidade e estabilidade proteica.

A produção e expressão proteica passará pela estirpe de *E. coli Rosetta*, sendo que a purificação passará por processos de IMAC e a caracterização das proteínas recombinantes e análise das suas estruturas conformacionais por técnicas de DLS e FPLC.

Capítulo 2

Materiais e Métodos

2.1. Esterilização de materiais, soluções e meios de cultura

Todos os materiais termoresistentes, soluções e meios de cultura bacterianos foram esterilizados recorrendo ao uso de autoclaves com ciclos de 20 min a 121 °C à pressão de 1 bar.

2.2. Estirpes

A tabela 2.1 apresenta a listagem de todas as estirpes usadas neste trabalho.

Tabela 2.1: Estirpes microbianas usadas durante este trabalho

Estirpe	Fonte
<i>E. coli</i> NZ5 α	NzyTech
<i>E. coli</i> Rosetta	Grupo de Trabalho

2.3. Vetores

A tabela 2.2 apresenta a listagem de todos os vetores usados neste trabalho. Estes vetores são os responsáveis pela clonagem e expressão dos genes de interesse.

Tabela 2.2: Vetores de clonagem e expressão usados durante este trabalho

Nome	Descrição	Uso	Fonte
pGEM-T Easy	Cadeias poli-T para ligação dos produtos de PCR; Marcador de 25eleção Amp; Gene LacZ para "screening" azul/branco dos clones	Ligação de produto de PCR por Taq DNA Polimerase	Promega
pETM-10	Origem de replicação f1; Marcador de 25eleção Kan; Gene promotor T7/lac; Contém insert EFH modificado de pET-24d e tag H-His/C-His	Expressão <i>E. Coli</i> dos genes	Grupo de Trabalho
pETM-20	Origem de replicação f1; Marcador de 25eleção Amp; Gene promotor T7/lac; Contém	Expressão <i>E. Coli</i> dos genes	Grupo de Trabalho

insert RAGE modificado de pET-32, *tags* N-Trx e N-His e região de clivagem TEV

2.4. Primers

O uso dos *primers* vem ao ajuste de cada uma das estratégias, pelo que na amplificação por PCR. Os *primers* desenhados e usados estão listados e descritos na tabela 2.3.

Tabela 2.3: *Primers* usados durante este trabalho. A “sublinhado” estão identificados os sítios de reconhecimento de enzimas de restrição usados nas clonagens. A “duplo sublinhado” está identificado o sítio de reconhecimento usado na clivagem por TEV protease. A vermelho o correspondente ao codão STOP

<i>Primers</i>	Sequência (5' → 3')	T _m [°C]
FTL_FW	TCTATT <u>CCATGG</u> GATCCGCGGAACAAAGCGGGAAGAGCC	70 °C
FTL_RV	AATAGACTCGAG <u>TCA</u> AAGGGACAAGTACATGCTAAAG	61,5 °C
TEV_α_FW	<u>GAGAATCTTTATTTTCAGGGC</u> GGTAAAGCTTTTGATGACG	75 °C
TEV_β_RV	<u>GCCCTGAAAATAAAGATTCTC</u> GCTTACTTTGGCTCCCAAGG	79 °C
α_FW	TCTATT <u>CCATGG</u> GATCCGGTAAAGCTTTTGATGACGG	76 °C
All_β TEVRV	AATAGACTCGAGTCAGCTTACTTTGGCTCCCAAGTCCTACTATTACAG TCTGGCTCTTCCCGCTTTGTTCCGCGCCCTGAAAATAAAGATTCTCAAGGG ACAAGTACATGCTAAAG	90 °C

2.5. Meios e condições de crescimento

As estirpes de *E. Coli*, como microrganismos bacterianos, cresceram em meio, líquido Luria-Bertani (LB) com uma rotação de 200 rpm ou em meio sólido (LB-agar) e a uma temperatura de 37°C.

Meio LB	1% (w/v) Cloreto de Sódio
	1% (w/v) Extrato de Levedura
	0.5% (w/v) Triptona
	pH a 7.5 com 10 N NaOH

Para seleção das células transformadas é suplementado o meio LB com os respetivos antibióticos (Tabela 2.4).

Tabela 2.4: Tabela resumo da suplementação do meio LB com antibióticos para seleção das culturas das estirpes de *E. coli* transformadas (Amp = Ampicilina; Kan = Kanamicina; Cloram = Cloranfenicol)

Estirpes	Vetores	Antibióticos	Concentração
<i>E. coli</i> NZ5u	pGEM-T Easy	Amp	100 µg/ml
<i>E. coli</i> Rosetta	pETM-10	Kan + Cloram	50 µg/ml + 10 µg/ml
	pETM-20	Amp + Cloram	100 µg/ml + 10 µg/ml

2.6. Conservação de culturas bacterianas

As culturas bacterianas são mantidas a 4 °C em meio apropriado e seletivo em placas de agar invertidas e seladas com parafilm. Para conservação a longo tempo foram preparados stocks permanentes, uma colónia de bactéria é crescida *overnight* em meio LB seletivo até um máximo de 18 h. Posteriormente, 1 ml dessa cultura é diluído em meio fresco numa razão de 1:10 continuando o crescimento por mais 5-6 h. Em *eppendorfs* devidamente esterilizados, adicionamos 1 ml da cultura de crescimento e 0,3 ml de glicerol 70% esterilizado, obtendo uma concentração final que ronda os 75%. É misturado, vai ao vortex e após 10 min de incubação em gelo está pronto para o armazenamento a -80 °C. Para recuperar as células armazenadas, ainda congeladas, raspar e espalhar em placas de agar com o meio apropriado à estirpe.

2.7. Estratégias de clonagem de sequências de DNA

Para a amplificação de fragmentos de DNA recorreu-se à técnica de PCR com o uso da Vent DNA Polimerase (BioLabs). Independentemente da estratégia em causa e a sequência da cadeia a amplificar foi preparada uma 28írus28 mix com uma quantidade de 100 a 300 ng DNA *template* e uma concentração final de 0,5 μM *primers* (para *forward* e *reverse*), 0,3 μM de mix de dNTP's (NZYTech), 2 mM de cofator da Taq polimerase (MgSO_4), tampão Taq 1x concentrado, 0,5-1 U de Vent – Taq Polimerase e água ultra pura (uH_2O) até perfazer o volume final de 50 μl .

2.7.1. Reação PCR de construção e amplificação da sequência β +TEV+ α

A amplificação das cadeias β +TEV, TEV+ α e β +TEV+ α seguiu a base da 28írus28 mix, carregada no termociclador e sujeita a uma série de 30 ciclos que envolveram uma desnaturação da amostra a uma temperatura de 95 °C temporizada a 2 min + 45 s, uma temperatura de *annealing* apropriada aos *primers* de 65 °C durante 45 s e uma fase de extensão de 1 min a 72 °C. A extensão final é concluída com uma fase a 72 °C durante 10 min.

Para a construção β +TEV usaram-se os *primers* FTL_FW e TEV_ β _RV, para a construção TEV+ α os *primers* TEV_ α FW e FTLRV e para a construção final de β +TEV+ α recorreu-se aos *primers* FTLFW e FTLRV.

2.7.2. Reação PCR de construção e amplificação da sequência α +TEV+ β

A ampliação da cadeia α +TEV+ β seguiu a base da 28 vírus28 mix, carregada no termociclador e sujeita a uma série de 30 ciclos que envolveram uma desnaturação da amostra a uma temperatura de 95 °C temporizada a 2 min + 45 s, uma temperatura de *annealing* apropriada aos *primers* de 55 °C durante 45 s e uma fase de extensão de 1 min a 72 °C. A extensão final é concluída com uma fase a 72 °C durante 10 min.

Para a construção da sequência α +TEV+ β usaram-se apenas dois *primers*, α _FW e All_ β TEVRV.

2.8 PCR de colónia

Esta técnica permite a verificação da correta inserção dos produtos de interesse no plasmídeo de ligação. Com a ajuda de um palito estéril, uma pequena quantidade da colónia em análise é colocada no fundo do tubo de PCR. Seguidamente as células são levadas ao microondas e aí ocorrem 2 ciclos de 45 s a 900 W. As amostras são imediatamente colocadas em gelo. É preparada uma 29irus29 mix por colónia com 2 µl de Tampão 10x, 0,6 µl de dNTP's a 10 mM, 0,8 µl de 50 mM MgCl₂, 0,4 µl de cada *primer* (*forward* e *reverse*), 0,4 de Taq DNA Polimerase e perfazer com uH₂O até um volume final de 20 µl. Durante o processo de PCR cada amostra foi sujeita a uma série de 30 ciclos que envolveram uma desnaturação da amostra a uma temperatura de 95 °C temporizada a 2 min + 45 s, uma temperatura de *annealing* apropriada aos *primers* de 55 °C durante 45 s e uma fase de extensão de 1 min a 72 °C. A extensão final é concluída com uma fase a 72 °C durante 10 min.

2.9. Purificação de DNA por kit comercial

A purificação de produtos de PCR e/ou de DNA genómico de plasmídeo foi feita com recurso a kit comercial (Qiagen) e seguindo os protocolos disponibilizados pelo fabricante. Deixando uma nota de alteração de um passo protocolar, a exceção aos protocolos ocorre no passo de eluição, que acabou por ser feita em 30 µl e não em 50 µl como é proposto nos protocolos. Desta forma é possível aumentar a concentração final de DNA.

2.10. Purificação de DNA em gel de agarose por kit comercial

A purificação de DNA genómico em gel de agarose foi conseguida com recurso ao uso de um kit comercial QIAquick Gel Extraction Kit (Qiagen) e seguindo o protocolo disponibilizado pelo fabricante, com exceção ao passo de eluição, que passou a ser feita em 30 µl e não em 50 µl como proposto no protocolo disponibilizado, como forma de aumento de rendimento da concentração final de DNA.

2.11. Extração de DNA genómico de plasmídeo por “Miniprep-Caseiro”

Esta técnica permite de uma forma mais prática e “artesanal” extrair DNA genómico das estirpes de *E. coli* repicadas para novas placas de crescimento meio seletivo *overnight*. No dia seguinte faz-se a recolha da biomassa, com ajuda de uma ansa estéril, que é ressuspensa em 200 µl de água ultra pura esterilizada com recurso ao vortex. Seguidamente adiciona-se 200 µl de Solução I e mistura-se invertendo o *ependorf* 5x. Nesta fase ocorre a lise celular. Posteriormente, adicionam-se 200 µl de Solução II voltando a misturar invertendo novamente os *ependorfs* 5x. Neste passo irá observar-se um precipitado branco e opcionalmente pode-se incubar em gelo durante 5 min de forma a favorecer esta precipitação. De seguida centrifuga-se durante 3 min à velocidade máxima e temperatura ambiente e transferir, com um certo cuidado para não levar precipitado, o sobrenadante para novos *ependorfs*. São então adicionados 500 µl de isopropanol voltando a inverter por mais 5x os *ependorfs*. Voltar a centrifugar por mais 3 min à velocidade máxima e descartar o sobrenadante. Por fim, deixar a secar durante 15 min e ressuspender em 30 µl de TE Buffer ou água ultra pura com RNase se pretendido (incubar durante 30 min a 37 °C) e guardar a -20 °C.

Solução I	1% (w/v) SDS 0,2 M NaOH
Solução II	3 M Acetato de Potássio 11,5% (w/v) Ácido Acético (pH a 5,5)

2.12. Quantificação de DNA

A concentração de ácidos nucleicos foi determinada via NanoDrop 1000 Spectrophotometer (Thermo Scientific) por análise de 1-1,5 µl de amostra. A concentração da amostra é dada em µg/µl e calculada a uma absorvância de 260 nm. É ainda possível medir a pureza da amostra pelo rácio da absorvância de 260 nm por 280 nm, em que um valor de aproximadamente 1,8 para o DNA indica uma solução de ácidos nucleicos puros.

2.13. Eletroforese de DNA em gel de agarose

A análise dos fragmentos de DNA é feita por eletroforese em gel de agarose, permitindo avaliar e estimar o tamanho das sequências ampliadas pela técnica do PCR. O gel é feito pela dissolução de uma suspensão de agarose, com uma percentagem de concentração compreendida entre 1% e 1,7% de acordo com o tamanho esperado, em solução tampão 1x TAE. Para coloração dos ácidos nucleicos foi usado Green Safe Premium.

A solidificação ocorre num molde apropriado em que com a aplicação de um pente obtém-se um gel com um alinhamento de poços que permitem carregar as amostras a analisar. O gel é colocado numa tina com tampão TAE 1x, e com a sua posição entre dois elétrodos e que por aplicação de um campo elétrico as amostras migram para o polo positivo.

O gel é visualizado e fotografado em Molecular Imager ChemiDoc™ XRS + Imaging System e analisado em software Image Lab 4.0. Para validação de resultados foi usado um marcador molecular de DNA, NZYDNA Ladder III.

Tabela 2.5: Tamanhos moleculares das bandas do marcador NZYDNA Lader III

Banda	Tamanho Molecular (bp)
1	1000
2	7500
3	6000
4	5000
5	4000
6	3000
7	2500
8	2000
9	1400
10	1000
11	800
12	600
13	400
14	200

2.14. Ligação ao vetor pGEM-T Easy

A ligação de fragmentos de DNA ao pGEM T-Easy Vetor é precedida de um passo de extensão de poli A's do produto de PCR. A otimização e quantificação do *insert* para a ligação é feita através da seguinte fórmula, para um rácio molar de 3:1.

$$\frac{ng\ vetor \cdot Kb\ insert}{Kb\ vetor} \cdot rácio\ molar = ng\ insert$$

Adicionar para uma concentração final tampão de 1x Taq DNA polimerase, 1mM MgCl₂, 0,2 mM dATP's, 0,5 U de Taq DNA polimerase e perfazer com água ultra pura até um volume de 10 µl e levar 10 min a 72 °C.

Para a ligação ao vetor (50 ng), ajustar o volume do produto PCR em função concentração, e adicionar para uma concentração final de 1x de tampão de ligação, 0,5 – 1 U de T4 DNA Ligase e por fim perfazer com água ultrapura até um volume de 10 µl. Repousar *overnight* a uma temperatura compreendida entre 4 – 10 °C.

