

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E EVOLUÇÃO

“Expressão heteróloga, purificação e caracterização das proteínas humanas DCRA (**D**own Syndrome Critical **R**egion **G**ene **A**) e DSCR8 (**D**own Syndrome Critical **R**egion **G**ene **8**)”

Elisete Márcia Corrêa

São Carlos
2004

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E EVOLUÇÃO

**“Expressão heteróloga, purificação e caracterização das proteínas humanas
DCRA (**D**own Syndrome **C**ritical **R**egion **G**ene **A**) e DSCR8 (**D**own
Syndrome **C**ritical **R**egion **G**ene **8**)”**

Elisete Márcia Corrêa

Tese de doutorado apresentada ao Programa de Pós-Graduação em Genética e Evolução do Centro de Ciências Biológicas e da Saúde da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Doutor em Ciências, área de concentração: Genética e Evolução.

São Carlos
2004

**Ficha catalográfica elaborada pelo DePT da
Biblioteca Comunitária/UFSCar**

C824eh

Corrêa, Elisete Márcia.

Expressão heteróloga, purificação e caracterização das proteínas humanas DCRA (Down Syndrome Critical Region Gene A) e DSCR8 (Down Syndrome Critical Region Gene 8) / Elisete Márcia Corrêa . -- São Carlos : UFSCar, 2004.
142 p.

Tese (Doutorado) -- Universidade Federal de São Carlos, 2004.

1. Genética molecular. 2. Down Síndrome de. 3. DCRA.
4. DSCR8. 5. Expressão heteróloga. I. Título.

CDD: 574.87328 (20^a)

Orientador

Prof. Dr. Flávio Henrique da Silva

“A procura da verdade é difícil e é fácil, já que ninguém poderá desvendá-la por completo ou ignorá-la inteiramente. Contudo, cada um de nós poderá acrescentar um pouco do nosso conhecimento sobre a natureza e, disto, uma certa grandeza emergirá”.

Aristóteles, 350 AC.

Dedico este trabalho a todos os portadores da síndrome de Down.

Agradecimentos

Gostaria de expressar meus agradecimentos a todos aqueles que direta ou indiretamente contribuíram para a realização deste trabalho, mas principalmente àqueles que, por qualquer razão, não foram mencionados.

Ao prof. Dr. Flávio Henrique da Silva, meu orientador, pela orientação e por ter me proporcionado um amadurecimento profissional e pessoal.

Ao prof. Otávio H. Thiemman pelo uso do laboratório, sugestões e sobretudo pelo incentivo nos momentos difíceis.

A profa. Ana Paula U. Araújo pelas sugestões e colaboração em vários momentos.

A profa. Heloisa S. S. Araújo e Gláucia Maria Machado Santeli pelo uso do laboratório e pelas sugestões.

Ao prof. Odécio Cáceres pela paciência e apoio em todos os momentos.

A Elisandra M. Rodrigues pela ajuda ilimitada e apoio incondicional.

A Raquel Kelly Bortoleto Rugs e a Suzana A. Sculaccio pela ajuda nas análises e confecção dos gráficos.

Aos meus amigos de laboratórios Andréa, Daniela, César, Patrícia, Wilson, Henrique, Reis, Camillo, Sandra, Rodrigo, Wendy, as meninas (Juliana, Andréia e Viviane), Márcia, Betiana e Paulo, pelo estímulo e ajuda ilimitada. Devo muito a vocês e não há palavras que expresse o meu mais profundo agradecimento.

A Andréa e Daniela pela ajuda ilimitada e amizade sincera.

Ao Fernando Corrêa e ao Alexandre Cassago pela experiência de ensinar e aprender.

Aos Amigos Jorge Iulec e Flávia Munim pelo estímulo.

Aos meus pais e irmãos pelo apoio irrestrito.

Aos docentes e funcionários do programa de Pós Graduação em Genética e Evolução e do Departamento de Genética e Evolução, que sem dúvida contribuíram para minha formação profissional.

Lista de Ilustrações

Figura 1. Mapeamento molecular das quatro regiões mais importantes para o fenótipo da Síndrome de Down.....	20
Figura 2. Representação esquemática da localização da DSCR no cromossomo 21 humano.....	21
Figura 3. Representação esquemática da técnica de cristalização conhecida como gota suspensa. A) Gota sobre a lamínula. B) Modelo de um poço de cristalização.....	47
Figura 4. Seqüência de aminoácidos presente nas construções DCRA-formas truncadas.....	54
Figura 5. Representação esquemática da clonagem dos fragmentos do DCRA no plasmídeo pEGFP.....	54
Figura 6. Amplificação do gene DCRA a partir da biblioteca de cDNA de cérebro fetal humano.....	60
Figura 7. Análise da expressão de recombinantes pET28aDCRA.....	61
Figura 8. Análise da expressão de recombinantes pET32aTrx-DCRA.....	64
Figura 9. Análise da expressão de recombinantes pET32aTrx-DCRA _{Nco/Bam}	65
Figura 10. Análise da expressão de recombinantes pET32aTrx-DCRA _{HindIII}	69
Figura 11. Análise da expressão de recombinantes pET32aTRX-DCRA-GFP e das frações solúvel e insolúvel.....	71
Figura 12. Análise de restrição do pET41aDCRA.....	73
Figura 13. Análise da expressão de recombinantes pET37bCBD-DCRA.....	75
Figura 14. Análise da expressão de recombinantes pET41aGST-DCRA.....	74
Figura 15. Análise da expressão de recombinantes pET43cNusA-DCRA.....	75
Figura 16. Análise da expressão de recombinantes pMalcMBP-DCRA.....	76
Figura 17. Análise da expressão de recombinantes pTYB2-DCRA-inteína.....	77
Figura 18. Análise da expressão de recombinantes pCRT7C/N-DCRA.....	78
Figura 19. Análise da expressão de recombinantes pBAD-DCRA.....	79
Figura 20. A Análise da expressão de recombinantes pET28aDSCR8 e ensaio de solubilidade.....	80
Figura 21 . Análise de restrição do pET32aDSCR8.....	81
Figura 22. A Análise da expressão de recombinantes pET32aDSCR8 e das frações solúvel e insolúvel.....	82

Figura 23. Análise da purificação da DSCR8 em cromatografia de afinidade Ni-NTA <i>Superflow</i>	83
Figura 24. Seqüência de aminoácidos da proteína DSCR8.....	84
Figura 25. Distribuição subcelular da GFP-DSCR8 em célula CHO-K1 transfectada ...	84
Figura 26. Análise da purificação da Trx-DCRA-GFP em cromatografia de afinidade Ni-NTA <i>Superflow</i>	91
Figura 27. Perfil cromatográfico da eluição da tiorredoxina-DCRA-GFP na coluna de troca aniônica Mono-Q HR5/5.....	92
Figura 28. Análise da purificação em cromatografia de troca aniônica.....	93
Figura 29. Perfil cromatográfico da eluição da tiorredoxina-DCRA-GFP na coluna Superose 12 HR 10/30.....	94
Figura 30. Análise da purificação da Trx-DCRA-GFP em cromatografia de exclusão molecular Superose.....	94
Figura 31. Análise da clivagem da proteína tiorredoxina-DCRA-GFP.....	96
Figura 32. Histograma representativo do raio hidrodinâmico da proteína tiorredoxina-DCRA-GFP obtido pela técnica de espalhamento dinâmico de luz.....	99
Figura 33. Seqüência de aminoácidos codificados pelo gene DCRA.....	101
Figura 34. Alinhamento múltiplo de seqüências similares a DCRA realizado no programa <i>Multalin</i> usando a matriz Blosum62-12-2.....	102
Figura 35. Representação esquemática da formação do retrômero.....	104
Figura 36. Representação esquemática da estrutura secundária da DCRA.....	107
Figura 37. Análise do antissoro através de <i>Western-Blot</i>	108
Figura 38. Distribuição subcelular de GFP e GFP-DCRA nas linhagens CHO-K1, COS-7, e HEK 293 transientemente transfectadas.....	111
Figura 39. Distribuição subcelular de GFP-DCRA na linhagem CHO-K1 transientemente transfectada.....	112
Figura 40. Distribuição subcelular das formas truncadas da GFP-DCRA em células CHO-K1 transientemente transfectadas.....	114
Figura 41 Distribuição subcelular de Flag-DCRA-cmyc e DCRA-GFP em células transientemente transfectadas.....	115
Figura 42. Vetor pET28a (Novagen).....	118
Figura 43. Vetor pET32a (Novagen).....	119
Figura 44. Vetor pET37b (Novagen).	120