2.15. Transformação de células competentes *E. coli*

A transformação das células competentes *E. Coli* NZY5α (NZYTech) foi feita por processos físicos, recorrendo a choques térmicos. Adicionar o volume de DNA preparado a 50 µl de células competentes e misturar cuidadosamente com auxílio da pipeta num *ependorf* e incubar em gelo durante 30 min. De seguida dar um choque térmico de 40 s a 42 °C em banho de água e incubar em gelo por mais 2 min. Adicionar 950 µl de meio SOC às células e incubar por 1 h a suspensão a 37 °C com 200 rpm.

Por último, espalhar em placas LB-agar-ampicilina com X-Gal e IPTG a diferentes diluições e crescer *overnight* a 37 °C.

SOC	2% (w/v) Triptona
	0.5% (w/v) Extrato de levedura
	10 mM NaCl

2.5 mM Kcl
10 mM MgSO₄.7H₂O
10 mM MgCl₂.6H₂O
20 mM Glucose

2.16. Digestão com enzimas de restrição *Nco*I e *Xho*I

As reações de digestão com endonucleases de restrição permitem retirar o *insert* alvo do nosso vetor secundário e transformar em vetor final de expressão. Para este trabalho como endonucleases de restrição foram selecionadas *Nco*I e *Xho*I (New England Biolabs) e usadas numa concentração final de 0,1 U num volume de reação contendo DNA vetorial purificado, água ultrapura e tampão devidamente recomendado por percentagem de atividade (NEBuffer) das duas enzimas na reação de dupla digestão. A reação ocorre a 37 °C *overnight*.

2.17. Ligação aos vetores de expressão pETM-10 e pETM-20

A ligação aos vetores finais de expressão é definida pelos locais de restrição criados pelas endonucleases. Nestes vetores a criação dos locais de ligação implica que haja uma desfosforilação, feita pela ação da *Srimp Alkaline Phosphatase* (SAP) (Fermantas). Esta desfosforilação do vetor de clonagem de DNA previne a recirculação durante ligação.

À mix da reação de digestão com as enzimas de restrição foi adicionado SAP para uma concentração final de 1 U de plasmídeo, sendo as condições da reação iguais as da digestão enzimática, 37 °C *overnight*.

2.18. Preparação e transformação de células competentes *E. coli*

A preparação de células competentes frescas de *E. Coli Rosetta* começa com a inoculação em 5 ml de meio LB seletivo, de colónias frescas crescidas em placas de agar, com crescimento *overnight* a 37 °C com uma agitação de 200 rpm. Inocular 50 ml de meio LB fresco com 0,5 ml do pré-inóculo e crescer a 37 °C, 200 rpm até uma densidade ótica (595 nm) compreendida entre 0,4 – 0,6. Em seguida, arrefecer as células em gelo por 10 min e centrifugar 40 ml a 4000 rpm a 4 °C durante 10 min. Ressuspender o *pellet* em 20 ml de

MgCl₂ frio a 0,1 M e voltar a centrifugar novamente durante 10 min a 4000 rpm a 4 °C. Por fim, ressuspender novamente as células em 2 ml de CaCl₂ frio a 0,1 M e deixar a repousar em gelo por um período de 2 h. A partir deste ponto as células podem ser usadas diretamente para transformação.

Para a transformação o processo é simples, basta apenas usar o DNA vetorial devidamente purificado e colocar em contacto com as células durante um período de tempo compreendido entre 5 – 10 min. Finalmente, plaquear o volume de transformação em placas de agar seletivas.

2.19. Crescimento e expressão proteica em “*small-scale*”

Após a seleção de colónias positivas pela técnica de PCR, segue-se a fase de avaliação do crescimento e expressão proteica em pequena escala.

O primeiro consiste na inoculação de colónias das células de expressão, no nosso caso *E. Coli Rosetta*, em 10 ml de meio LB seletivo, num falcon de 50 ml. É importante não esquecer para cada estratégia o uso de um controlo negativo. Crescer a 37 °C com uma agitação de cerca de 200 rpm durante um período de tempo aproximadamente de 3 – 4 h, que deve coincidir com uma D.O_{595nm} igual a 0,4. Colocar agora os inóculos a uma temperatura de 18 °C por um período de 30 min mantendo sempre a mesma agitação. De seguida, induzir com IPTG para uma concentração final de 0,2 mM e deixar *overnight* a 18 °C. No dia seguinte, retirar 100 µl de amostra de cada cultura, centrifugar à temperatura ambiente e a uma velocidade máxima e ressuspender o pellet em 50 ml de solução tampão TA 1x. Por fim, ferver as amostras durante 10 min a 100 °C e correr em gel SDS-PAGE.

Da restante cultura, recuperar todas as células por centrifugação e posteriormente congelar o *pellet* a -20 °C para sonicação. Possibilita, mais tarde por análise, compreender por comparação a quantidade existente de proteína na fração solúvel e insolúvel.

2.20. Produção proteica em “*scale-up*”

A produção de frutalina em grande escala começa com uma pré-cultura de 15 ml de meio LB seletivo em frascos de cultura de 50 ml, em que é inoculada biomassa celular de placas frescas de agar. Crescer *overnight* a 37 °C com uma agitação de 200 rpm. No dia seguinte, preparar frascos de cultura (matrizes de 1 L) com 250 ml de meio LB seletivo e inocular com a pré-cultura para uma diluição estima de 1:100. Incubar os frascos a 37 °C com uma agitação de 200 rpm até uma OD_{595nm} de 0,4. Baixar a temperatura da incubadora para 18 °C e manter a mesma agitação por mais um período de 30 min, período de tempo suficiente para baixar até aos 18 °C. O próximo passo passa por induzir as culturas bacterianas com IPTG, a uma concentração final de 0,2 mM. Voltar a incubar *overnight* a 18 °C a 200 rpm. No dia seguinte, recolher as células por centrifugação em tubos de centrifugação de 1L a 4500 rpm a 4 °C durante 15 min. Remover o meio sobrenadante e ressuspender o pellet em 15 ml de solução tampão PBS. Transferir para um falcon de 50 ml, centrifugar novamente a 4500 rpm a 4 °C durante 25 min. Por fim, descartar o sobrenadante e guardar o *pellet* a – 20 °C.

2.21. Lise celular por sonicação

Para este processo, em primeiro é necessário ressuspender o *pellet* previamente guardado a – 20 °C em 20 ml solução de lise celular. A lise celular é feita com recurso a um sonicador (Branson Sonifier) em 6 ciclos com um total 2,5 min com 25 s *on* e 30 s *off*. Colocar em *ependorfs* e centrifugar à rotação de 12000 rpm durante 30 min, recuperar o sobrenadante e rejeitar o *pellet*. Por fim a preparar a amostra para purificação, filtrando toda a suspensão com filtros de 0,2 µm e com recurso a uma seringa de rosca. É importante guardar aliquotas da suspensão antes e depois de centrifugar.

Tampão de Lise	50 mM Tris pH 8.0
	250 mM NaCl
	20 mM Imidazol
	5 mM MgCl ₂
	1 mM PMSF

2.22. Purificação por cromatografia de afinidade por iões metálicos

A purificação da proteína de interesse (frutalina recombinante) foi feita recorrendo à técnica de cromatografia de afinidade por iões metálicos imobilizados (*Immobilized metal-ion affinity chromatography*, IMAC) por afinidade ao Níquel. Previamente a coluna de purificação foi devidamente preparada, sendo lavada com 10 ml de uma solução de lavagem de NaOH a 1M e que ficou em contacto entre 1 a 2 h, seguido de uma passagem de 50ml de *binding buffer* (tampão de ligação) e por fim entre 5 – 10 volumes de água destilada. Pronta para uso, a coluna a calibração da coluna segue o seguinte protocolo: passar 2 CV de água destilada filtrada a uma velocidade de 1 ml/min (1x10), desta forma é removido o etanol da coluna, adicionar 3 ml de níquel a 100 mM por meio de uma seringa de rosca, passar mais 2 CV de uH₂O filtrada a uma velocidade de 1 ml/min e por fim 4 CV de *binding buffer* a 1 ml/min.

A purificação propriamente dita, começa com a passagem da amostra devidamente filtrada a uma velocidade mais baixa de 0,5 ml/min (1x6). Seguidamente fazer uma “lavagem” com 6 CV de *binding buffer* e 4 CV de *washing buffer* (tampão de lavagem) a uma velocidade de 1 ml/min. É importante para cada passagem, tanto da amostra inicial como das soluções de purificação, fazer uma recolha do *flow-through*.

A ultima etapa consiste na passagem de *elution buffer* (tampão de eluição) a uma velocidade de 0,5 ml/min e fazer a recolha do *flow through* em falcons num volume aproximado de 5 ml enquanto fazemos a monitorização da quantidade de proteína que está a sair pelo método de *Bradford*, numa placa de 96 poços por exemplo (100 µl de *Bradford* para 5 µl de amostra).

No fim, é importante voltar a lavar a coluna, começando com a passagem de 2 CV de solução EDTA a 100mM, 2 CV de dH₂O filtrada, 2 CV de uma solução de guanidina em PBS, de novo mais 2 CV de dH₂O filtrada e por fim passar 2 CV de etanol a 20 % e deixar.

<i>Binding Buffer</i>	50 mM Tris pH 8.0
	150 mM NaCl
	20 mM Imidazol

<i>Washing Buffer</i>	50 mM Tris pH 8.0
	150 mM NaCl
	40 mM Imidazol

<i>Elution Buffer</i>	50 mM Tris pH 8.0
	150 mM NaCl
	300 mM Imidazol

2.23. Diálise

O processo de diálise implica o uso de membranas adequadas ao tamanho molecular da proteína em estudo. Esta, contendo no seu interior, a fração purificada da proteína, fica emersa num recipiente contendo solução tampão PBS 1x com agitação baixa a 4 °C *overnight*. O volume de PBS usado deve ser cerca de 30x o volume de amostra a dialisar.

2.24. Clivagem por *tobacco etch virus* protease

A clivagem por TEV protease foi feita *in vitro* num falcon de 50 ml *overnight* a 4 °C e com agitação lenta. A quantidade de TEV protease usada em cada reação é definida por um fator de 1/20 dependendo da quantidade proteica presente na amostra a dialisar. Por exemplo, numa amostra de 2 ml contendo uma concentração proteica de 4,5 mg/ml, no volume total existe 9 mg de proteína pelo que, para garantir uma clivagem, teremos de ter 9 x 1/20 mg de TEV protease, ou seja 0,45 mg, no volume final da reação, que para um stock de 1 mM corresponde a 450 µl.

2.25. Análise proteica por eletroforese de géis SDS-PAGE

Para analisar a expressão e o processo de purificação proteica recorreu-se ao uso de géis SDS-PAGE, compostos por uma fase de gel “*stacking*” e “*resolving*” ambos de acrilamida a 4 % e 12 % respetivamente.

4% Gel “ <i>Stacking</i> ” de Acrilamida	3.84% (v/v) Acrilamida (Bio-Rad) 0.104% (v/v) Bis-Acrilamida (Bio-Rad) 0.12 M Tris-Hcl, pH
--	---

	6.8
	0.1% (w/v) SDS
	0.05% (w/v) APS (Merck)
	0.005% (v/v) TEMED (NzyTech)
12% Gel “ <i>Resolving</i> ” de Acrilamida	11.68% (v/v) Acrilamida
	0.32% (v/v) Bis-Acrilamida
	0.375 M Tris-Hcl, pH 8.8
	0.1% (w/v) SDS
	0.05% (w/v) APS 0.005% (v/v) TEMED

As amostras de análise SDS-PAGE são preparadas pela mistura de 16 µl das amostras da purificação mais 4 µl de 5x tampão de amostra, e são incubadas durante 5 min a 100 °C, para desnaturação. Deste volume, são carregados 15 µl de cada amostra a analisar nos poços dos géis. A corrida é feita em 500 ml de 1x tampão de corrida a 30 mA com uma duração aproximada de 1 h, para dois géis em simultâneo.

Tampão de amostra	0.15 m M Tris-Hcl, pH a 6.8
	50% (v/v) Glicerol
	10% (w/v) SDS
	25% (w/v) β-mercaptoethanol
	0.01% (w/v) Bromophenol Blue
Tampão de corrida	25 mM Tris base
	192 mM Glicina
	0.1% (w/v) SDS

Para validação de resultados foi usado um marcador molecular de proteína, Page Ruller Unstained Broad Range Protein Ladder.

Tabela 2.6: Tamanhos das bandas moleculares do marcador Page Ruller Unstained Broad Range Protein Ladder

Banda	Peso Molecular (kDa)
1	250
2	150
3	100
4	70
5	50
6	40
7	30
8	20
9	15
10	10
11	5

2.26. Revelação de géis SDS-PAGE pelo método de *Coomassie Blue*

Após término da eletroforese, a revelação de cada gel foi feita através de um método que recorre ao uso de *coomassie blue*. Cada gel é colocado num recipiente com dH₂O e aquecida durante um ciclo de 30 segundos num micro-ondas e posteriormente colocado em agitação constante durante 3 minutos. Este processo é repetido, de igual forma, por mais dois ciclos. Seguidamente, a dH₂O é substituída pela solução corante *coomassie blue*, e cada gel é igualmente aquecido durante 30 segundos num micro-ondas e posteriormente colocado em agitação constante durante cerca de 30 minutos. No final, o corante *coomassie blue* é removido e é adicionada novamente dH₂O para descoloração de cada gel.

Solução Corante <i>Coomassie Blue</i>	60-80 mg Coomassie Brilliant
	1 L dH ₂ O
	2-4 h agitação magnética
	3 ml de Hcl concentrado a 37%

2.27. Caracterização por *Dinamic Light Scattering*

A técnica de DLS é uma forma prática de caracterização e de leve caracterização estrutural de moléculas proteicas dispersas em solução. Existem outros métodos de caracterização de partículas, sendo que a técnica de DLS apresenta-se como um potencial em estudos recentes como sendo um método simples para análise e medida dos tamanhos e estabilidade das partículas em suspensão. Na prática, com recurso a uma cuvette, efetua-se a medição de cerca de 0,5 ml da suspensão amostral com uma concentração entre 0,5 a 0,75 mg/ml. Após ajuste do programa ao material correspondente da cuvette a usar e ao correspondente “branco” ou dispersante da amostral, cada amostra passa por uma medição de análise de 10 ciclos. O equipamento usado pertence à Malvern e correspondente a um modelo Nanosizer.

2.28. Caracterização por *Fast Protein Liquid Chromatography*

A cromatografia líquida é um método amplamente usado na recuperação e purificação de proteínas. Esta técnica permite analisar ou purificar proteínas ou mistura de proteínas. A separação ou a sua caracterização por cromatografia é possível pelo diferente nível de afinidade das partículas moleculares pelos materiais constituintes da coluna. Para análise, 1 ml correspondente amostra de frutalina recombinante em PBS é injetada numa coluna (Sephacryl) por um sistema de FPLC (Pharmacia Biotchnology). A calibração e a eluição da coluna foram feitas com PBS a um caudal de 1 ml/min.

Para enquadramento da análise dos resultados, foram ainda injetadas amostras padrão de BSA e de frutalina nativa.

Capítulo 3

Resultados e Discussão

O estudo e a investigação nas diferentes áreas científicas são marcos importantes na história da ciência, independentemente do resultado final poder ser o mais ou menos desejado. A verdade é que cada estudo pode fornecer e influenciar, numa óptica construtiva, novos resultados, conhecimentos e ideias a novos potenciais projetos de investigação, o que por sua vez já se torna numa mais valia e contributo para o avanço científico.

3.1. Construção de uma frutalina recombinante em *E. coli*

Este estudo pretende caracterizar uma proteína de interesse, a frutalina, partindo de ideias e estudos provenientes de investigações anteriores e que tiveram resultados e conclusões promissoras do que é expressão de frutalina recombinante em sistemas como *E. coli*.

A fase inicial deste estudo passou pela construção e clonagem, das sequências genéticas pretendidas e projetadas. Sendo a frutalina uma lectina, sequencialmente é constituída por duas cadeias, uma mais curta de 20 aminoácidos (cadeia β) e uma outra mais longa de 133 aminoácidos (cadeia α), unidas entre si por uma região ligante à qual denominamos de *linker* e que é constituído por 4 aminoácidos. Estudos corroboram a hipótese de que a atividade da frutalina está muitas vezes associada à correta excisão do *linker* tetra peptídico ligante entre as cadeias β e α . (Oliveira *et al.*, 2008) No entanto, outros estudos indicam que quando se trata de uma frutalina recombinante e expressa em sistemas vivos, *P. pastoris*, tal processo de clivagem não ocorre naturalmente, resultando numa forma de frutalina com baixa afinidade de ligação a carboidratos. (Oliveira *et al.*, 2011). Desta forma, um dos grandes objetivos passa pela alteração genética desta região ligante por uma outra que de certa forma seja de simples processamento de clivagem, *in vitro*, envolvendo menos custos e com recursos mais acessíveis. Assim sendo, o primeiro passo consistiu na alteração do referido *linker* nativo da frutalina por uma outra região ligante, *tobacco etch 43irus* (TEV), de reconhecimento para atuação de uma protease, TEV protease neste caso.

É importante mostrar e descrever a base da sequência nucleotídica por detrás das construções e aquela que se apresenta como frutalina *template* de processamento e que serviu de molde estrutural das versões recombinantes em estudo.

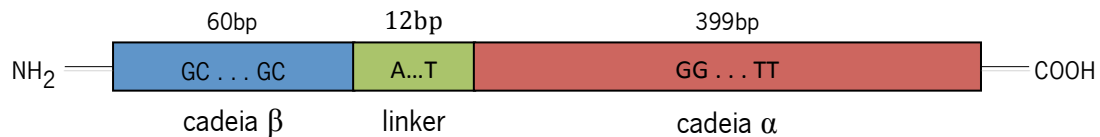


Figura 3.1: Representação esquemática e estrutural da frutalina *template*

A imagem da figura 3.1 representa a estrutura da frutalina *template* e onde estão representadas as duas cadeias constituintes, β e α , ligadas entre si por uma região ligante à qual designamos de *linker*. Como complemento apresenta-se a sequência nucleica correspondente:

Sequência Frutalina *Template* (471 nucleótidos)

```

1 GCGGAACAAA GCGGGAAGAG CCAGACTGTA ATAGTAGGAC CTTGGGGAGC CAAAGTAAGC
61 ACTAGCTCCA ATGGTAAAGC TTTTGATGAC GGTGCATTCA CCGGAATCAG AGAAATCAAC
121 CTTTCATATA ATAAGGAGAC CGCCATTGGG GACTTCCAAG TTATTTACGA CTTGAATGGA
181 AGGCCATTTG TTGACAAAG TCATACTAGT TTTATAAAAG GCTTCACACC AGTGAAGATT
241 TCTCTAGACT TTCCAAGCGA GTATATAGTA GAAGTGAGCG GACACACTGG TAAAGTGAGT
301 GGGTATGTAG TAGTACGCTC TTTGACATTC AAGACTAATA AAAAAACCTA TGGACCATAT
361 GGAGTTACAA GCGGCACACC TTTCAACCTC CCAATCGAGA ATGGCTTAGT TGTGGATTTC
421 AAAGGAAGTA TCGGCTACTG GATGGACTAC TTTAGCATGT ACTTGTCCCT T
    
```

A sequência de frutalina que serviu de base e molde a este estudo tem por fonte uma construção já descrita e expressa em *E. coli*, denominada de pETMFh8FTL, correspondente ao estudo e descrição da ação de parceiros de fusão na solubilidade proteica. (*Costa et al., 2013*)

Outro dos objetivos deste estudo passou pela caracterização das formas recombinantes da frutalina com alterações conformacionais, ou seja, alterações ao nível da disposição das cadeias β e α relativamente à região ligante. Os produtos finais são representados de seguida, nos quais a presença de zonas de corte *NcoI/BamHI* e *XhoI* resultam da utilização dos *primers* desenhados especificamente para as estratégias propostas. A utilização dos *primers* torna-se fundamental, possibilitando assim rigor e controlo de qualidade das regiões específicas de interesse e a delimitar.

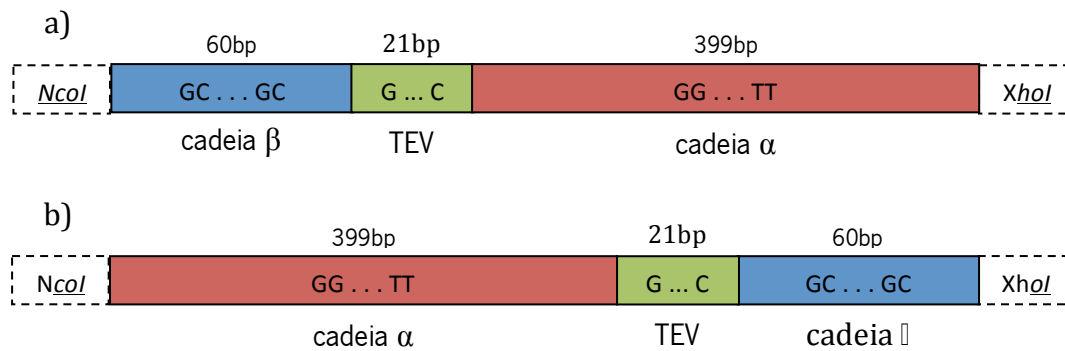


Figura 3.2: Representação esquemática e estrutural das formas recombinantes finais de frutalina. a) Forma recombinante final para a estratégia "1" com alteração da região ligante; b) Forma recombinante final para a estratégia "2" com alteração da região ligante e da disposição das cadeias β e α relativamente à região TEV

Para que, durante a análise, a identificação correspondente de cada forma recombinante de frutalina, seja mais clara podemos fazer corresponder a cada estratégia uma designação. A partir deste momento podemos associar, para as imagens da figura 3.2, uma estratégia. Faremos corresponder à 'estratégia 1' a construção de frutalina com a disposição na forma $\beta + TEV + \alpha$ e à qual podemos designar de FTL_1 enquanto que, a 'estratégia 2' corresponderá à construção de frutalina com a disposição na forma $\alpha + TEV + \beta$ e à qual podemos designar por FTL_2 . É de salientar que ao longo da análise serão ser feitas as referências descritivas de aspetos relevantes e suscetíveis de caracterização.

3.1.1. Construção por PCR de FTL_1 , sequência $\beta + TEV + \alpha$

A primeira das estratégias consistiu na construção e amplificação da sequência da frutalina pela ordem estrutural $\beta + linker + \alpha$, lembrando ainda a referência à alteração da região ligante por uma outra zona de clivagem TEV. Estas modificações estruturais levam a uma conformação do produto final na forma $\beta + TEV + \alpha$. Esta construção implicou subprocessos, ou seja, uma fase intermédia de construção para a sequência $\beta + TEV$ e outra fase de construção da para a sequência $TEV + \alpha$. Posteriormente, por um terceiro processo de reação em cadeia da polimerase e de ligação obtém-se como produto final $\beta + TEV + \alpha$, correspondente à designação de FTL_1

Para cada processo de reação por polimerase existe um fator comum e que corresponde aos já referidos *primers*. Para a construção e amplificação da subcadeia β +TEV recorreu-se a dois *primers* distintos, FTL_FW e β _TEV_RV e que têm como sequência alvo:

Sequência alvo dos *primers*: sequência β +TEV (81 nucleótidos)

1 GCGGAACAAA GCGGGAAGAG CCAGACTGTA ATAGTAGGAC CTTGGGGAGC CAAAGTAAGC
61 ACTAGCTCCA AT

Como resultado final é esperado uma sequência com um tamanho molecular de cerca de 92 nucleótidos. Este tamanho molecular corresponde aos 81 bp da sequência β mais região TEV juntamente com a contabilização da introdução na construção de uma zona de restrição de endonucleases correspondente aos 11 bp da sequência *NcoI/BamHI* anexada ao *primer* FTL_FW em *N*-terminal.

Produto final esperado: sequência *NcoI/BamHI*+ β +TEV (92 nucleótidos)

1 CCATGGGATC CGCGGAACAA AGCGGGAAGA GCCAGACTGT AATAGTAGGA CCTTGGGGAG
61 CCAAAGTAAG CGAGAATCTT TATTTTCAGG GC

Por análise à imagem da figura 3.3 comprova-se o resultado esperado para a construção da sequência β +TEV, com recurso aos *primers* FTL_FW e β _TEV_RV, com um tamanho esperado de 92bp.

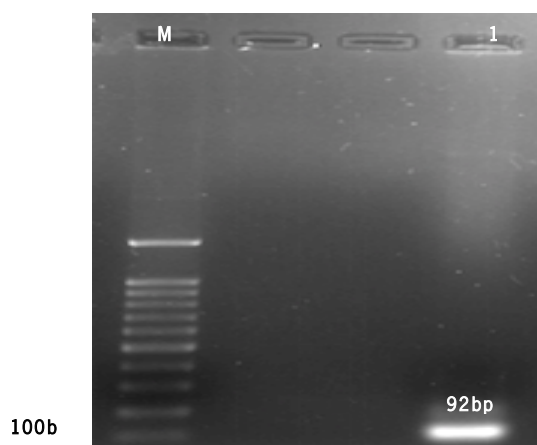


Figura 3.3: Construção e amplificação do produto final β +TEV. (M) Marcador NZY Ladder; (1) Banda correspondente à sequência β +TEV (tamanho molecular esperado de 92 bp)

Pelo mesmo processo mas com recursos diferentes ocorreu a construção e amplificação da sub sequência TEV+ α . Para esta construção foram usados outros dois *primers* distintos dos anteriores, α _FW e FTL_RV, tendo como sequência alvo.

Sequência alvo dos *primers*: sequência TEV+ α (411 nucleótidos)

```

1 ACTAGCTCCA ATGGTAAAGC TTTTGATGAC GGTGCATTCA CCGGAATCAG AGAAATCAAC
61 CTTTCATATA ATAAGGAGAC CGCCATTGGG GACTTCCAAG TTATTTACGA CTTGAATGGA
121 AGGCCATTTG TTGGACAAAG TCATACTAGT TTTATAAAAG GCTTCACACC AGTGAAGATT
181 TCTCTAGACT TTCCAAGCGA GTATATAGTA GAAGTGAGCG GACACACTGG TAAAGTGAGT
241 GGGTATGTAG TAGTACGCTC TTTGACATTC AAGACTAATA AAAAAACCTA TGGACCATAT
301 GGAGTTACAA GCGGCACACC TTTCAACCTC CCAATCGAGA ATGGCTTAGT TGTGGATTCT
361 AAAGGAAGTA TCGGCTACTG GATGGACTAC TTTAGCATGT ACTTGTCCTT
    
```

O produto final esperado corresponderá a uma sequência com cerca de 429 nucleótidos correspondentes dos 420 bp da sequência da cadeia α com a região TEV (TEV+ α) mais a contabilização da introdução na construção das sequências correspondentes a um codão STOP (3 bp) mais uma zona de restrição *XhoI* com 6 bp, anexadas ao *primer* FTL_RV em *C-terminal*.

Produto final esperado: sequência TEV+ α +STOP/*XhoI* (429 nucleótidos)

```

1 GAGAATCTTT ATTTTCAGGG CGGTAAAGCT TTTGATGACG GTGCATTAC CCGAATCAGA
61 GAAATCAACC TTTCATATA TAAGGAGACC GCCATTGGGG ACTTCCAAGT TATTTACGAC
121 TTGAATGGAA GGCCATTTGT TGGACAAAGT CATACTAGTT TTATAAAAGG CTTACACACCA
181 GTGAAGATTT CTCTAGACTT TCCAAGCGAG TATATAGTAG AAGTGAGCGG ACACACTGGT
241 AAAGTGAGTG GGTATGTAGT AGTACGCTCT TTGACATTCA AACTAATAA AAAAACTAT
301 GGACCATATG GAGTTACAAG CGGCACACCT TTCAACCTCC CAATCGAGAA TGGCTTAGTT
361 GTTGGATTCA AAGGAAGTAT CGGCTACTGG ATGGACTACT TTAGCATGTA CTTGTCCCTT
421 TGACTCGAG
    
```

O resultado da construção da sequência final é comprovado por observação da banda com um tamanho esperado de 429 bp na imagem da figura 3.4. Este produto de PCR será usado, juntamente com o anterior, como *template* base da seguinte reação de PCR para a construção e amplificação da sequência final β +TEV+ α , de designação FTL₁.