Figura 45. Vetor pET41a (Novagen).....	121
Figura 46. Vetor pET43a (Novagen).....	122
Figura 47. Vetor pTBY2 (NEB).....	123
Figura 48. Vetor pMal-c2x (NEB).....	124
Figura 49. Vetor pBad (Invitrogen).....	125
Figura 50. Vetor pCX-TOPO (Invitrogen).....	126
Figura 51. Vetor pCRT7/CT TOPO (Invitrogen).....	127
Figura 52. Vetor pCRT7/NT TOPO (Invitrogen).....	127
Figura 53. Vetor pEGFP-C1 (Clontech).....	128
Figura 54. Vetor pEGFP-N1 (Clontech).....	128
Figura 55. Vetor p3XFLAG-Myc-CMV (Sigma).....	129
Figura 56. Vetor pcDNA3 (Invitrogen).....	130
Figura 57. Vetor pBluescript II KS ⁺ (Stratagene).....	131
Tabela 1. Composição dos tampões utilizados para solubilização de proteínas superexpressas acordo com LINDWALL et al. (2000).....	41
Tabela 2. Vetores de expressão construídos para a expressão da DCRA em <i>E. coli</i>	87
Tabela 3: Predição <i>in silico</i> de motivos estruturais na proteína DCRA.....	105

Sumário

RESUMO	i
ABSTRACT	ii
1 INTRODUÇÃO	14
1.1 A Síndrome de Down.....	14
1.2 Mapeamento molecular do cromossomo 21.....	18
1.2.1 Delineamento molecular da região crítica (DSCR).....	19
1.3 Genes DCRA e DSCR8.....	23
2 OBJETIVOS	25
3 MATERIAL E MÉTODOS	26
3.1 Obtenção da ORF DCRA a partir de biblioteca de cDNA de cérebro fetal humano..	26
3.2 Clonagem da ORF DCRA no vetor de propagação.....	27
3.2.1 Construção do vetor pBluescript II KS ⁺ -DCRA.....	27
3.3 Subclonagem das ORFs DCRA e DSCR8 nos vetores de expressão para <i>E. coli</i>	28
3.3.1 Construção do pET 28a-DCRA	28
3.3.1.2 Construção do pET 28a-DSCR8	28
3.3.2 Construção do vetor pET32a-DCRA.....	29
3.3.2.1 Construção do vetor pET32a-DSCR8.....	31
3.3.3 Construção dos vetores pET 37b-DCRA, pET 41a-DCRA	31
3.3.4 Construção do vetor pET 43c-DCRA.....	32
3.3.5 Construção do vetor pMALC2-DCRA	33
3.3.6 Construção do vetor pTYB2-DCRA.....	34
3.3.7 Construção do vetor pBAD-TOPO-DCRA.....	35
3.3.8 Construção do vetor pCX TOPO-DCRA.....	36
3.3.9 Construção dos vetores pCRT7/CT TOPO-DCRA e pCRT7/NT TOPO-DCRA...	36
3.4 Subclonagem da ORF DCRA no vetor de expressão para células de mamífero.....	37
3.4.1 Construção do vetor pcDNA3-DCRA.....	37
3.5 Hospedeiros usados.....	37
3.6 Análise dos recombinantes.....	38
3.7 Expressão da proteína recombinante.....	39
3.8 Purificação de proteínas.....	40

3.8.1 Fracionamento com sulfato de amônio.....	41
3.8.2 Cromatografia por Afinidade.....	42
3.8.3 Cromatografia de Troca iônica.....	43
3.8.4 Cromatografia de filtração em gel.....	44
3.9 Clivagem com a enteroquinase.....	45
3.10 Espalhamento Dinâmico de Luz (DLS).....	45
3.11 Ensaio de cristalização.....	46
3.12 Análise da seqüência de aminoácidos <i>in silico</i> usando o pacote <i>ExpASy</i> para similaridade, modificações pós-traducionais, topologia, estrutura primária e secundária.....	48
3.13 Produção de anticorpos policlonais contra a DCRA humana em camundongos.....	49
3.13.1 Preparação das amostras para imunização.....	50
3.13.2 Imunização dos animais para produção do antissoro.....	50
3.13.3 Imunodeteção.....	51
3.14 Localização subcelular em células de mamífero.....	52
3.14.1 Construção do plasmídeo pEGFPC1-DCRA.....	52
3.14.2 Construção do plasmídeo pEGFPC1-DCRA-formas truncadas.....	53
3.14.3 Construção do vetor p3XFLAG- <i>myc</i> -CMV25-DCRA.....	55
3.14.4 Construção do vetor pEGFPN1-DCRA.....	56
3.14.5 Construção do plasmídeo pEGFPC1-DSCR8.....	56
3.14.6 Cultura de células Transfecção Celular.....	57
3.14.6.1 Linhagens de células utilizadas.....	57
3.14.6.2 Transfecção Celular.....	57
3.14.7 Análise da localização subcelular da DCRA, DCRA-formas truncadas e DSCR8	58
4 RESULTADOS E DISCUSSÃO	60
4.1 Amplificação da ORF DCRA a partir da biblioteca de cDNA de cérebro fetal humano.....	60
4.2 Subclonagem e expressão da DCRA no vetor pET28a	60
4.3 Subclonagem e expressão da DCRA no vetor pET32a.....	63
4.3.1 Subclonagem e expressão da Tiorredoxina-DCRA _{Nco/Bam} no vetor pET32a.....	64
4.3.2 Subclonagem e expressão da Tiorredoxina-DCRA _{Hind III} no vetor pET32a.....	68
4.3.3 Subclonagem e expressão da Tiorredoxina-DCRA-GFP no vetor pET32a.....	69
4.4 Subclonagem e expressão da DCRA nos vetores pET37b, pET41a e	

pET43c.....	72
4.5 Subclonagem e expressão da DCRA no vetor pMALc2.....	75
4.6 Subclonagem e expressão da DCRA no vetor pTYB2.....	77
4.7 Subclonagem e expressão da DCRA nos vetores PCRT7/CN-TOPO, pBAD-TOPO e pCX-TOPO.....	78
4.8 Subclonagem e expressão da DSCR8 no vetor pET28a.....	79
4.8.1 Subclonagem, expressão e ensaios de purificação da DSCR8 no vetor pET32a.....	80
4.8.2 Análises <i>in silico</i> e de localização subcelular da DSCR8.....	83
4.9 Expressão da proteína DCRA em <i>E. coli</i>	85
5 Purificação da proteína Trx-DCRA-GFP.....	90
5.1 Cromatografia NI-NTA <i>Superflow</i>	90
5.2 Cromatografia Mono-Q HR5/5.....	91
5.3 Cromatografia Superose 12 HR10/30.....	93
6 Clivagem da proteína Trx-DCRA-GFP com a enteroquinase.....	95
7 Espalhamento Dinâmico de Luz.....	98
8 Ensaios de cristalização.....	99
9 Predições <i>in silico</i>	100
9.1 Estrutura primária da proteína DCRA.....	101
9.2 Similaridade.....	101
9.3 Predição de domínio.....	102
9.4 Modificações pós-traducionais.....	104
9.5 Topologia.....	105
9.5.1 PSORT II.....	106
9.5.2 TargetP.....	106
9.6 Estrutura secundária.....	106
10 Anticorpos.....	108
11. Localização subcelular da DCRA em células de mamífero.....	110
12 CONCLUSÕES	117
13 ANEXOS	118
13.1 Mapas de plasmídeos.....	118
13.2 Meios de cultura.....	132
14. REFERÊNCIAS BIBLIOGRÁFICAS	134

RESUMO

A síndrome de Down é a causa mais freqüente de retardo mental que afeta milhões de pessoas em todo o mundo e resulta de uma trissomia completa ou parcial do cromossomo 21 (HC21). Casos raros de trissomia parcial permitiram identificar uma pequena região do HC21 comum a todos os portadores denominada de Região Crítica da Síndrome de Down. Os genes DCRA e DSCR8 foram mapeados nesta região e codificam proteínas de função desconhecida. Com o objetivo de contribuir para a caracterização destas proteínas DSCR8 foi expresso em fusão com seqüência codificadora da tiorredoxina na linhagem *de Escherichia coli* Rosetta (DE3) o que possibilitou a obtenção de Trx-DSCR8 na fração solúvel em quantidade suficiente para os ensaios iniciais de purificação em cromatografia de afinidade Ni-NTA. O DCRA foi clonado em fase com seqüências codificadores da tiorredoxina e GFP o que permitiu que a solução protéica fosse parcialmente purificada e que ensaios de cristalização e imunização pudessem ser realizados. A seqüência de aminoácidos da proteína DCRA possui 57% e 49 % de identidade respectivamente com proteínas de *Anopheles gambiae* e de *Drosophila melanogaster*, ambas de função desconhecida. Análises *in silico* da seqüência de aminoácidos demonstram que DCRA possui um provável domínio de Vps26. A localização subcelular da DCRA em fusão com a GFP foi observada preferencialmente no compartimento citoplasmático em todas as linhagens celulares analisadas. Esses resultados contribuem para melhor conhecimento da DCRA e da DSCR8 e estimulam futuras análises para o entendimento da função celular destas proteínas, bem como suas implicações na síndrome de Down.

ABSTRACT

Down syndrome is the most frequent cause of mental retardation affecting millions of people worldwide and results from full or partial trisomy of chromosome 21 (HC21). Rare cases of partial trisomy allowed the identification of a small region in HC21 common to all carriers called Down Syndrome Critical Region. The genes DCRA and DSCR8, mapped to this region, encode two proteins of unknown function. With the aim of contributing to a better characterization of these proteins, DSCR8 was subcloned and expressed in fusion with thioredoxin in *Escherichia coli* Rosetta (DE3). Recombinant Trx-DSCR8, observed in the soluble fraction, was subjected to the initial purification assays by affinity chromatography in a Ni-NTA column. The gene DCRA was cloned and expressed in fusion with the proteins thioredoxin and GFP allowing the partial purification of the protein and the achievement of crystallization and immunological assays. The amino acid sequence of DCRA showed 57% and 49% of identity to a protein of unknown function from *Anopheles gambiae* and *Drosophila melanogaster* respectively. Also, *in silico* analyses revealed that DCRA contains a putative Vps26 domain. Subcellular localization of DCRA in fusion with GFP was observed preferentially in the cytoplasm of the cell lines tested. These results contribute to a better characterization of DCRA and DSCR8 and open up possibilities towards the understanding of the cellular role of these proteins and their relationship with the Down syndrome.

Expressão heteróloga, purificação e caracterização das proteínas humanas DCRA (Down Syndrome Critical Region Gene A) e DSCR8 (Down Syndrome Critical Region Gene 8)

1 Introdução

1.1 A Síndrome de Down

O primeiro relato científico da Síndrome de Down (SD) foi há 137 anos (DOWN, 1866), quando o inglês, John Langdon Down, publicou na Inglaterra um ensaio que descreve um grupo de crianças com fenótipos comuns, porém distintos de outras crianças com retardo mental.

No seu relato Down atribuiu essa condição à tuberculose materna e se referia aos portadores como “Mongolóides” por achar que as crianças se pareciam com as pessoas da Mongólia. Essa denominação etnicamente incorreta foi questionada e o termo foi mudado para *Down's Syndrome*. Pouco tempo depois a comissão americana de revisão de termos científicos sugeriu a mudança para *Down Syndrome* (LESHIN, 2000).

Quase um século depois Waardenburg sugeriu que a síndrome era uma consequência de uma anormalidade cromossômica e em 1959 Lejeune e colaboradores confirmaram a presença da trissomia de um cromossomo do grupo G, originalmente identificado como número 21, em nove crianças portadoras dos fenótipos descritos por Down (LEJEUNE et al., 1959).

Os trabalhos com microscopia de fluorescência, no entanto mostraram que na verdade o cromossomo presente em triplicata é o formalmente chamado de número 22 (BURNS & BOTTINO, 1991). Como os cromossomos são numerados por ordem decrescente de tamanho e tendo em vista a classificação da trissomia do 21 ter se tornado tão difundida na literatura, houve concordância em renumerar os dois membros do Grupo G, com o 22 (menor dos autossomos) ficando antes do 21 (segundo menor dos autossomos) (BURNS & BOTTINO 1991 op cit).

A SD ocorre em todos os grupos étnicos e em ambos os sexos. Do ponto de vista citológico pode ser encontrado: o (1) trissômico para o 21, com um complemento somático total de 47 cromossomos (aproximadamente 95% dos casos de Síndrome de Down resultam de trissomia completa do 21, devido a não disjunção cromossômica de origem materna (ANTONARAKIS, 1991)); (2) translocação em cerca de 5% dos portadores, no

qual o 21 extra está total ou parcialmente ligado a outro cromossomo acrocêntrico, usualmente nos grupos D (13-15), G (21-22) (CAPONE, 2001) ou mais frequentemente no cromossomo 14 ou 21 (THULINE & PUESCHEL, 1982; HOOK, 1982). Um cariótipo de uma translocação apresenta 46 cromossomos. Portanto, em ambos os tipos anteriores há um número 21 extra, no primeiro caso há um cromossomo a mais e no segundo o 21 se ligou a uma parte de outro cromossomo. A terceira possibilidade e a mais rara delas ocorre quando a não disjunção se dá no ovo fecundado, resultando em dois diferentes tipos de células com 46 e 47 cromossomos no mesmo indivíduo. Essa condição é conhecida como mosaicismo (MIKKELSEN, 1977) com a frequência de 2 a 4% dos casos.

O mosaicismo pode ocorrer quando o zigoto tem três cromossomos 21 resultante de uma trissomia e durante as divisões celulares uma ou mais linhagens celulares perdem a cópia extra do cromossomo 21 ou quando o zigoto inicia as divisões celulares com um par de cromossomos 21 e durante o curso dessas divisões ocorre à duplicação de um cromossomo 21 (LESHIN, 2000).

A Síndrome de Down é a desordem cromossômica não letal mais comum e afeta 1 em cada 650 a 1.000 recém-nascidos (HOOK, 1982). Contudo, a incidência da Síndrome depende da idade materna, período gestacional em que é realizado o diagnóstico, término da gestação e casos de perda no pré-natal (MORRIS, MUTTON & ALBERMAN, 2000).

ANTONARAKIS (1993) estudou mais de 400 famílias procurando determinar a associação entre a idade materna, origem de uma cópia extra do cromossomo 21 e se essa origem era de natureza meiótica ou mitótica. Os resultados demonstraram que o erro na meiose que leva a trissomia é na grande maioria (67-70%) de origem materna, ocorre durante o desenvolvimento fetal da mãe (meiose I) e está associado com a idade da progenitora acima de 31,2 anos; enquanto que em somente 25% dos casos surgem durante a ovulação (meiose II) e raramente (5%) ocorrem durante a espermatogênese. Apesar de um predomínio de erros ocorrerem durante a meiose observou-se que em 5% dos indivíduos a trissomia era devido a um evento ocorrido na mitose. Esses resultados sugerem que a não disjunção não é uniforme em todos os cromossomos e gestações.

Ao contrário da clara associação entre a idade materna e o fator de risco para a SD, a existência de um efeito positivo da idade paterna ainda é controversa (FISCH et al., 2003).

JYOTHY et al. (2001) estudaram a relação entre a idade dos progenitores e a origem de uma cópia extra do cromossomo 21 e concluíram que mulheres com 30,34 anos apresentam uma taxa de erro na não disjunção dos cromossomos de 79,24% enquanto, que

homens com 31,04 anos apresentam 20,76%. Portanto, a incidência da SD está inequivocamente relacionada à idade materna. Após os 40 anos a mulher tem maior probabilidade de ter filhos com Down. Aos 30 anos o risco é de 1 para 1.000, enquanto que aos 40 anos é de 9 em 1.000 (PENROSE, 1933). PENROSE (1961) apud GUSMÃO, TAVARES & MOREIRA (2003) afirma que o aumento na taxa de trissomia do 21 em mães com mais de 35 anos é devido à no mínimo 2 eventos, um dependente da idade materna e o outro independente. O primeiro dependente da idade compreenderia o processamento anormal de um bivalente vulnerável na meiose I. O segundo evento, independente da idade materna, poderia estar associado a uma configuração vulnerável de quiasmas e a mutação de genes que atuam no processo de disjunção meiótica. (HASSOLD & SHERMAN, 2000, apud, GUSMÃO, TAVARES & MOREIRA, 2003, op cit.).

Embora a maioria dos trabalhos associe maternidade tardia aos ricos da SD, recentes estudos têm sugerido que polimorfismos em genes envolvidos com o metabolismo do folato poderiam ser fatores de risco materno para a Síndrome de Down (CHARLOTE et al., 2000). O folato é importante para a síntese *de novo* do nucleotídeo pirimidínico (dTMP) e também para as reações normais de metilação, incluindo a metilação do DNA. A associação entre a deficiência de folato e a hipometilação do DNA sugere que deficiências nutricionais e/ou genéticas que afetam o metabolismo do folato contribuiriam para o risco da não-disjunção e a Síndrome de Down (CHARLOTE et al., 2000 *op cit.*).