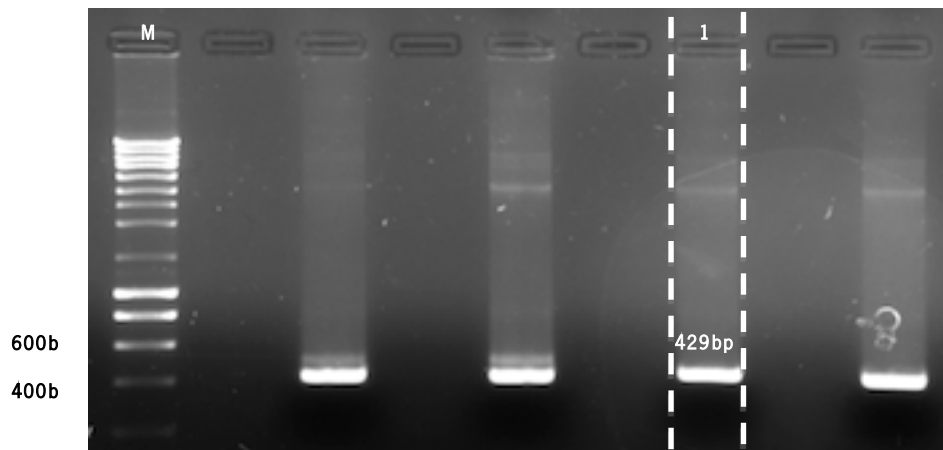


Figura 3.4: Construção e amplificação do produto final TEV+ α . (M) Marcador NZY Ladder III; (1) Banda correspondente à sequência TEV+ α (tamanho molecular esperado de 429 bp)

A construção do produto final FTL₁ correspondente à sequência β +TEV+ α que culmina com uma última reação de PCR, processo este que usa como sequências *template de* base os dois sub produtos obtidos nas construções anteriores, como já referido anteriormente e com um tamanho de banda molecular de 500 bp.

Produto final esperado: sequência *Nco*I+ β +TEV+ α +STOP/*Xho*I (500 Nucleótidos)

```

1  CCATGGGATC  CGCGGAACAA  AGCGGGAAGA  GCCAGACTGT  AATAGTAGGA  CCTTGGGGAG
61  CCAAAGTAAG  CGAGAATCTT  TATTTTCAGG  GCGGTAAGC  TTTTGATGAC  GGTGCATTCA
121 CCGGAATCAG  AGAAATCAAC  CTTTCATATA  ATAAGGAGAC  CGCCATTGGG  GACTTCCAAG
181 TTATTTACGA  CTTGAATGGA  AGGCCATTTG  TTGGACAAAG  TCATACTAGT  TTTATAAAAG
241 GCTTCACACC  AGTGAAGATT  TCTCTAGACT  TTCCAAGCGA  GTATATAGTA  GAAGTGAGCG
301 GACACACTGG  TAAAGTGAGT  GGGTATGTAG  TAGTACGCTC  TTTGACATTC  AAGACTAATA
361 AAAAAACCTA  TGGACCATAT  GGAGTTACAA  GCGGCACACC  TTTCAACCTC  CCAATCGAGA
421 ATGGCTTAGT  TGTGGATTC  AAAGGAAGTA  TCGGCTACTG  GATGGACTAC  TTTAGCATGT
481 ACTTGTCCCT TGACTCGAG
    
```

O resultado final esperado corresponderá a um produto com um tamanho de sequência com 500 nucleótidos e que representam a construção β +TEV+ α com 480 bp mais as sequências introduzidas e correspondentes às zonas de restrição das endonucleases *Nco*I/*Bam*HI com 11 bp anexada N-terminal na cadeia β e resultante do recurso ao *primer* FTL_FW mais uma zona de restrição da endonuclease *Xho*I com 6 bp anexada a C-terminal na

cadeia α juntamente com um codão STOP (3 bp) resultante do uso do *primer* FTL_RV. O resultado da construção da sequência final pode ser facilmente comprovado pela análise à imagem da figura 3.5.

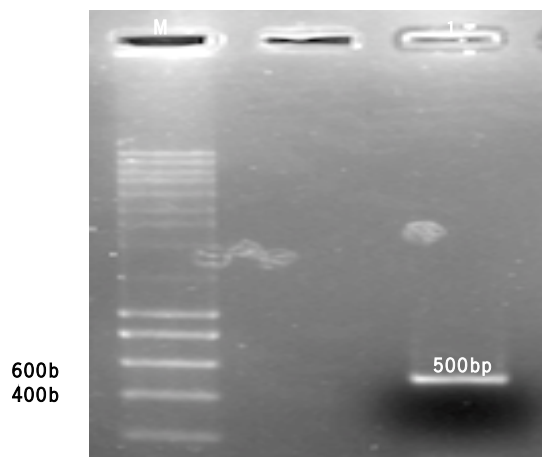


Figura 3.5: Construção e amplificação do produto final FTL₁, com sequência correspondente β +TEV+ α . (M) Marcador NZY Ladder III; (1) Banda correspondente à sequência do produto final FTL₁ (tamanho molecular esperado de 500 bp)

Esta é uma reação crítica na medida em que aconteceu entre dois produtos de PCR tem partilham entre si 21 bp. Estas bases em comum correspondem à sequência da zona de clivagem TEV pelo que no decorrer do processo poderia ocorrer duplicação dessas mesmas bases partilhadas pela ação da Vent polimerase. Comprovou-se que tal fenómeno não ocorreu e que esta reação crítica foi conseguida com puro sucesso.

3.1.2. Construção por PCR de FTL₂, sequência α +TEV+ β

A segunda estratégia consistiu na construção e amplificação de uma sequência de frutalina com alterações conformacionais, ou seja, com alterações da disposição das cadeias β e α relativamente à região ligante que de igual modo sofrerá uma alteração para uma sequência TEV. Desta forma o resultado final será uma sequência representativa de α +TEV+ β , correspondente à designação de FTL₂.

Ao contrário da estratégia anterior esta construção envolve apenas um processo de PCR com recurso à utilização de dois *primers* devidamente pensados para o propósito. A sequência

alvo da frutalina *template* pelos *primers* correspondeu apenas a aquela que é a sequência da cadeia α com os seus 399 nucleótidos constituintes.

Tem como produto final esperado uma sequência de bases com um peso molecular de 500 nucleótidos, este processo tem uma particularidade e que está relacionada com um dos *primers* envolvidos na construção. Pela sequência alvo percebemos que apenas a cadeia α é usada da frutalina *template* o que deixa questões relativamente à restante estrutura TEV+ β . Aqui entra a particularidade do *primer* All_TEV_ β _RV que tem na sua constituição a sequência para a região TEV juntamente com a cadeia β que agregados pelo processo de PCR com a cadeia α perfazem aquele que é o produto final, α +TEV+ β .

Sequência produto final *NcoI*+ α +TEV+ β +*STOP/XhoI*. (500 Nucleótidos)

1	<u>CCATGGGATC</u>	CGGTAAAGCT	TTTGATGACG	GTGCATTAC	CGGAATCAGA	GAAATCAACC
61	TTTCATATAA	TAAGGAGACC	GCCATTGGGG	ACTTCCAAGT	TATTTACGAC	TTGAATGGAA
121	GGCCATTTGT	TGGACAAAGT	CATACTAGTT	TTATAAAAGG	CTTCACACCA	GTGAAGATT
181	CTCTAGACTT	TCCAAGCGAG	TATATAGTAG	AAGTGAGCGG	ACACACTGGT	AAAGTGAGTG
241	GGTATGTAGT	AGTACGCTCT	TTGACATTCA	AGACTAATAA	AAAAACCTAT	GGACCATATG
301	GAGTTACAAG	CGGCACACCT	TTCAACCTCC	CAATCGAGAA	TGGCTTAGTT	GTTGGATTCA
361	AAGGAAGTAT	CGGCTACTGG	ATGGACTACT	TTAGCATGTA	CTTGTCCTT	<u>GAGAATCTTT</u>
421	<u>ATTTTCAGGG</u>	<u>CGCGGAACAA</u>	AGCGGGAAGA	GCCAGACTGT	AATAGTAGGA	CCTTGGGGAG
481	CCAAAGTAAG	<u>CTGACTCGAG</u>				

As 500 bases representam a construção α +TEV+ β com 480 bp mais as sequências introduzidas e correspondentes às zonas de restrição das endonucleases *NcoI/BamHI* com 11 bp anexada *N*-terminal na cadeia β e resultante do recurso ao *primer* FTL_FW mais uma zona de restrição da endonuclease *XhoI* com 6 bp anexada a *C*-terminal na cadeia α juntamente com um codão STOP (3 bp) resultante do uso do *primer* FTL_RV. O resultado da construção da sequência final pode ser comprovado pela análise à imagem da figura 3.6.

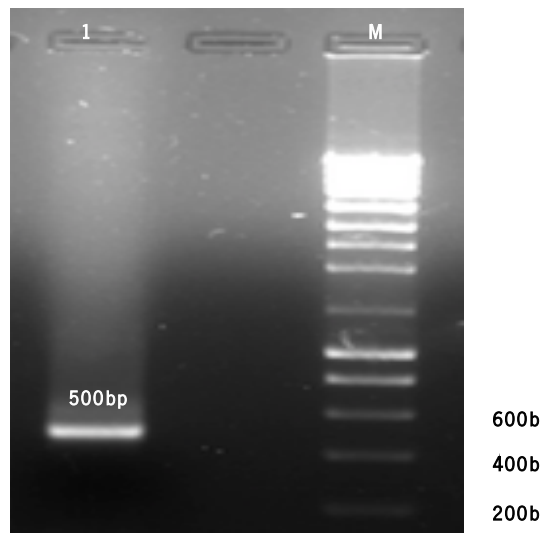


Figura 3.6: Construção e amplificação do produto final FTL_2 , com sequência correspondente $\alpha+TEV+\beta$. (M) Marcador NZY Ladder III; (1) Banda correspondente à sequência do produto final FTL_2 (tamanho molecular esperado de 500 bp)

3.2. Clonagem em vetor intermédio pGEM-T Easy e transformação em células competentes *E. coli*

Purificadas as sequências de interesse, a linha a seguir correspondeu à clonagem dos produtos finais confirmados de PCR, para cada uma das estratégias, em vetor plasmídeo intermédio. Esta ligação permite transpor os produtos para um sistema viável de replicação. A ligação do produto de ambas as estratégias ao vetor pGEM-T Easy implica um passo adicional e que consiste em expor a sequência de ligação a 72 °C durante 10min com adição de Taq DNA Polimerase que atua como um extensor das chamadas cadeias poli-A às extremidades sequenciais que funcionaram como regiões de ligação na reação com o vetor. A ligação ocorre à *posteriori* com recurso ao uso de T4 DNA Ligase.

Culturas de meio LB, suplementado com ampicilina a 100 $\mu\text{g}/\text{ml}$, foram usados para seleção de *E. coli* recombinante com as sequências de interesse. A transformação permitiu ser criado um sistema viável de replicação com marcadores genéticos de seleção que conferem resistência antibiótica à *E. coli* hospedeira, garantindo uma seleção na cultura e que permite o isolamento, a recolha e purificação do plasmídeo de interesse.

O rastreio de clones positivos dos múltiplos clones contendo a sequência das formas recombinantes da frutalina foi avaliado com recurso a PCR de colónias numa primeira fase, como mostra a figura 3.7.

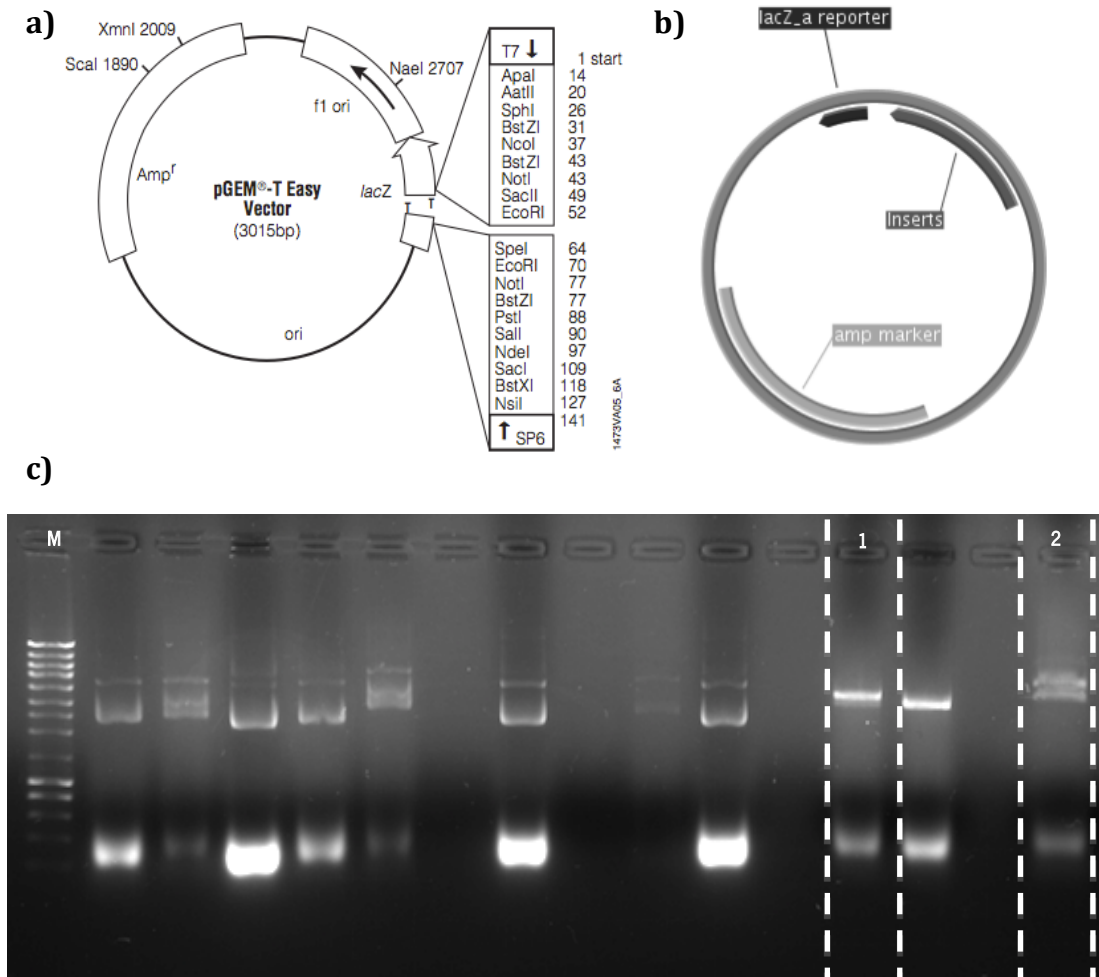


Figura 3.7: Rastreamento de clones positivos da transformação em vetor pGEM-T Easy dos genes de ambas as seqüências. (a) Representação esquemática da constituição do vetor pGEM-T Easy; (b) Representação esquemática da construção de ambas as estratégias em vetor pGEM-T Easy; (c) Despistagem das construções em vetor pGEM-T Easy para ambas as estratégias: (M) Marcador NZY Ladder III; (1) Clone selecionado para construção FTL₁ (β +TEV+ α) em pGEM-T Easy; (2) Clone selecionado para construção FTL₂ (α +TEV+ β) em pGEM-T Easy