A SD é a mais freqüente forma de retardo mental causada por uma desordem cromossômica. Poderia ser entendida como uma síndrome complexa de origem genética e epigenética que apresenta conseqüências neurobiológicas somada a várias manifestações neuro desenvolvimentais típicas (CAPONE, 2001).

As características clínicas da SD são congênitas e incluem principalmente: retardo mental (como já mencionado), hipotonia (fraqueza) muscular, baixa estatura, anomalia cardíaca, perfil achatado, orelhas pequenas com implantação baixa, olhos com fendas palpebrais oblíquas, língua grande protusa e subsulcada, encurtamento dos quintos dígitos, aumento de distância entre o primeiro e o segundo artelho e prega única nas palmas. Anormalidades do sistema imune, do trato gastrointestinal, risco aumentado de leucemia e em pacientes adultos, freqüentemente, alterações atróficas características da doença de Alzheimer (EPSTEIN, 1986; TSUKARA et al., 1998).

Outras desordens também encontradas e que exigem monitoramento ou intervenção são: perda auditiva em 90% dos portadores de SD (MAZZONI, et al., 1994; KATTAN et al., 2000; SHOTT et al., 2001); desordem oftalmológica (80% das crianças com idades

entre 5 a 12 anos) e problemas ortodônticos (ROIZEN & PATTERSON 2003); menor densidade óssea (SEPULVEDA et al., 1995); redução do metabolismo (ALLISON et al., 1995) o que contribui para a maior incidência de obesidade e outros distúrbios que ocorrem na população em geral, mas que são mais freqüentes nos portadores de SD tais como artrite, Diabetes Mellitus (menos de 1%), apnéia obstrutiva do sono, vitiligo e dermatite (ROIZEN & PATTERSON op. cit. 2003).

A trissomia exerce também um poderoso efeito sobre o quociente de inteligência (QI). Há uma significativa redução do QI nos primeiros anos de vida. No adulto o QI varia entre 25 a 55 o que equivale a uma escala de retardo moderado a severo. Contudo, há indivíduos com muito pouco ou nenhum déficit de QI (NIZETIC, 2001).

Em relação à fertilidade pode-se dizer que a menarca e as características do ciclo menstrual são similares tanto para a paciente com SD como para qualquer outra adolescente (GOLDSTEIN, 1998). Os homens, no entanto, apresentam capacidade reprodutiva diminuída enquanto que as mulheres, embora seja raro, podem gerar crianças normais ou afetadas (BOVICELLI et al., 1982).

Apesar dos inúmeros problemas apresentados, da expectativa de vida para os pacientes com SD ter aumentado nos últimos tempos, dos aspectos clínicos dessa síndrome ser conhecidos e suas conseqüências no desenvolvimento humano também, ainda não se sabe qual é a razão da não disjunção cromossômica, quais mecanismos básicos estão por trás do efeito da idade materna (PERSEN & MIKKELSEN 2000), ou porquê a trissomia do 21 é mais compatível com a vida do que as outras trissomias descritas e o que determina a natureza e a severidade dessa anomalia.

Segundo MAO et al. (2003) há duas hipóteses principais para explicar como a trissomia do cromossomo 21 resulta na SD. A primeira é a hipótese da *instabilidade no desenvolvimento* que diz que o desequilíbrio global de genes do 21 é responsável pela diversidade fenotípica observada na SD (SHAPIRO, 1997). A outra hipótese sugere que a presença de um pequeno número de genes numa região crítica do CH21 contribuiria para o complexo fenótipo da SD (KORENBERG et al., 1990). Contudo até o presente momento essas hipóteses não são necessariamente mutuamente exclusivas.

Assim, um dos grandes desafios é integrar os conhecimentos moleculares sobre os genes do cromossomo 21 com a complexa alteração fenotípica observada na Síndrome de Down.

1.2 Mapeamento molecular do cromossomo 21

O cromossomo 21 (CH21) e seus genes ocupam lugar de destaque na genética humana devido ao seu pequeno tamanho e envolvimento não somente com a Síndrome de Down, mas também com outras doenças genéticas tais como doença de Alzheimer, desordens autoimunes, epilepsia, e susceptibilidade a desordem afetiva bipolar (GUIMERA et al., 1997). Conseqüentemente tem sido alvo de intensos estudos, mas somente quarenta e um anos após a comprovação de sua implicação com a SD, HATTORI et al. (2000) publicaram na revista Nature a seqüência genômica completa do 21q. Essa revelação foi o fruto de um consórcio entre Japão-Alemanhã e colaborações de laboratórios norte-americanos, franceses, ingleses e suíços.

É o menor cromossomo autossômico humano, representa cerca de 1 a 1,5% do genoma. O braço longo (21q) possui cerca de 33,5 Mb e apresenta grandes diferenças na densidade gênica, uma vez que a metade distal de q possui maior quantidade de genes do que a metade proximal (HATTORI et al., 2000 op cit). É na região distal 21q22 que se acredita conter loci responsáveis pelos fenótipos clínicos da síndrome (SHAPIRO, 1999).

O braço curto (21p) consiste de uma região organizadora de nucléolo que contém múltiplas cópias de genes que codificam RNAs ribossômicos, seqüências de DNA altamente repetitivas (ANTONARAKIS, 1998) e pelo menos 1 gene que codifica uma putativa tirosina fosfatase de função desconhecida (HATTORI et al., id. ibid 2000). Além disso, os genes do 21p não são essenciais para o desenvolvimento normal, já que duplicações ou deleções desta região raramente resultam em manifestação clínica detectável (CAPONE, 2001).

No trabalho original Hattori descreve que o cromossomo 21 é composto de 225 genes e 59 pseudogenes. Desses, 127 correspondem a genes conhecidos, 98 representam putativos novos genes preditos *in silico*. Desses 98, 13 são similares a proteínas conhecidas, 17 são ORFs anônimas caracterizando domínios modulares e os 68 restantes são unidades de transcrição desconhecidas sem similaridades com proteínas já descritas.

A anotação inicial do CH21 foi apenas o ponto de partida, refinamentos posteriores detectaram que muitos genes estavam incompletos, não tinham validação experimental ou não foram relatados (GARDINER et al., 2000; REYMOND et al., 2001). Uma atualização dessa notação realizada em abril de 2003 estimou 329 genes dos quais 165 foram confirmados experimentalmente, 150 foram baseados em bancos de EST (Expressed Sequence Tag) e 14 foram preditos *in silico* (ROIZEN & PATTERSON 2003).

Uma análise mais refinada sobre a classificação funcional dos genes do 21 publicada alguns meses depois do término do seqüenciamento, por GARDINER & DAVISSON (2000), revelou que 17 genes codificam reguladores de transcrição, 8 codificam quinases e 4 são moléculas de ubiquitinação. Entre os genes do 21, aproximadamente 35% são homólogos a *Drosophila*, 35% a *C. elegans* e 18% a *Sacharomyces cerevisiae*.

Outra impressionante observação citada por ROIZEN & PATTERSON (2003 op cit.) é que alguns dos genes do 21 formam conjuntos de genes envolvidos com a mesma via metabólica ou sistema biológico, por ex. pelo menos 16 genes ou genes preditos participam do metabolismo energético e no mínimo 9 genes participariam do desenvolvimento do sistema nervoso.

Portanto, não há dúvidas que o seqüenciamento do CH21 terá implicações para as pesquisas sobre SD, uma vez que muito pouco é conhecido sobre a patofisiologia molecular da Síndrome de Down.

1.2.1 Delineamento molecular da região crítica (DSCR)

Desde 1970 numerosas tentativas foram realizadas visando determinar se regiões específicas do cromossomo 21, que quando em triplicata, são responsáveis pelo fenótipo da SD (SHAPIRO, op. cit. 1997).

Embora a maioria dos casos de SD seja causada por três cópias do CH21 há um número raro de pacientes clinicamente caracterizados que apresentam somente trissomia parcial do 21, trissomia parcial sem os fenótipos da SD ou fenótipos da SD sem qualquer anormalidade cromossômica aparente (ANTONARAKIS, 1998).

Essa diversidade fenotípica parecia indicar a presença de uma região específica que provavelmente conteria os genes responsáveis pela anomalia. Esses relatos sobre a trissomia parcial e as propostas de NIEBUH, 1974 *apud* KORENBERG et al. (1994) de que a banda q22 seria “patogênica” para a SD colaboraram para a construção de um mapa fenotípico para a referida síndrome.

Análises moleculares posteriores realizadas por RAHMANI et al. (1989) sugeriram que uma pequena região da banda q22 próxima ao marcador D21S55 era crítica para a

expressão fenotípica da SD. Essa região cromossômica tem sido denominada de *Região Crítica da Síndrome de Down* (DSCR).

DELABAR et al. (1993) realizaram o mapeamento molecular de vinte e quatro características da SD e concluíram que 13 das 24 características principais estavam presentes em uma região de 4 Mb ao redor do marcador D21S55 entre DS21S17 e EST2.

SINET et al. (1994), estudando as correlações fenótipo-genótipo definiram as regiões mínimas que estão ligadas à expressão de vinte e três características clínicas dessa síndrome, do seguinte modo: a região D21S55 ou região 1 da Síndrome de Down (DCR1) que corresponde a 1/20 do braço longo acima da região 21q22.2-21q22.3 proximal está relacionada com quatro características da doença: retardo mental, crescimento retardado, hipotonia muscular e hiperlaxia de junções, e oito das dezoito anormalidades morfológicas mais comuns na face, mãos e pés. A sobreposição da DCR1, a região D21S55-MX1 ou DCR2 (1/10 do braço longo), que compreende a sub-banda 21q21.2 até ¼ da parte proximal da região 21q22.3, está envolvida com as anormalidades contempladas na DCR1 somado aos defeitos cardíacos congêntos, e mais cinco anormalidades morfológicas adicionais (Figura 1).

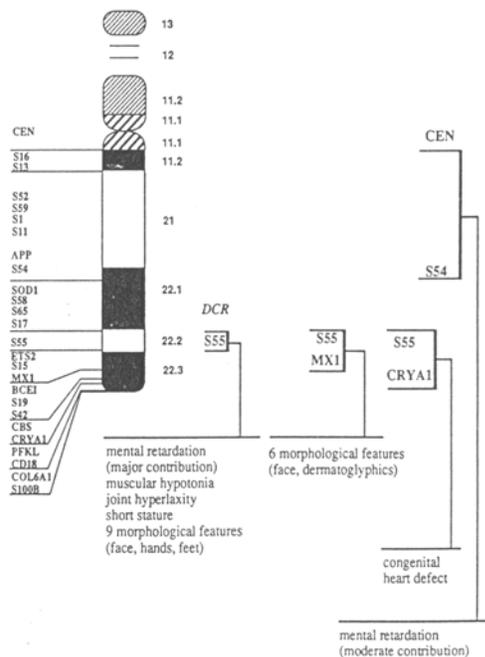


Figura 1. Mapeamento molecular das quatro regiões mais importantes para o fenótipo da Síndrome de Down. (Fonte: SINET et al., 1994).

LUCENTE et al. (1995) redefiniram o limite da área crítica de 4 Mb para 2,5 Mb entre os marcadores D21S17 e ERG (em 21q22). Posteriormente esse intervalo foi restringido para uma região de 1,6 Mb entre o sítio de Not I, LA68 (D21S396 distal a D21S17) e ERG por OHIRA et al. (1996) (Figura 2).

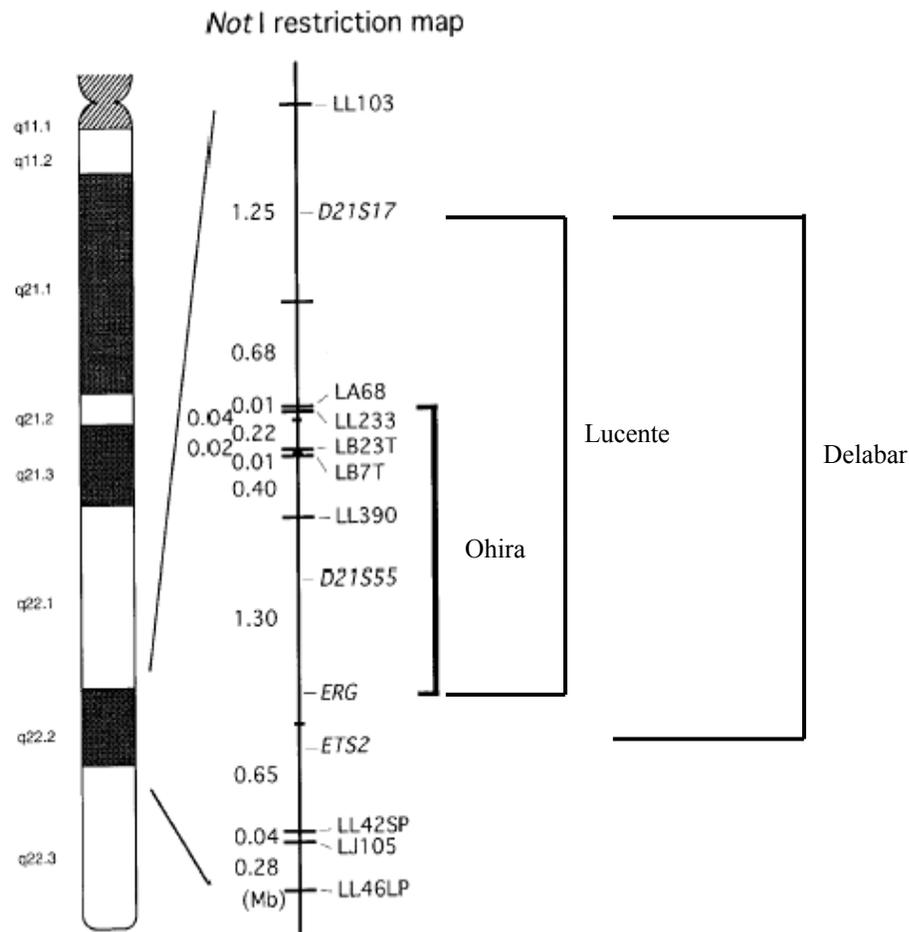


Figura 2. Esquema da representação da localização da DSCR no cromossomo 21 humano, sugerida por Delabar (ao redor do marcador D21S55 entre DS21S17 e EST2), Lucent (D21S17 e ERG (em 21q22)) e Ohira (sítio de *Not* I, LA68 (D21S396 distal a D21S17) e ERG) (Fonte: OHIRA et al. 1996).

Portanto, análises de genes contidos na DSCR trariam importantes contribuições para o entendimento da SD. Entretanto, não se pode excluir a possibilidade de nenhum segmento fora da DSCR conter genes relevantes para a etiologia dessa síndrome (KORENBERG, et al., 1994; NIZETIC, 2001). Para ANTONARAKIS (2001) o conceito

da região crítica deve ser reavaliado através de estudos com um maior número de pacientes com trissomia parcial, pela análise sistemática dos fenótipos e pelo diagnóstico das regiões genômicas em triplicata por meio de abordagens atuais.

Além disso, a hipótese do efeito de dose gênica (*gene dosage effects*) tem sido desafiada (GREBE-PLATZER et al., 1999) e a desregulação da expressão de genes localizados em outros cromossomos tem sido sugerida (BIRCHLER et al., 2001), a qual poderia ser secundária ao desequilíbrio cromossômico ou uma consequência direta da síndrome (ENGIDAWORK & LUBEC, 2001). Para CHRAST et al. (2000) os vários fenótipos da Síndrome de Down poderiam ser diretamente originados da superexpressão dos genes do cromossomo 21 ou indiretamente da interação do produto destes genes com o genoma, transcriptoma e proteoma.

Assim a identificação de genes do CH21 e a caracterização de suas funções são particularmente importantes para o entendimento da SD. Por esta razão os genes localizados na região crítica e superexpressos em tecidos especificamente afetados são de especial interesse.

Um desses genes é o DSCR1 clonado por FUENTES et al. (1995). Esse gene localizado em 21q22.1-q22-2 é altamente expresso no cérebro, coração e músculo esquelético. Compreende 7 éxons, dos quais 4 (éxons 1-4) são alternativos. A proteína MCIP1 codificada por DSCR1 pertence a uma família de três membros de proteínas conservadas evolutivamente: DSCR1, ZAKI-4 e DSCR1L2 e interage física e funcionalmente com calcineurina A (FUENTES et al., 2000). A interação entre MCIP1 e calcineurina modularia vias dependentes de calcineurina que controlariam o crescimento e programas seletivos de expressão gênica em músculo estriado (ROTHERMEL et al., 2000). Além disso, a localização celular e atividade da MCIP1 são provavelmente reguladas por quinases e/ ou fosfatases (PFISTER, et al., 2002). Vega et al. (2002) sugerem que MCIP1 pode ser fosforilada por MAPK e que quando fosforilada torna-se substrato para calcineurina (uma fosfatase com função crítica em muitos processos fisiológicos tais como ativação de células T, apoptose, hipertrofia cardíaca) competindo com outros substratos tais como fator nuclear de ativação das células T (NFAT).