Para uma confirmação mais precisa da correcta transformação das construções para cada uma das estratégias recorreu-se ao processo de digestão enzimática dos clones identificados e seleccionados, referenciados na figura 3.7 em (1) e (2) para FTL_1 e FTL_2 respectivamente. Desses clones as respectivas sequências são agora extraídas do vetor pGEM-T Easy, após purificação do material genético, por ação das endonucleases de restrição *NcoI* e *XhoI* como enzimas de ação digestiva sobre os pontos de restrição usados em cada construção. A confirmação da extração dos *inserts* β +TEV+ α e α +TEV+ β pode ser observada por análise à imagem da figura 3.8.

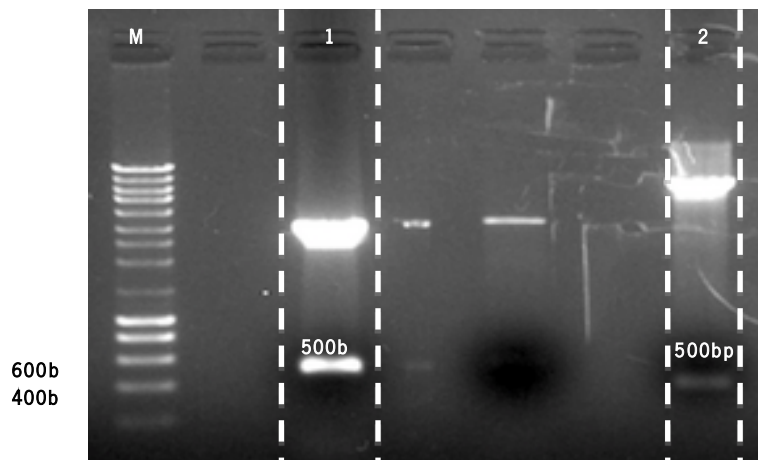


Figura 3.8: Extração dos inserts do vetor pGEM-T Easy por digestão enzimática *NcoI* e *XhoI*. (M) Marcador NZY Ladder III; (1) Confirmação por extração da sequência da estratégia FTL_1 (β +TEV+ α com tamanho molecular esperado de 500 bp); (2) Confirmação da extração da sequência da estratégia FTL_2 (α +TEV+ β com tamanho molecular esperado de 500 bp)

A imagem da Figura 3.8 permite comprovar o sucesso da primeira fase de clonagem e transformação dos produtos de PCR de interesse em vetor pGEM-T Easy. Ambas as bandas foram devidamente purificadas e clonadas em dois diferentes vetores de expressão.

3.3. Clonagem em vetores de expressão pETM-10 e pETM-20 e transformação em células competentes *E. coli Rosetta*

O foco passa agora pela clonagem das sequências de interesse FTL_1 e FTL_2 , devidamente isoladas e purificadas em vetor de expressão proteica. Para a clonagem foram

adotados dois vetores de expressão, pETM-10 e pETM-20. Nesta fase temos duas estratégias para dois vetores o que perfaz um total quatro tipos de expressão de frutalina recombinante. Para a clonagem e ligação dos *inserts* alvo, os plasmídeos pETM-10 e pETM-20 foram também eles sujeitos à ação das endonucleases para as sequências de restrição projetadas nas construções e na posterior excisão dos produtos clonados em pGEM-T Easy, *NcoI* e *XhoI*.

A constituição e descrição de ambos os vetores está representada na imagem da figura 3.9 e permite perceber alguns dos aspectos relevantes que levaram à sua escolha.

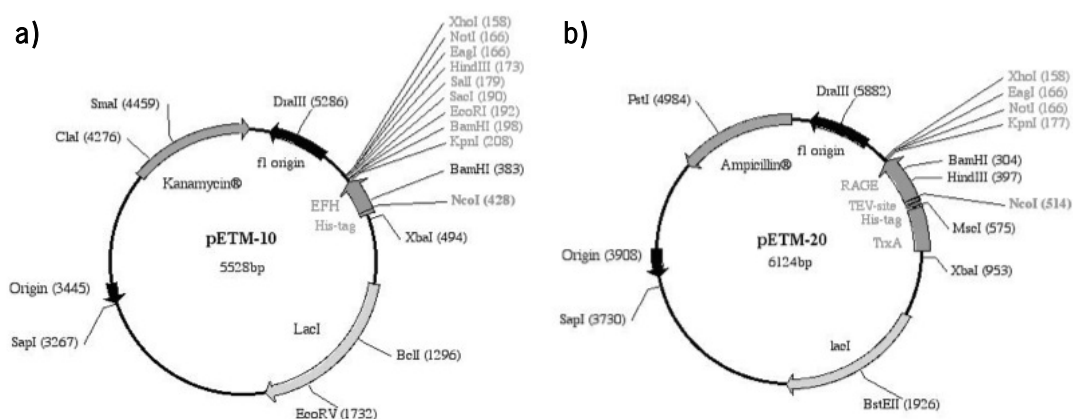


Figura 3.9: Representação esquemática da constituição dos vetores de expressão selecionados, pETM-10 e pETM-20. (a) Descrição pETM-10: origem de replicação f1, marcador de selecção Kan, gene promotor T7/lac, *insert* EFH modificado de pET-24d e *tag* N-His/; (b) Descrição pETM-20: origem de replicação f1, marcador de selecção Amp, gene promotor T7/lac, *insert* RAGE modificado de pET-32, *tags* N-Trx e N-His e região de clivagem TEV

Os vectores pETM são idênticos diferenciando apenas nos parceiros de fusão na posição N-terminal. A posição N-terminal dos parceiros de fusão mostra ser ótima opção para a expressão proteica quando comparada com a posição C-terminal. É descrito que alguns parceiros de fusão, nomeadamente o Trx presente no pETM-20, confere melhorias na expressão e solubilidade proteica, sendo largamente utilizados na produção de proteínas recombinantes. (Costa et al., 2013) A presença de as zonas His-tag em N-terminal apresentam-se como fundamentais naquele que será o processo de purificação proteica. Estes plasmídeos possuem ainda marcas de resistência a antibióticos para a ampicilina (pETM-10) e para kanamicina (pETM-20). A imagem da figura 3.10 confirma o sucesso das transformações de ambas as estratégias em ambos os vetores de expressão.

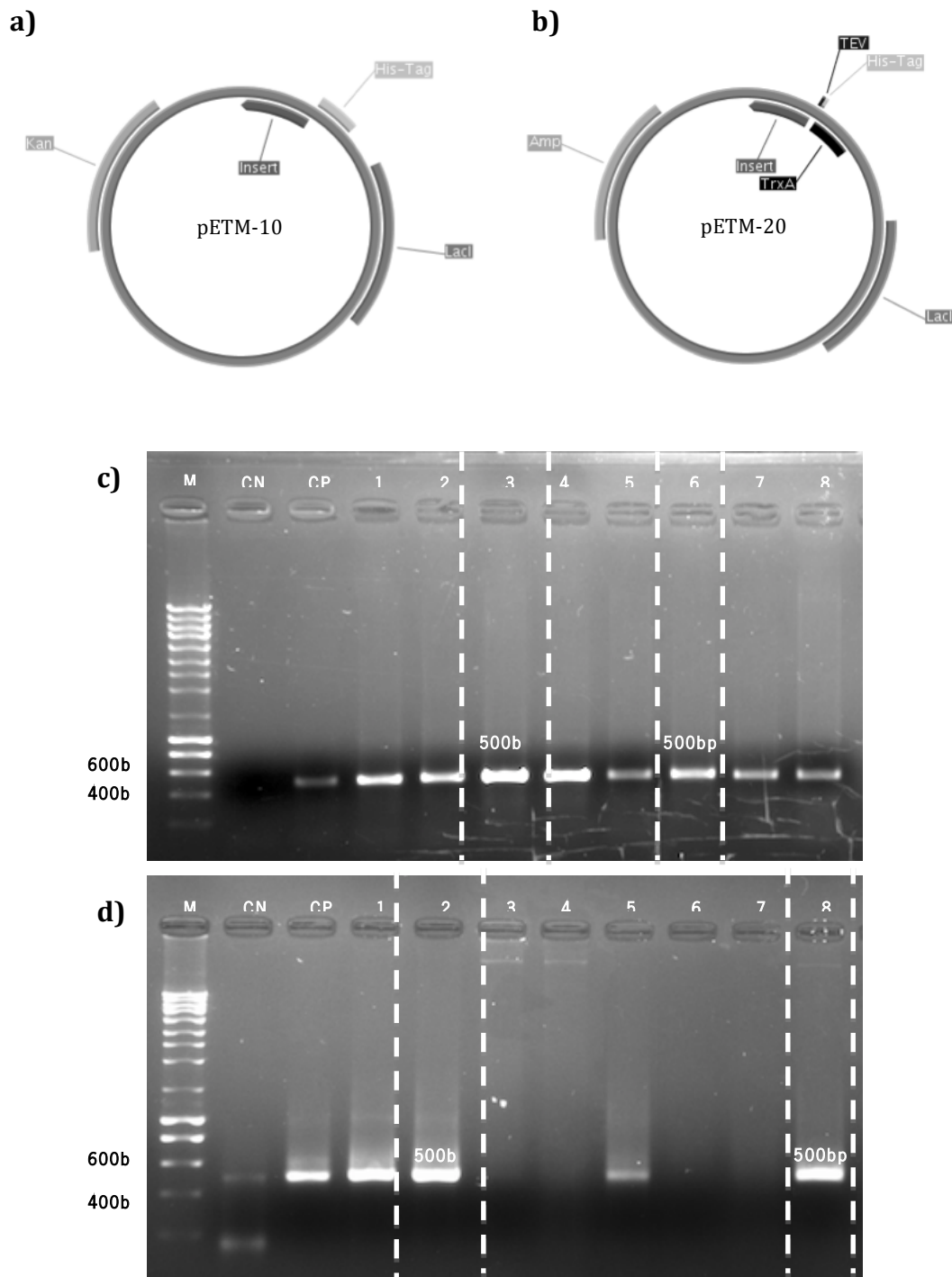


Figura 3.10: Construções FTL₁ e FTL₂ nos plasmídeos de expressão pETM-10 e pETM-20, por análise de PCR de colônia de *E. coli* transformada. (a) Esquema da construção das clonagens em vetor pETM-10; (b) Esquema da construção das clonagens em vetor pETM-20; (c) Confirmação por PCR de colônia da clonagem de FTL₁ nos vetores pETM-10 (clones de 1 - 4, com seleção do clone 3) e pETM-20 (clones de 5 - 8, com seleção do clone 6); (d) Confirmação por PCR de colônia da clonagem de FTL₂ nos vetores pETM-10 (clones de 1 - 4, com seleção do clone 2) e pETM-20 (clones de 5 - 8, com seleção do clone 8); (M) Marcador molecular NZY Ladder III; (CN) Controlo negativo; (CP) Controlo positivo

A confirmação da construção e da presença das sequências de interesse nos vetores de expressão foi feita pela excisão dos *inserts* clonados. Como recurso usaram-se as enzimas de restrição desenhadas para as estratégias (*NcoI* e *XhoI*) resultado que podemos observar e analisar pela seguinte imagem da figura 3.11.