Outros genes isolados e clonados na região crítica são: DCRB ou DSCR4 (NAKAMURA, A., HATTORI, M., SAKAKI, Y. 1997 a); DCRR1 (EKI et al., 1997), TPRD (TSUKAHARA, F. et al., 1998); WUB (EGEO, A. et al., 1998); DSCR5 e DSCR6 (SHIBUYA et al., 2000); DSCR2 também denominado c21LRP está localizado na região crítica dois entre os marcadores D21S55-MX1 e provavelmente relacionado com

proliferação celular (VIDAL-TABOADA et al., 2000); DSCR9 e DSCR10 identificados por predição computacional seguida por clonagem do cDNA. São preferencialmente expressos em testículos, codificam respectivamente proteínas com 149 e 87 aminoácidos de função desconhecida e provavelmente são genes primata-específicos (TAKAMATSU et al., 2002).

1.3 Genes DCRA e DSCR8

O gene designado por DCRA (**D**own **S**yndrome **C**ritical **R**egion **G**ene **A**) ou DSCR3 (**D**own **S**yndrome **C**ritical **R**egion **G**ene **3**), foi clonado da região 21q22.2 por NAKAMURA A., HATTORI, M., SAKAKI, Y. (1997 b), possui o tamanho 3.252 pb e consiste de 8 éxons. Seu cDNA codifica uma fase aberta de leitura de 297 resíduos de aminoácidos.

Análises via *Northen blot* detectaram um transcrito de DCRA de aproximadamente 3.3 kb em cérebro fetal e adulto, coração, pulmão, fígado, rins, placenta e um transcrito de 4 kb no músculo esquelético. Fato semelhante também ocorreu no dcra de camundongo (esse gene possui 92% de identidade com o DCRA humano) no qual transcritos de 2,2 Kb e 1,4 kb foram detectados em tecidos embrionários (de 7 a 17 dias) e adultos de cérebro, coração, músculo esquelético quando foi utilizada uma sonda humana. Estes diferentes transcritos poderiam ser resultantes de um processamento alternativo ou de diferentes inícios de transcrição (NAKAMURA, A., HATTORI, M., SAKAKI, Y. *id. ibid.*).

Segundo NAKAMURA, A., HATTORI, M., SAKAKI, Y (op. cit.) DCRA exhibe identidade com o gene H β 58 de camundongo. O gene H β 58 codifica um mRNA de 2,7 kb expresso durante o desenvolvimento embrionário, fetal e em muitos tecidos somáticos adultos. Além disso, é fortemente conservado entre as aves e os mamíferos (>96% de identidade) (LEE, et al., 1992). O fato do gene H β 58 estar relacionado com o desenvolvimento embrionário reforça a hipótese de que o DCRA poderia estar envolvido com o desenvolvimento humano.

O DSCR8 (*Down Syndrome Critical Region Gene 8*) também é conhecido como MTAG2, C21orf65 ou 21_38480505. Foi formalmente designado DCR1-24.0 por Dahamane et al. (1998) e anotado logo após a catalogação do cromossomo 21 (TOYODA

et al., 2002). É composto por 4 éxons, codifica uma ORF de 276 pb e codifica uma proteína de 91 aminoácidos de função desconhecida.

Dois anos após a publicação da anotação do CH21, TOYODA et al. (op. cit.) usaram a ausência de conservação e o baixo nível de expressão para justificar a eliminação do DSCR8 da lista de genes do cromossomo 21 previamente descritos como codificadores de proteína (GARDINER et al., 2003).

REYMOND et al. (2002) refinaram a anotação do CH21 utilizando abordagens experimentais e algoritmos múltiplos e encontraram seqüências similares a DSCR8 em região sintênica de camundongo. Além disso, identificaram este transcrito através da construção de mapa de transcrição do cromossomo 21q22.2 e detectaram a expressão deste gene em testículo, cérebro adulto e fetal, útero, placenta, glândula salivar, músculo, traquéia, rim fígado e glândula adrenal.

Dando suporte contrário a afirmativa de TOYODA está ainda o fato do transcrito DSCR8 apresentar 100% de identidade com MMA-1a (*malignant melanoma-associated 1*) um EST (*expressed sequence tag*) presente em linhagens celulares de melanoma humano. A expressão do MMA-1a foi detectada exclusivamente em testículo após 30 ciclos de PCR, mas exibiu uma significativa expressão em melanoma primário (20%) e melanoma metastático (30%) após somente um ciclo da PCR além de ser detectado em outras linhagens de células tumorais e sarcomas (de WIT, et al., 2002).

Portanto, embora a função real da DCRA permaneça desconhecida e a existência da proteína DSCR8 ainda seja questionada, acredita-se que a superexpressão destes genes cause certas mudanças fisiológicas que poderiam estar envolvidas com a SD. Portanto, a identificação e a caracterização funcional dos genes mapeados na região crítica do cromossomo 21 são fundamentais para o entendimento da patogênese da SD e das anormalidades que caracterizam o fenótipo clínico. Com base nesse contexto a motivação deste trabalho é contribuir para o entendimento da função biológica do produto protéico do DCRA e do DSCR8 visando auxiliar no entendimento da função desses genes bem como suas implicações na Síndrome de Down.

2 Objetivos

O presente trabalho teve como objetivo geral à expressão heteróloga, purificação e caracterização das proteínas humanas DCRA e DSCR8.

Para tanto os seguintes objetivos específicos foram desenvolvidos:

- 1- Clonagem do gene DCRA a partir de biblioteca de cDNA de cérebro fetal humano.
- 2- Subclonagem em vetores de expressão e a determinação da melhor construção para a expressão da DCRA e da DSCR8 em sistema procarioto.
- 3- Purificação das proteínas recombinantes através da cromatografia de afinidade.
- 4- Análise da seqüência de aminoácidos *in silico* usando *ExpASy site* para similaridade, modificação pós-traducionais, topologia, estrutura primária e secundária.
- 5- Produção de anticorpos policlonais contra a DCRA humana em camundongos.
- 6- Localização subcelular das proteínas DCRA e DSCR8 em células de mamífero.

O produto amplificado foi analisado em gel de agarose 1%, em tampão TAE [1X] e visualizado pela coloração com brometo de etídeo para confirmação da amplificação.

O produto de PCR foi purificado em coluna S 400 Micro Spin™ (Amersham Biosciences) de acordo com as instruções do fabricante.

3.2 Clonagem da ORF DCRA no vetor de propagação

3.2.1 Construção do vetor pBluescript II KS⁺-DCRA

O pBSIIS⁺ (Stratagene) é um vetor de 2961 pb derivado do pUC19. Possui o sítio de múltipla clonagem flanqueado pelos promotores T3 e T7 e a porção do gene LacZ que possibilita a α -complementação para seleção de colônias azuis e brancas.

O pBSIIS⁺ foi obtido pelo método da lise alcalina das células bacterianas utilizando-se o Kit CONCERT (GIBCO BRL). Aproximadamente 3 μ g do DNA plasmidial foram clivados com *Eco* RV.

Após 3 horas de clivagem a 37°C o DNA plasmidial foi separado em gel de agarose 1% e purificado utilizando o Kit Concerttm Rapid Gel Extraction System (GIBCO BRL). O DNA recuperado foi quantificado e utilizado na mistura de ligação.

A mistura de ligação vetor-inserto foi feita na razão equimolar de 3:1, utilizando-se aproximadamente 90ng/ μ L do inserto DCRA e 100 ng/ μ L do vetor pBSIIS⁺.

A mistura de ligação foi utilizada para transformar as células de *E. coli* DH5 α quimicamente competente por tratamento com cloreto de cálcio (AUSUBEL et al., 1994).

Os recombinantes (colônias brancas) foram analisados através da clivagem com as enzimas de restrição *Nde* I e *Bam* HI. A junção (inserto:vetor) e a integridade do gene foram confirmadas pelo seqüenciamento.

3.3 Subclonagem das ORFs DCRA e DSCR8 nos vetores de expressão para *E. coli*

3.3.1 Construção do pET 28a-DCRA

O vetor pET 28a (Novagen) é capaz de produzir uma proteína de fusão com His-tag, no N e no C-terminal e codificar uma seqüência de reconhecimento para a protease trombina.

O aminoácido histidina não é um resíduo comum na seqüência da proteína uma vez que sua freqüência é de somente 2,1% em proteínas globulares (KLAPPER, 1977). Assim há uma alta probabilidade de que a proteína superexpressa contendo resíduos adicionais de histidina no C ou no N-terminal possa ser preferencialmente separada das demais principalmente porque o grupo imidazólico do aminoácido histidina liga-se a Ni imobilizado no grupo NTA (seção 3.8.2).

O inserto DCRA presente no vetor pBSIIKS⁺ foi clivado com as enzimas *Nde* I e *Bam* HI e ligado ao vetor pET 28a linearizado com as mesmas enzimas.