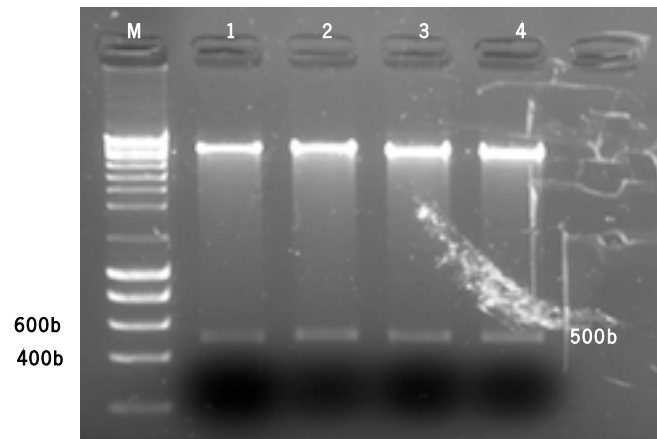


Figura 3.11: Extração dos “*inserts*” dos vetores de expressão por digestão enzimática das endonucleases *NcoI* e *XhoI*. (M) Marcador NZY Ladder III; (1) Confirmação da extração da sequência β +TEV+ α do vetor pETM-20; (2) Confirmação da extração da sequência α +TEV+ β do vetor pETM-20; (3) Confirmação da extração da sequência β +TEV+ α do vetor pETM-10; (4) Confirmação da extração da sequência α +TEV+ β do vetor pETM-10. (tamanhos moleculares esperados de 500 bp)

Os *inserts* foram sequenciados para confirmação de que as modificações genéticas na sequência da frutalina foram amplificadas com total identidade nos produtos de PCR. Os resultados mostraram uma mutação para estratégia FTL₁, correspondente à sequência β +TEV+ α , em ambos os vetores. A mutação ocorreu numa base nucleotídica pertencente à cadeia α na sua posição 364. Ocorreu a alteração na transcrição uma base timina constituinte da sequência *template* por uma base adenina. Esta alteração tem influência nos processos de co-tradução, codificando uma asparagina em vez de uma tirosina. Por outro lado para a estratégia FTL₂ (sequência α +TEV+ β) houve uma confirmação total de identidade com a sequência de nucleótidos constituintes da frutalina.

Para transformação dos vectores e respetiva expressão proteica foi escolhido como sistema de produção a estirpe de *E. coli*, *Rosetta*.

Tabela 3.1: Vetores de expressão e proteínas de fusão usados neste trabalho, e respetiva designação

Vetor pETM	Proteínas de fusão	Peso Molecular [Da]	Designação
pETM-10	MK-His ₆ - β_TEV_α	18980.4	His ₆ -FTL ₁
	MK-His ₆ - α_TEV_β		His ₆ -FTL ₂
pETM-20	M-TrxA-His ₆ -TEV- β_TEV_α	32180.46	Trx-His ₆ -TEV-FTL ₁
	M-TrxA-His ₆ -TEV- α_TEV_β		Trx-His ₆ -TEV-FTL ₂

3.4. Expressão primária de frutalina recombinante

Para cada proteína alvo foi averiguada a expressão proteica e analisados os resultados como forma de perceber se o processo ocorre de acordo com o esperado relativamente às condições de produção e de indução para as proteínas alvo. Os meios de produção são suplementados com os antibióticos correspondentes às marcas de seleção específicas dos plasmídeos de expressão e da estirpe *Rosetta*. Para a produção His₆-FTL o meio de cultura foi suplementado com o antibiótico correspondente à marca de seleção do vetor pETM-10, a kanamicina sendo que para a produção Trx-His₆-TEV-FTL o meio de cultura foi suplementado com o antibiótico correspondente à marca de seleção do vetor pETM-20, a ampicilina. Independentemente da estratégia qualquer meio de produção foi ainda suplementado com cloramfenicol, correspondente à marca de resistência da própria estirpe de *E. coli Rosetta*.

As condições de expressão foram também testadas em pequena escala. Resultados indicam que a temperatura, o tempo de indução, a concentração de IPTG e a interação entre a concentração de IPTG e o tempo de indução têm efeitos significativos sobre o nível de expressão de frutalina solúvel recombinante. (*Oliveira et al., 2009b*) As condições de indução estão correlacionadas com o peso molecular das proteínas alvo. As proteínas com pesos moleculares elevados apresentam melhor performance a temperaturas de indução mais baixas. (*Costa et al., 2013*) A temperatura de indução usada para His₆-FTL e Trx-His₆-TEV-FTL foi de 18°C.

Na tabela 3.1 estão sumarizadas as diferentes frutalinas recombinantes presentes na análise SDS-PAGE da figura 3.12. Para análise, é importante referir a observação teórica esperada, pelo que é previsto para a construção das duas estratégias no vetor pETM-10 (His₆-FTL) uma expressão proteica com um tamanho molecular de aproximadamente 18,9 kDa

enquanto que para a construção em pETM-20 (Trx-His₆-TEV-FTL) é prevista uma expressão das bandas proteicas a rondar os 32 kDa. A imagem da figura 3.12 comprova o sucesso da correta expressão das diferentes proteínas recombinantes em estudo.

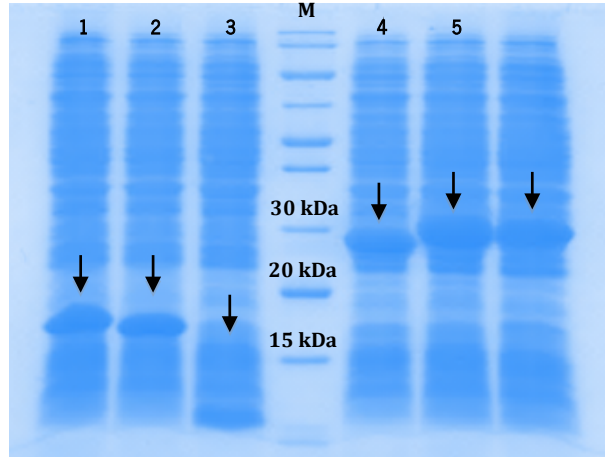


Figura 3.12: Análise de His₆-FTL e Trx-His₆-TEV-FTL das amostras obtidas da expressão em *E. coli Rosetta*, por coloração de *Coomassie Blue* e 12% SDS-PAGE. (1) Expressão de His₆-FTL₁; (2) Expressão de His₆-FTL₂; (3) Controle com vetor de expressão pETM-10; (M) Marcador proteico; (4) Controle com vetor de expressão pETM-20; (5) Expressão de Trx-His₆-TEV-FTL₁; (6) Expressão de Trx-His₆-TEV-FTL₂

A análise da produção e expressão a pequena escala permite ainda perceber outros fatores. As proteínas recombinantes podem ser direcionadas para os diferentes compartimentos celulares por péptidos de sinalização ou por meio de proteínas de fusão, pelo que diferentes locais de localização das proteínas têm diferentes vantagens e desvantagens. A construção design de proteínas de fusão possibilita uma expressão com localização no citoplasma bacteriano e boas características de solubilidade. É uma técnica vantajosa na medida que ocorre um bom nível de expressão da proteína recombinante com solubilidade e estabilidade aprimorada pelos parceiros de fusão e *tags* de afinidade comumente utilizados para técnicas de purificação proteica. Como desvantagens requer sítios de clivagem específicos aos péptidos de fusão e pode resultar num baixo rendimento nas culturas de produção. (Palomares *et al.*, 2004)

Análise por SDS-PAGE permitiu avaliar a expressão intracelular das formas recombinantes da frutalina. Foi avaliada em que fase a expressão era mais notória, se em fase solúvel ou em fase insolúvel. Para efeitos de análise das amostras foi usado para eficiência da

lise celular um protocolo de sonicação com respectiva recolha celular por centrifugação. Pela análise À imagem da figura 3.13 mostra-se que efetivamente se obteve proteína solúvel com significativa expressão.

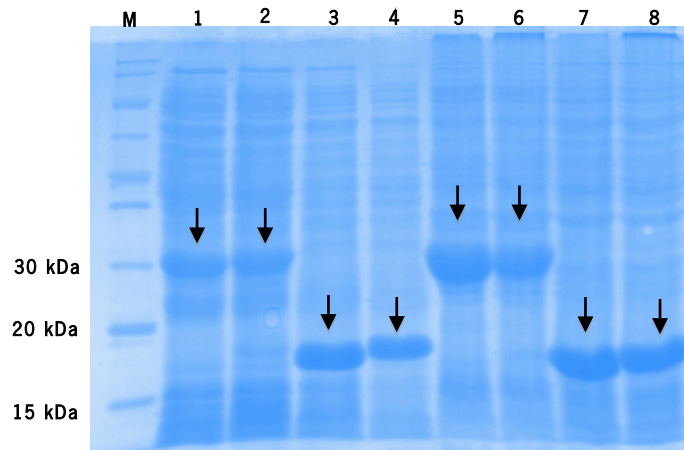


Figura 3.13: Análise de His₆-FTL (tamanho molecular esperado de 18,9 kDa) e Trx-His₆-TEV-FTL (tamanho molecular esperado de 32 kDa) das amostras das fases solúveis e insolúveis obtidas da expressão em *E. coli* Rosetta, por coloração de Coomassie Blue e 12% SDS_PAGE. (M) Marcador proteico; (1) Amostra da expressão Trx-His₆-TEV-FTL₁ (fase solúvel); (2) Amostra da expressão Trx-His₆-TEV-FTL₂ (fase solúvel); (3) Amostra da expressão His₆-FTL₁ (fase solúvel); (4) Amostra da expressão His₆-FTL₂ (fase solúvel); (5) Amostra da expressão Trx-His₆-TEV-FTL₁ (fase insolúvel); (6) Amostra da expressão Trx-His₆-TEV-FTL₂ (fase insolúvel); (7) Amostra da expressão His₆-FTL₁ (fase insolúvel); (8) Amostra da expressão His₆-FTL₂ (fase insolúvel)

3.5. Produção e purificação de frutalina recombinante expressa em *E. coli* Rosetta

O rastreio a pequena escala foi reprodutível quando observamos por comparação com a purificação das proteínas recombinantes alvo, frutalina recombinante, produzida em condições de cultura ótimas (indução com IPTG 1 mM a 18°C até uma $DO_{595nm} \approx 0,4$), foi purificada a partir de culturas de *E. coli* Rosetta por cromatografia de afinidade por iões metálicos, neste caso por níquel.

Como principal foco passam a ser as estratégias contendo as construções FTL₂, correspondente à sequência α +TEV+ β , com expressão em ambos os vetores, pETM-10 e pETM-20. Esta escolha advém dos resultados da sequenciação que mostraram 100% de correspondência para a construção. Um calculo prévio mostra que o peso molecular esperado

para a His₆-FTL₂ e Trx-His₆-TEV-FTL₂ é aproximadamente 18,9 kDa e 32 kDa, respectivamente. Desta forma, por análise de SDS-PAGE, é possível confirmar que tanto a His₆-FTL₂ (figura 3.14) como a Trx-His₆-TEV-FTL₂ (figura 3.15) foram corretamente expressas e purificadas, por IMAC-Ni, com algum nível de pureza.

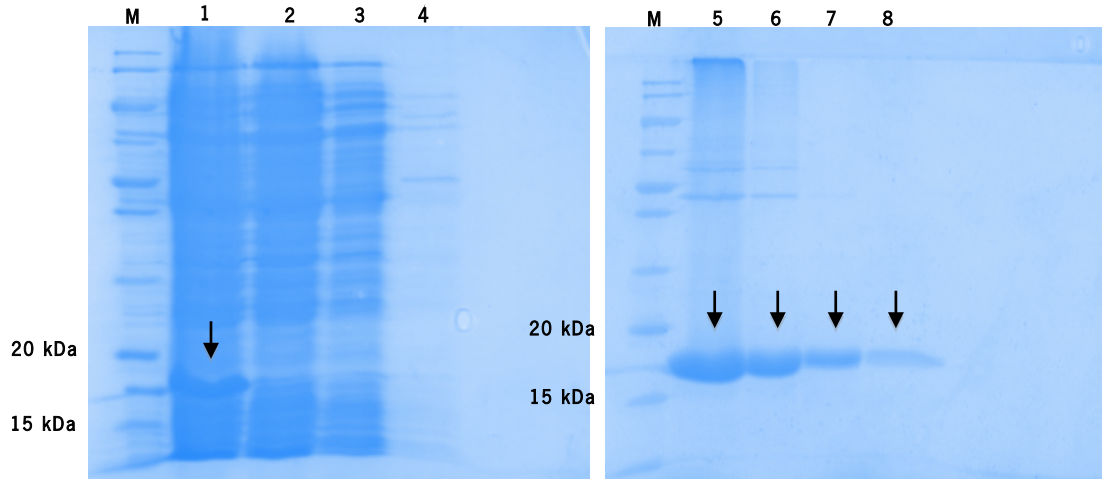


Figura 3.14: Análise de His₆-FTL₂ das amostras obtidas da purificação por afinidade a Níquel (IMAC-Ni), por coloração de *Coomassie Blue* e 12% SDS-PAGE. (M) Marcador proteico; (1) Amostra filtrada; (2) Amostra "flow-through"; (3) Amostra da passagem de solução "binding"; (4) Amostra da passagem de solução "washing"; (M) Marcador proteico; (5) Amostra da 1° eluição; (6) Amostra da 2° eluição; (7) Amostra da 3° eluição; (8) Amostra da passagem com solução EDTA

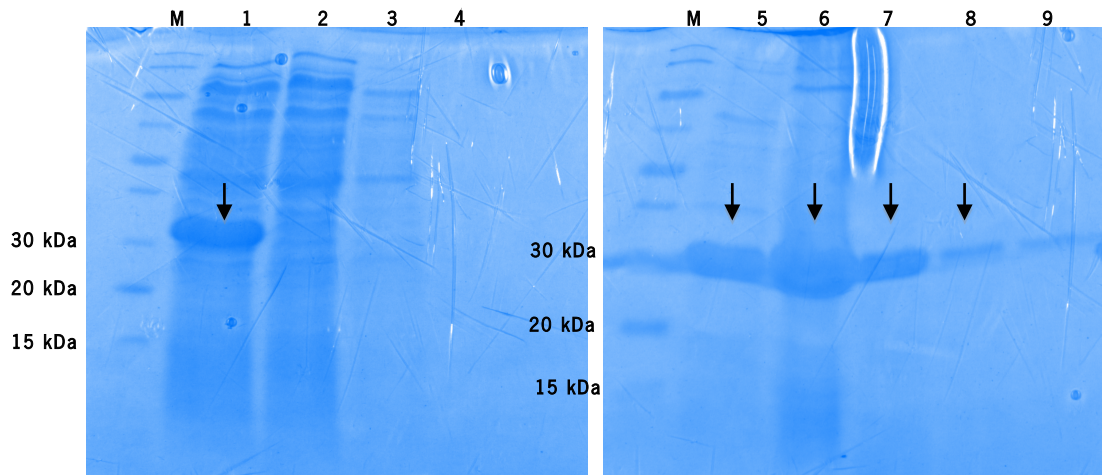


Figura 3.15: Análise de Trx-His₆-TEV-FTL₂ das amostras obtidas da purificação por afinidade a Níquel (IMAC-Ni), por coloração de *Coomassie Blue* e 12% SDS-PAGE. (M) Marcador proteico; (1) Amostra filtrada; (2) Amostra "flow-through"; (3) Amostra da passagem de solução "binding"; (4) Amostra da passagem de solução "washing"; (M) Marcador proteico; (5) Amostra da 1° eluição; (6) Amostra da 2° eluição; (7) Amostra da 3° eluição; (8) Amostra da 4° eluição; (9) Amostra da passagem com solução EDTA

Para remoção dos *tags* e clivagem de Trx-His₆-TEV-FTL₂ a amostras selecionadas já purificada foram digeridas por ação da TEV protease (figura 3.16). Desta forma é possível separar os *tags* Trx-His₆-TEV da FTL₂ ao mesmo tempo que clivamos a região ligante entre as cadeias β e α . Para interpretação a TEV protease atuará nos dois pontos de clivagem que encontra pelo que podemos observar bandas correspondentes a: Trx-His₆-TEV \approx 14,3 kDa, cadeia α +TEV \approx 15,9 kDa, cadeia β \approx 2,1 kDa, Trx-His₆-TEV-cadeia α +TEV \approx 30,1 kDa e ainda cadeia α +TEV+ β \approx 17,9 kDa. Será possível ainda observar a banda correspondente à TEV protease com um tamanho molecular de aproximadamente \approx 27 kDa.