Todos os passos de linearização dos vetores pBSIIKS⁺ e pET28a, isolamento e recuperação do inserto e do vetor do gel de agarose, preparação da mistura de ligação e transformação foram realizados de modo semelhante ao descrito em 3.1 e 3.2.1.

O clone recombinante pET28aDCRA foi confirmado por análise de restrição e após sua verificação 100ng desse DNA plasmidial foram utilizados para transformação de *E. coli* BL21 (DE3).

A análise da indução da expressão da proteína nas células de *E. coli* BL21(DE3) e *Epicurian coli* BL21(DE3)–Codon PlusTM (Stratagene) com concentrações de 0,4 a 1mM de IPTG foi feita em SDS-PAGE 15% (LAEMMLI,1970).

3.3.1.2 Construção do pET 28a-DSCR8

O inserto DSCR8 clonado no vetor pGEM-T (gentilmente cedido pela Prof. Dra. Anamaria A. Camargo (Instituto Ludwig-SP)) foi clivado com as enzimas *Nde* I e *Eco* RI e ligado ao vetor pET 28a linearizado com as mesmas enzimas.

Todos os passos de linearização dos vetores pGEM-T e pET28a, isolamento e recuperação de inserto e do vetor do gel de agarose, preparação da mistura de ligação e transformação foram realizados de modo semelhante ao descrito em 3.1.

Os recombinantes pET28aDSCR8 foram confirmados por análise de restrição e após sua verificação 100ng desse DNA plasmidial foram utilizados para transformação das células de *E. coli* BL21 (DE3) e de Rosetta (DE3).

A análise da indução da expressão da proteína com concentrações de 0,4 a 1mM a de IPTG a 37°C, 30°C e 25°C foi feita em SDS-PAGE 15% (LAEMMLI,1970).

3.3.2 Construção do vetor pET32a-DCRA

O vetor pET 32a (Novagen) permite a expressão da proteína alvo em fusão com o C-terminal de uma seqüência codificadora de 109 aminoácidos da proteína tiorredoxina (Trx.TagTM) em fase com His-tag, sítio para trombina, S-tag, sítio para enteroquinase e uma His-tag opcional no C-terminal, o que facilita a purificação em cromatografia de afinidade.

O pET32a foi utilizado para a clonagem da DCRA em quatro situações, sendo que a primeira foi para a clonagem da DCRA em fusão com a seqüência codificadora do C-terminal da Trx; a segunda para a fusão Trx-DCRA sem o N-terminal; a terceira para a fusão Trx-N-terminal da DCRA e a última para a fusão Trx-DCRA-GFP.

Para a primeira construção o vetor pET32a foi clivado inicialmente com *Nco* I, tratado com a nuclease *Mung Bean* de acordo com as especificações de SAMBROOK & RUSSEL (2001) e depois digerido com *Bam* HI.

O gene DCRA foi retirado do vetor pET28a através da linearização deste com as enzimas *Nde* I (Fermentas), preenchimento com dNTPs pela atividade da *Klenow* polimerase (Amersham Biosciences) de acordo com as especificações do fabricante, e *Bam* HI (GIBCO BRL).

O inserto e o vetor após a digestão foram purificados de um gel de agarose 1% utilizando-se o Kit Nucleo Spin (Macherey-Nagel).

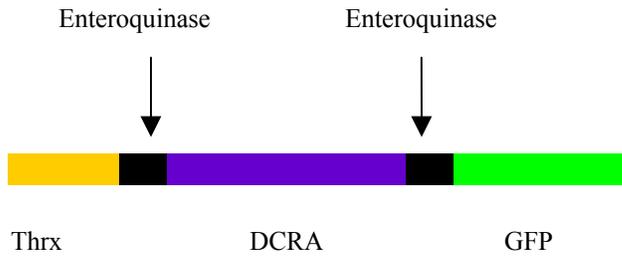
Em seguida, plasmídeo e inserto foram ligados durante a noite a 16°C. A mistura de ligação foi utilizada para transformar *E. coli* DH5α cálcio competente.

Para a segunda construção o gene DCRA foi retirado do vetor pET28a com as enzimas *Nco* I e *Bam* HI e subclonado no vetor pET32a clivado com as mesmas enzimas.

Essa construção (denominada Tiorredoxina-DCRA_{Nco/Bam}) permitiu a retirada de 126 pb iniciais do gene DCRA, devido à presença de um sítio de *Nco* I no gene, mas preservou a fase de leitura.

Para verificar a contribuição dessa fração de 126 pb na expressão gênica, o fragmento contendo os primeiros 180 pb do DCRA foi clonado em fase com uma seqüência que codifica o C-terminal da tiorredoxina (terceira construção). Assim a região codante foi amplificada com o *primer* sense que anela no códon ATG iniciador e o reverso que flanqueia o final do gene. A seguir o amplificado foi clivado com *Hind* III (o DCRA possui um sítio reconhecido pela *Hind* III na 179ª pb), resultando em um fragmento de 180 pb respectivamente com a extremidade 5' cega e a 3' coesiva, purificado e ligado ao vetor pET 32a linearizado com as enzimas *Eco* RV e *Hind* III.

Finalmente o DCRA foi clonado em fase com seqüências que codificam respectivamente o C-terminal da tiorredoxina e o N-terminal da GFP. Essa construção (denominada Tiorredoxina-DCRA-GFP) fez parte do projeto de iniciação científica do aluno Fernando Corrêa, intitulado “Construção de um novo plasmídeo para expressão de proteínas heterólogas em fusão com a GFP (Green Fluorescent Protein)”. O plasmídeo pEGFP-N1 vector (Clontech) foi usado como molde para a amplificação da região codante da *Green Fluorescent Protein* (EGFP) com os *primers* sense (5' atc cga cga cga cga caa gat ggt gag caa ggg cga gga 3') possuindo a seqüência atc seguido do sítio de reconhecimento da enteroquinase (sublinhado) e reverso 5' tga gct cga gtt act tgt aca gct cgt cca tgc cga g 3' (sítio para *Xho* I sublinhado). As bases atc no oligonucleotídeo sense reconstituem o sítio de *Eco* RV e a base C evita que a GFP fique em fase de leitura com a Trx. Esse amplificado foi clonado no vetor pET32a previamente clivado com as enzimas *Eco* RV and *Xho* I. Essa construção denominada de pET32aGFP foi linearizada com *Eco* RV para a clonagem do produto de PCR do gene DCRA amplificado com o *primer* sense (5' atg ggg acc gcc ctg gac atc 3') e o *primer* reverso que retirou o códon de finalização original do gene (5' cct atc ctg cag agc ttc ag 3'). A retirada do códon de terminação possibilitou que DCRA fosse lida em fase com GFP como representado no esquema abaixo.



A análise da indução da expressão da proteína nas células de *E coli* AD494 e de BL21 (DE3) (somente para a construção Tiorredoxina-DCRA-GFP) foi feito em SDS-PAGE 15% (LAEMMLI,1970).

3.3.2.1 Construção do vetor pET32a-DSCR8

A ORF DSCR8 foi retirada do vetor pGEM-T com as enzimas *Bam* HI e *Eco* RI e clonada no vetor pET 32a linearizado com as mesmas enzimas.

O inserto e o vetor após a digestão foram purificados de um gel de agarose 1% utilizando-se o Kit Nucleo Spin (Macherey-Nagel).

Em seguida plasmídeo e inserto foram ligados durante a noite a 16⁰C. A mistura de ligação foi utilizada para transformar células de *E. coli* DH5 α cálcio competente.

Os recombinantes foram analisados através da clivagem com as enzimas de restrição *Bam* HI e *Eco* RI e a integridade do gene foi confirmada pelo seqüenciamento.

A análise da expressão da proteína na linhagem Rosetta (DE3) foi feita em SDS-PAGE 15% (LAEMMLI,1970).

3.3.3 Construção dos vetores pET 37b-DCRA, pET 41a-DCRA

O vetor pET 37b (Novagen) permite a expressão de proteínas em fusão com CBD (domínio de ligação a celulose) de aproximadamente 13 kDa que pode ser retirada pela clivagem proteolítica com trombina ou fator Xa, além de uma seqüência sinalizadora que envia a proteína híbrida para o espaço periplasmático. Esse domínio CBD possui afinidade por matrizes de celulose e a ligação entre este e a celulose é dada pela interação

hidrofóbica em pH neutro, a eficiente eluição é conseguida com etileno glicol, que é inócuo para a maioria das proteínas.

Já o vetor pET 41a (Novagen) permite a clonagem em fusão com uma seqüência codificadora de 220 aminoácidos da GST (Glutathione S. Transferase), além das caudas de histidina e de seqüência de clivagem reconhecida pela enteroquinase. A vantagem desse vetor é que a proteína de fusão pode ser purificada em matriz com glutathione immobilizada.