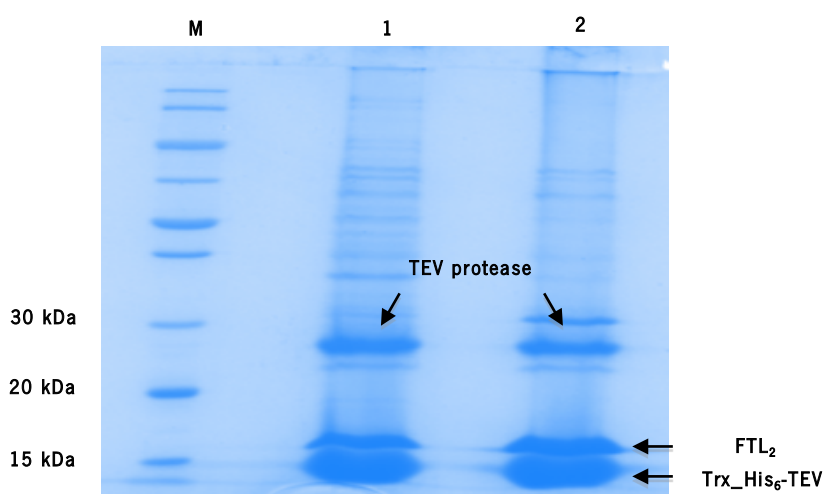


Figura 3.16: Análise da clivagem de Trx-His₆-TEV da FTL₂ por TEV protease (27 kDa). (M) Marcador molecular; (1) 1^o Amostra clivada Trx-His₆-TEV da FTL₂; (2) 2^o amostra clivada Trx-His₆-TEV da FTL₂

Pela análise da imagem da figura 3.16 confirma-se a correta clivagem das zonas de reconhecimento da TEV protease. Observa-se a banda correspondente à clivagem e separação Trx-His₆-TEV com um tamanho molecular de aproximadamente 14,3 kDa.

As proteínas clivadas foram purificadas dos *tags* de fusão e das proteases por cromatografia de afinidade por níquel (IMAC-*Ni*) mas agora numa perspectiva reversa, ou seja, a afinidade de ligação His-_{tag} ao níquel permite fazer a separação entre Trx-His₆-TEV das restantes cadeias recuperando a proteína de interesse imediatamente antes das fases de eluição. A imagem da figura 3.17 confirma a correta clivagem por ação da TEV protease bem como o sucesso no uso e recurso à técnica de purificação reversa. Por observação, é nítida a separação da proteína de interesse das restantes frações proteicas.

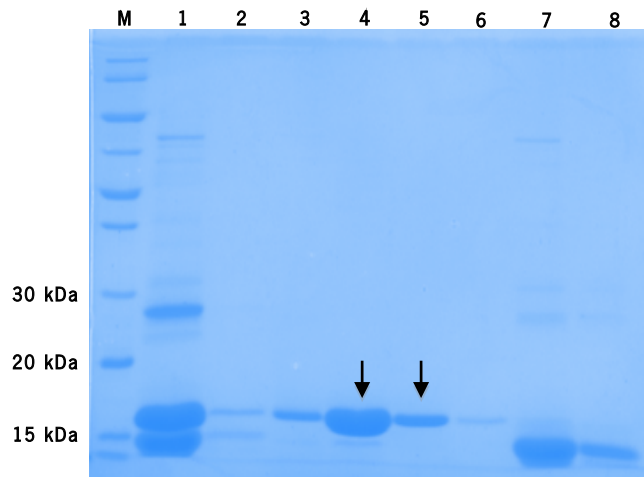


Figura 3.17: Análise de Trx-His₆-TEV-FTL₂ clivada das amostras obtidas da purificação reversa por afinidade a Níquel (IMAC-N), por coloração de *Coomassie Blue* e 12% SDS-PAGE. (M) Marcador molecular; (1) Amostra Trx-His₆-TEV-FTL₂ clivada; (2) 1^o passagem da amostra Trx-His₆-TEV-FTL₂ clivada; (3) 2^o passagem da amostra Trx-His₆-TEV-FTL₂ clivada; (4) Amostra da 1^o passagem de solução "binding"; (5) Amostra da 2^o passagem de solução "binding"; (6) Amostra da 1^o passagem de solução "washing"; (7) Amostra da 1^o eluição; (8) Amostra da passagem com solução EDTA

O processo de purificação reversa permitiu com clareza observar as diferentes frações resultantes do processo de clivagem. Nas correspondentes fases de eluição verificou-se a saída de Trx-His₆-TEV com um peso molecular a rondar os 14,3 kDa sendo que imediatamente nas fases anteriores apurou-se a saída daquela que corresponde à proteína de interesse e que apresenta um tamanho molecular superior e de 17,9 kDa no caso de ter a cadeia a proteína correspondente a α +TEV+ β . Nesta análise não se descartou a hipótese de a banda corresponder apenas à sequência α +TEV em que o seu tamanho molecular se encontra nesta gama, com cerca de 15,9 kDa. Isto significaria um erro no processo de *refolding* entre as cadeias α e β da frutalina após clivagem. Um cálculo mais experimental do tamanho das bandas observadas e marcadas na figura 3.17, indica um valor aproximado de 15,3 kDa e que reforça a ideia anterior. Por outro lado, o tamanho correspondente à cadeia β de 2,1 kDa torna-se muito pequeno para análise por géis SDS.

Deve notar-se que a atividade biológica da frutalina na sua forma nativa e recombinante são diferentes. A elevada afinidade da frutalina para resíduos de galactose, dependa da excisão correta de um péptido ligante entre as cadeias β e α . (Oliveira et al., 2011)

3.6. Caracterização proteica por *Dinamic Light Scattering* e *Fast Protein Liquid Chromatography*

A caracterização estrutural de Trx-His₆-TEV-FTL₂ clivada pode ser analisada por interpretação dos resultados obtidos por DLS e pela análise ao cromatograma de exclusão molecular obtido por sec-FPLC.

O ensaio de DLS da figura 3.18 permitiu caracterizar estruturalmente esta versão de frutalina em termos de conformação, se está organizada em tetrâmeros e ainda se, se encontra numa população homogénea, dependendo do número de picos observados. Na análise por DLS era esperado um raio_{nm} igual a 3,70, que representaria uma conformação tetramérica da frutalina recombinante. Os resultados mostram a existência de uma população homogénea correspondente a um único pico gráfico em que a proteína de análise apresenta um valor de raio_{nm} igual a 2,35 [± 0,49] o que por si representa um tamanho molecular máximo a rondar os 38 kDa, valor longe do resultado estimado e que corresponderia a 72 kDa. Este valor refuta a hipótese desta frutalina recombinante apresentar uma estrutura em forma de tetrâmero, mas possivelmente uma estrutura em forma de dímero.

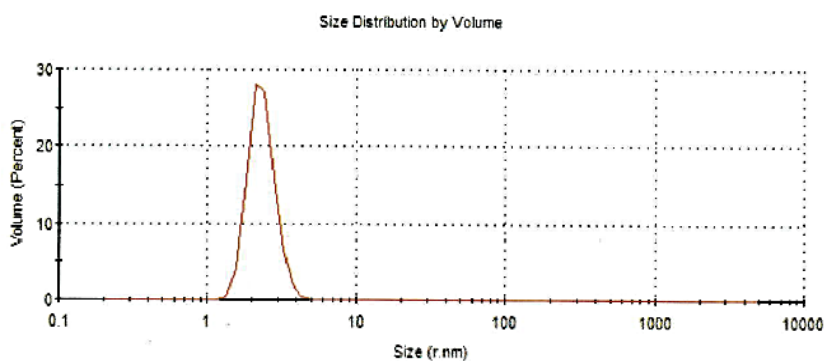


Figura 3.18: Gráfico do ensaio de DLS para uma frutalina recombinante, Trx-His₆-TEV-FTL₂ clivada. Distribuição “Tamanho vs Volume”, raio_{nm} esperado = 3,70 e raio_{nm} obtido = 2,35 ± 0,49

Esta análise pode ser corroborada por observação e interpretação dos resultados obtidos no cromatograma de exclusão molecular da figura 3.19. Esta técnica permite estimar o peso molecular da frutalina recombinante recorrendo à calibração da coluna com proteínas de peso molecular conhecido, neste caso a BSA e a frutalina nativa. Estes resultados não vão de encontro ao valor teórico esperado para a Trx-His₆-TEV-FTL₂ clivada. Em ensaios de exclusão

molecular sec-FPLC as proteínas eluem pelo tamanho, e que em primeiro lugar saem as mais pesadas pelo que, observando o gráfico a proteína clivada tem um peso inferior à BSA. Por outras palavras, observa-se um tempo na saída superior ao esperado e à proteína padrão BSA (66 kDa) e frutalina nativa (48-49 kDa). (Oliveira *et al.*, 2008) O tempo de saída registado não dista muito do correspondente ao da frutalina nativa indicando uma proximidade no tamanho molecular observado via DLS, reforçando a ideia de uma conformação em dímeros.

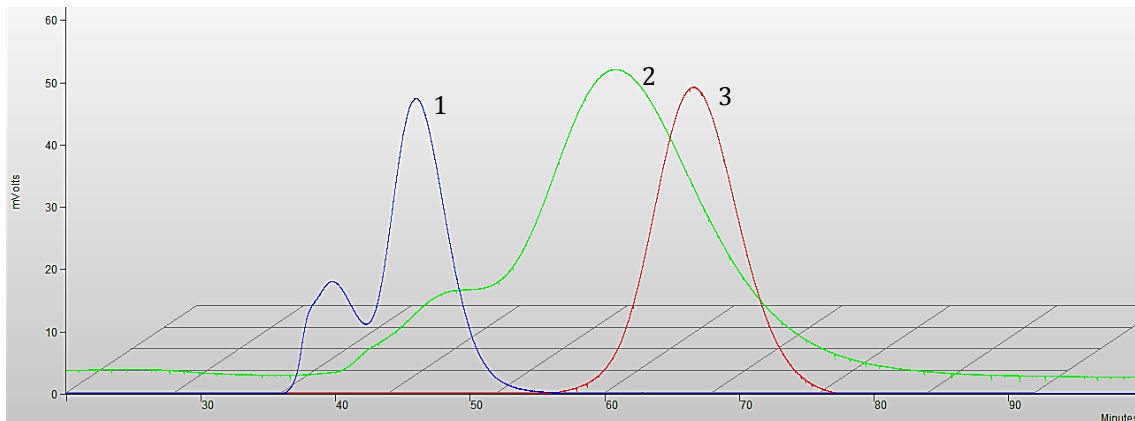


Figura 3.19: Cromatograma de exclusão molecular de BSA, frutalina nativa e Trx-His₆-TEV-FTL₂ clivada por sec-FPLC. (1) Amostra padrão de BSA; (2) Amostra padrão de frutalina nativa; (3) Amostra clivada Trx-His₆-TEV-FTL₂ [0,5 mg/ml]

Extrapolando um pouco, esta dimerização pode estar a ocorrer apenas entre cadeias α provenientes do processo de clivagem por TEV protease. (α +TEV \approx 15,9) kDa. A dimerização de α vai de encontro aos valores observados via DLS e FPLC, sendo esperado um peso molecular de aproximadamente 32 kDa. Ao comprovar-se tal hipótese seria um indicativo do mau processo de *refolding* após clivagem por TEV protease.

Capítulo 4

Conclusão e Perspetivas Futuras

4. Conclusão e Perspetivas Futuras

A construção e descrição das sequências de uma lectina recombinante, a frutalina, foram garantidas com sucesso. O objetivo principal deste estudo foi cumprido na medida em que as propostas de construção e expressão para das diferentes proteínas projetadas foram alcançadas e comprovadas. É certo que o resultado da sequenciação nucleica para a estratégia FTL₁, correspondente à construção da cadeia β +TEV+ α , apresentou uma identidade de 99% mostrando a existência de um erro mínimo na alteração de uma base nucleica suficiente para provocar uma modificação na tradução correspondente de aminoácido. Contudo, a substituição na tradução destes aminoácidos não foi avaliada, ficando por elucidar se esta alteração pode influenciar modificações pós-tradução e se por sua vez poderá afetar a atividade biológica desta frutalina recombinante. Contrariamente, o resultado da sequenciação para a estratégia FTL₂, correspondente à construção da cadeia α +TEV+ β , apresentou 100% de identidade.

Para a contabilização do sucesso da fase de clonagem, é importante referir que às construções estão associados quatro processos de PCR que permitiram a alteração com sucesso do *linker* nativo por uma zona TEV de reconhecimento à TEV protease bem como a alteração da disposição das cadeias constituintes em relação a esta região ligante. Os produtos finais resultam nas sequências de FTL₁ e FTL₂ que foram devidamente transformadas, numa primeira fase, em vetor pGEM-T Easy perfazendo um total de duas construções. Numa segunda fase FTL₁ e FTL₂ foram clonadas com sucesso em vetores de expressão pETM-10 e pETM-20 o que resulta num total de quatro clonagens finais, perfazendo um total de seis diferentes clonagens referentes às duas construções nos diversos vetores.

Ambas as clonagens das diferentes frutalinas recombinantes foram devidamente expressas em *E. coli Rosetta*, que mostra ser uma bactéria hospedeira viável como microrganismo de expressão de proteínas recombinantes, à imagem do provado por outros estudos. *E. coli*, como sistema de produção, permitiu obter quantidades proteicas viáveis para purificação, em concentrações moleculares que permitiram a sua caracterização.

A purificação, para as proteínas His₆-FTL₂ e Trx-His₆-TEV-FTL₂, foi devidamente conseguida com recurso à técnica de cromatografia de afinidade por Níquel (IMAC-*Ni*). Para a Trx-His₆-TEV-FTL₂, a remoção do *tag* de fusão Trx foi conseguida e observada por digestão pela ação de TEV protease *in vitro*. Após clivagem a recuperação da proteína de interesse foi feita com recurso à técnica de purificação reversa por IMAC-*Ni* que permitiu e comprovou a separação das frações Trx-His₆-TEV da fração da proteína de interesse.

Por último, a análise estrutural de Trx-His₆-TEV-FTL₂ clivada foi feita com base nas técnicas de DLS e FPLC e que mostrou ser um muito preliminar pelo que existe a necessidade de ser explorada. Não é propriamente um insucesso, caracterizou-se a proteína em que os valores observados sugerem que a proteína analisada não se encontra na forma estrutural tetramérica mas possivelmente numa estrutura de dímeros. Uma análise mais empírica sugere que estes valores podem refletir um mau *refolding* proteico ou até possíveis erros na clivagem da mesma.

Certamente, para este estudo, o caminho não acabaria aqui e seria ainda mais longo e com outros desafios que enriquecessem quantitativamente e qualitativamente o que para trás já foi conseguido mas, nem sempre o tempo o permite. Pensemos que o futuro será o reflexo do passado e que tem por base o presente e todas as matérias e conhecimentos de que hoje dispomos.

O círculo científico apresenta-se, nos dias de hoje, em permanente crescimento caracterizado pela necessidade da evolução e pela necessidade de perceber cada vez mais e de forma mais aprofundada as leis, fenômenos e processos. Faz parte da mentalidade científica perceber o porquê e que vantagens poderá trazer futuramente.

A área da Biologia Molecular não foge à regra e tem vindo, ao longo dos tempos, a ganhar maior importância na sociedade dos dias de hoje e é vista como uma dimensão que poderá trazer grandes vantagens futuramente naqueles que são os vários sectores da sociedade de hoje em dia.

Falando de um sector que é crítico a todos, como a saúde, e não falando apenas da prática em si mas de outras fontes também elas de extrema importância como a investigação, estudos específicos ao nível genético e proteico mostram o quanto vantajosos podem ser ao nível da biomedicina. Tem crescido aquilo que podemos chamar os números de patrocínios de estudos na área o que reflete a importância que a mesma tem presentemente mas também aquilo que representa e nos deixa expectante numa perspectiva futura.

A mesma realidade pode ser transposta para o estudo em questão. Nem sempre é possível avaliar à primeira tentativa aquilo que pode ser o potencial do objeto de investigação pelo que é necessário investir mais tempo e recursos para encontrar o caminho de aplicação viável do estudo. Em termos de perspectivas futuras é esta a principal mensagem a passar. Não sendo possível, numa primeira abordagem, fazer uma caracterização global daquilo que pode ser o potencial da frutalina como objeto de investigação, futuramente, é importante pegar nas

restantes construções e estratégias desenvolvidas neste trabalho e, por processos de produção e purificação, obter quantidades proteicas passíveis de caracterização e de utilização de estudos de atividade. A caracterização das mesmas pode ser feita com recurso às técnicas já usadas como *Dinamic Light Scattering* (DLS) ou *Fast Protein Liquid Chromatography* (FPLC). De igual forma, é importante elucidar para a proteína Trx-His₆-TEV-FTL₂ a clivagem entre as duas cadeias constituintes, β e α , para as condições testadas e se necessário otimizar a clivagem e o respectivo *refolding* proteico.

É igualmente importante avaliar a atividade proteica por estudos como de hemaglutinação, testes de imunodifusão e até testes *in vivo* antimicrobianos ou de aplicação em células de linhas tumorais como forma de avaliar o potencial anti tumoral.

Como projeto, foram tempos desafiantes que permitiram desenvolver capacidades e conhecimentos na área da investigação, nomeadamente para as disciplinas à biologia molecular.

Capítulo 5

Bibliografía

Referências Bibliográficas:

1. Ladik, J. J. (2004). "Molecular biology needs a theory." *Journal of Molecular Structure: THEOCHEM* 673(1-3): 59-64.
2. Rheinberger, H. J. (2009). "Recent science and its exploration: the case of molecular biology." *Stud Hist Philos Biol Biomed Sci* 40(1): 6-12.
3. Festa Lectins, versatile Lectins, versatile, P. (2007). "On some optimization problems in molecular biology." *Math Biosci* 207(2): 219-234.
4. Strasser, B. J. (1999). "Reviewed: Molecular Biology, Macroscopic History." *BioScience* Vol. 49(11): 929-931.
5. Morange, Michel. *History of molecular biology*. John Wiley & Sons, Ltd, 1998.
6. Echols, Harrison G., and Carol A. Gross. *Operators and promoters: The story of molecular biology and its creators*. Univ of California Press, 2001.
7. McPherson, Alexander. *Advances in Protein Chemistry and Structural Biology*. Academic Press, 2010.
8. Kennedy, J.F. and L.C.B.B Coelho et al., (1995). "Lectins, versatile proteins of recognition: a review." *Elsevier Science Limited Carbohydrate Polymers* 26: 219-230.
9. Teixeira, E. H., et al. (2012). "Biological Applications of Plants and Algae Lectins: An Overview."
10. Vasconcelos, I. M. and J. T. Oliveira (2004). "Antinutritional properties of plant lectins." *Toxicon* 44(4): 385-403.

11. Peumans, W. J. and Van Damme, J. M. (1995). "Lectins as Plant Defense Proteins." *Plant Physiol.* 109: 347-352

12. Astoul, C. H., et al. (2002). "The size, shape and specificity of the sugar-binding site of the jacalin-related lectins is profoundly affected by the proteolytic cleavage of the subunits>." *Biochem. J.* 367, 817–824.

13. Vandenborre, G., et al. (2011). "Plant lectins as defense proteins against phytophagous insects." *Phytochemistry* 72(13): 1538-1550.

14. Peumans, W. J., Hauseb, B., Van Damme, E. (2000). "The galactose-binding and mannose-binding jacalin-related lectins are located in diferent sub-cellular compartments." *FEBS Letters* 477: 186-192.

15. Liu, B., et al. (2010). "Plant lectins: potential antineoplastic drugs from bench to clinic." *Cancer Lett* 287(1): 1-12.

16. Bourne, Y. (2002). "Structural basis for the unusual carbohydrate-binding specificity of jacalin towards galactose and mannose." *Biochem. J.* 364, 173–180.

17. Fernandez-del-Carmen, A., et al. (2013). "Recombinant jacalin-like plant lectins are produced at high levels in *Nicotiana benthamiana* and retain agglutination activity and sugar specificity." *J Biotechnol* **163**(4): 391-400.

18. Fu, L. L., et al. (2011). "Plant lectins: targeting programmed cell death pathways as antitumor agents." *Int J Biochem Cell Biol* 43(10): 1442-1449.

19. Oliveira, C., et al. (2011). "Cytotoxic effects of native and recombinant frutalin, a plant galactose-binding lectin, on HeLa cervical cancer cells." *J Biomed Biotechnol* 2011: 568932.

20. Hannigan, M., Gorski, S. M., Macroautophagy: the key ingredient to a healthy diet?, *Autophagy* 5 (2009) 140–151
21. Oliveira, C., et al. (2014). "Recombinant production of plant lectins in microbial systems for biomedical application - the frutalin case study." *Front Plant Sci* 5: 390.
22. Mislovicova, D., Gemeiner, P., Kozarova, A., and Kozar, T. (2009). Lectinomics I. Relevance of exogenous plant lectins in biomedical diagnostics. *Biologia* 64, 1–19. doi: 10.2478/s11756-009-0029-3
23. Oliveira, C., et al. (2009a). "A comparative study of recombinant and native frutalin binding to human prostate tissues." *BMC Biotechnol* 9: 78.
24. Oliveira, C., et al. (2008). "Expression of frutalin, an α -D-galactose-binding jacalin-related lectin, in the yeast *Pichia pastoris*." *Protein Expression and Purification* 60(2): 188-193.
25. Oliveira, C., et al. (2009b). "cDNA cloning and functional expression of the alpha-D-galactose-binding lectin frutalin in *Escherichia coli*." *Mol Biotechnol* 43(3): 212-220.
26. Chen, R. (2012). "Bacterial expression systems for recombinant protein production: *E. coli* and beyond." *Biotechnol Adv* 30(5): 1102-1107.
27. François, B. (1999). "Recombinant Protein Expression in *E. coli*." *Current Opinion in Biotechnology*: 10:411-421.
28. Demain, A. L. and P. Vaishnav (2009). "Production of recombinant proteins by microbes and higher organisms." *Biotechnol Adv* 27(3): 297-306.
29. Sorensen, H. P. and K. K. Mortensen (2005a). "Advanced genetic strategies for recombinant protein expression in *Escherichia coli*." *J Biotechnol* 115(2): 113-128.

30. Sorensen, H. P. and K. K. Mortensen (2005b). "Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*." *Microb Cell Fact* 4(1): 1.
31. Brando-Lima, A. C., et al. (2006) Involvement of phosphatidylinositol-3 kinase–Akt and nuclear factor kappa-B pathways in the effect of frutalin on human lymphocyte. *International Immunopharmacology*. v. 6, p. 465-472.
32. Costa, S. J., et al. (2013). "The novel Fh8 and H fusion partners for soluble protein expression in *Escherichia coli*: a comparison with the traditional gene fusion technology." *Appl Microbiol Biotechnol* 97(15): 6779-6791.
33. Palomares, Laura A., Estrada-Mondaca, S., Ramírez, Octavio T., (2004). "Production of Recombinant Proteins." *Methods in Molecular Biology* 267.

Capítulo 6

Anexos

6.1. Estratégia 1: β +TEV+ α

1 CCATGGGATC CGCGGAACAA AGCGGGAAGA GCCAGACTGT AATAGTAGGA CCTTGGGGAG
 61 CCAAAGTAAG CGAGAATCTT TATTTTCAGG GCGGTAAAGC TTTTGATGAC GGTGCATTCA
 121 CCGGAATCAG AGAAATCAAC CTTCATATA ATAAGGAGAC CGCCATTGGG GACTTCCAAG
 181 TTATTACGA CTTGAATGGA AGGCCATTTG TTGGACAAAG TCATACTAGT TTTATAAAAG
 241 GCTTCACACC AGTGAAGATT TCTCTAGACT TTCCAAGCGA GTATATAGTA GAAGTGAGCG
 301 GACACACTGG TAAAGTGAGT GGGTATGTAG TAGTACGCTC TTTGACATTC AAGACTAATA
 361 AAAAAACCTA TGGACCATAT GGAGTTACAA GCGGCACACC TTTCAACCTC CCAATCGAGA
 421 ATGGCTTAGT TGTGGATTG AAAGGAAGTA TCGGCTACTG GATGGACTAC TTTAGCATGT
 481 ACTTGCCCT TCACTCGAG

Sequência a.a correspondente:

NcoI/BamHI/AEQSGKSQTVIVGPWGAKVSENLYFQGGKAFDDGAFTGIREINLSYNKETAIGDFQVIYDLNRPVFGQSHTSF
 IKGFTPVKISLDFPSEYIVEVSGHTGKVSQYVWRSLTFKTNKKTY
 GPYGVTSQTPFNLPINLPIENGLVGVFKGSGIYWMDYFSMYLSLE

6.2. Estratégia 2: α +TEV+ β

1 CCATGGGATC CCGTAAAGCT TTTGATGACG GTGCATTCAC CGGAATCAGA GAAATCAACC
 61 TTTCATATAA TAAGGAGACC GCCATTGGGG ACTTCCAAGT TATTTACGAC TTGAATGGAA
 121 GGCCATTTGT TGGACAAAGT CATACTAGTT TTATAAAAGG CTTCACACCA GTGAAGATTT
 181 CTCTAGACTT TCCAAGCGAG TATATAGTAG AAGTGAGCGG ACACACTGGT AAAGTGAGTG
 241 GGTATGTAGT AGTACGCTCT TTGACATTCA AGACTAATAA AAAAACTAT GGACCATATG
 301 GAGTTACAAG CGGCACACCT TTCAACCTCC CAATCGAGAA TGGCTTAGTT GTTGGATTCA
 361 AAGGAAGTAT CGGCTACTGG ATGGACTACT TTAGCATGTA CTTGTCCCTT GAGAATCTTT
 421 ATTTTCAGGG CGCGGAACAA AGCGGGAAGA GCCAGACTGT AATAGTAGGA CCTTGGGGAG
 481 CCAAAGTAAG TCACTCGAG

Sequência a.a correspondente:

NcoI/BamHI/GKAFDDGAFTGIREINLSYNKETAIGDFQVIYDLNRPVFGQSHTSFIKFTPVKISLDFPSEYIVEVSGHTGKVSQY
 VWRSLTFKTNKKTYGPYGVTSQTPFNLPINLPIENGLVGVFKGSGIYWMDYFSMYLSLENLYFQGAEQSGKSQTVIVGPWGAKVSLLE

Legenda:

ATG – Codão iniciação

NcoI/BamHI

Cadeia Alfa

TEV

Cadeia Beta

TCA – Codão STOP

XhoI