Os vetores pET 37b, pET41a previamente digeridos com *Nco* I foram tratados com a nuclease *Mung Bean* de acordo com as especificações de SAMBROOK & RUSSEL (2001) e depois digeridos com *Bam* HI.

O DCRA foi retirado do vetor pET28a com as enzimas *Nde* I (GIBCO BRL) e *Bam* HI (Amersham Biosciences). Após a digestão com *Nde* I, o sítio foi preenchido com dNTPs pela atividade da *Klenow* polimerase (Amersham Biociences), de acordo com as especificações do fabricante, exceto pelo tempo de incubação que foi de 30 minutos a 37°C e posteriormente digerido com *Bam* HI.

O inserto e o vetor após a digestão foram purificados de um gel de agarose 1% utilizando-se o Kit Nucleo Spin (Macherey-Nagel).

Em seguida plasmídeo e inserto foram ligados durante a noite a 16°C. A mistura de ligação foi utilizada para transformar células de *E. coli* DH5α cálcio competente.

Os recombinantes foram analisados através da clivagem com as enzimas de restrição e a integridade do gene foi confirmada pelo seqüenciamento.

A análise da indução da expressão da proteína em células de *E. coli* BL21(DE3) em fusão com seqüências codificadoras do domínio de ligação a celulose (CBD *tag*) e da proteína GST, respectivamente no pET 37b e pET41a foi feita em SDS-PAGE 15% (LAEMMLI,1970).

3.3.4 Construção do vetor pET 43c-DCRA

O vetor pET 43 (Novagen) permite a clonagem em fusão com uma seqüência que codifica 495 aminoácidos da proteína NusA de aproximadamente 55,0 kDa, além das seqüências *His-tag* e *S-tag*, em fase com a *Nus-tag* e sítios de clivagem para trombina e enteroquinase. A NusA de *E. coli* é um fator que medeia o término da transcrição (TSUGAWA et al. 1985) e é uma das proteínas parceiras que poderiam aumentar a

solubilidade de proteínas insolúveis, não somente pelas suas características, mas também devido ao seu alto nível de expressão (DAVIS et al., 1999).

O vetor pET43c previamente digerido com *Sac* I foi tratado com a nuclease *Mung Bean* de acordo com as especificações de SAMBROOK & RUSSEL (2001) e depois digerido com *Bam* HI.

A ORF DCRA foi amplificada a partir do vetor pET28a-DCRA com a enzima polimerase *Platinum High Fidelity* (GIBCO BRL) utilizando o oligonucleotídeo sense que iniciou a amplificação no códon ATG iniciador e o reverso que possibilitou a inserção de um sítio de clivagem para a endonuclease *Bam* HI na extremidade 3' do gene.

A seguir, inserto e vetor foram recuperados de gel de agarose 1% e ligados durante a noite a 16⁰C. A mistura de ligação foi utilizada para transformar *E. coli* DH5 α cálcio competente.

Os recombinantes foram analisados através da clivagem com enzima de restrição e a integridade do gene foi confirmada pelo seqüenciamento.

A análise da indução da expressão da proteína em *E. coli* BL21(DE3) em fusão com a proteína NusA foi feita em SDS-PAGE 10% (LAEMMLI,1970).

3.3.5 Construção do vetor pMALC2-DCRA

O vetor pMALC2 (New England) foi desenhado para criar fusões entre o inserto e o gene *malE* de *E. coli*, que codifica a proteína ligadora a maltose (MBP). Esse vetor possui o promotor *tac* que é induzível por IPTG. Permite a seleção de colônias recombinantes pelo X-Gal (5 bromo-4-cloro 3-indolil- β -D-galactosídeo) e a purificação em coluna de resina de afinidade a MBP.

O vetor pMALc foi inicialmente digerido com *Eco* RI preenchido com dNTPs pela atividade da *Klenow* polimerase e posteriormente clivado com *Sal* I.

O DCRA foi retirado do vetor pET28a (Novagen) com as enzimas *Nde* I (Fermentas) e *Sal* I (GIBCO BRL). Após a digestão com *Nde* I, o sítio foi preenchido com dNTPs pela atividade da *Klenow* polimerase, de acordo com as especificações do fabricante, e posteriormente digerido com *Sal* I.

O inserto e o vetor após a digestão foram purificados do gel de agarose 1% utilizando-se o Kit Núcleo Spin (Macherey-Nagel).

Em seguida, plasmídeo e inserto foram ligados durante a noite a 16⁰C. A mistura de ligação foi utilizada para transformar *E. coli* DH5 α cálcio competente.

A análise da indução da expressão da proteína em *E. coli* BL21(DE3) e BL21 codon plus com diferentes concentrações de IPTG foi feita em SDS-PAGE 10% (LAEMMLI,1970).

3.3.6 Construção do vetor pTYB2-DCRA

O vetor pTYB2-DCRA (NEB) foi desenhado para criar fusões em fase entre o C-terminal da proteína de interesse e o N-terminal da inteína. A porção C-terminal da inteína encontra-se em fusão com o CBD (*Chitin Binding Domain*) que permite a purificação em resina de afinidade a quitina. A inteína sofre autoclivagem no seu N-terminal sob baixa temperatura e na presença de grupos tiol dispensando a necessidade de proteases específicas para a retirada da CBD-tag.

O vetor pTYB2-DCRA foi clivado com *Nde* I e *Sma* I (Amersham Pharmacia Biothec) e recuperado de um gel de agarose 1% utilizando-se o Kit Núcleo Spin (Macherey-Nagel).

A ORF DCRA foi amplificada utilizando-se o vetor pET28a-DCRA como molde e a polimerase *Platinum High Fidelity* (GIBCO). O oligonucleotídeo sense (seqüência descrita na seção 3.1) utilizado na amplificação inseriu um sítio de *Nde* I e o reverso (5'tat cct gca gag ctt cag cgg3') possibilitou a retirada do códon de parada.

O produto de PCR amplificado foi purificado utilizando-se o Kit Concerttm Rapid PCR Purification System (GIBCO BRL), clivado com *Nde* I e recuperado do gel de agarose 1%. Em seguida plasmídeo e inserto foram ligados durante a noite a 16⁰C. A mistura de ligação foi utilizada para transformar *E. coli* DH5 α cálcio competente.

Os recombinantes foram analisados através da clivagem com enzima de restrição e a integridade do gene foi confirmada pelo seqüenciamento.

A indução da expressão da proteína em *E. coli* ER2566 em fusão com o N-terminal da inteína foi feita com 0,3 mM de IPTG e visualizada em SDS-PAGE 10% (LAEMMLI,1970). ER2566 é uma linhagem alternativa para expressão de vetores com promotor T7 e possui uma cópia cromossômica do gene codificador da RNA polimerase do T7 sob o controle do promotor *lac*.

3.3.7 Construção do vetor pBAD-TOPO-DCRA

O vetor pBAD-TOPO (Invitrogen) permite a rápida clonagem de produtos de PCR em fusão com uma cauda de seis histidinas. Possui um promotor, P_{BAD}, induzível por L-arabinose. A forte regulação do promotor pelo produto do gene AraC é útil para a expressão de genes potencialmente tóxicos ou essenciais (GUZMAN et al., 1995).

O vetor pBAD-TOPO é um *T-extended vector*, isto é, é um vetor linearizado que contém uma desoxitimidina (T) na extremidade 3'. Além disso, o sistema de clonagem utiliza a topoisomerase para catalisar a ligação entre o plasmídeo e o inserto.

A ORF DCRA foi amplificada a partir do pET28a-DCRA com a enzima polimerase *Platinum High Fidelity* utilizando o oligonucleotídeo sense que iniciou a amplificação a partir do códon ATG iniciador. O reverso 5' ctt gtc gtc gtc gtc tat cct gca gag ctt cag cgg3' possibilitou a inserção de uma seqüência reconhecida pela protease enteroquinase (sublinhado) e a retirada do códon de parada.

Para a posterior clonagem no vetor pBAD-TOPO (Invitrogen), era necessário que o inserto possuísse desoxiadenosina (A) na extremidade 3', assim após a purificação do produto de PCR, 2µg desse amplificado foi adicionado a um tubo contendo tampão da Taq polimerase 1X, dATP 1mM, Taq polimerase 0,5 U e incubado por 15 minutos a 72^oC.

A seguir o inserto foi ligado no vetor pBAD-TOPO de acordo com as especificações do fabricante.

A mistura de ligação foi utilizada para a transformação de células *E. coli* TOP10F' (*Kit One Shot Chemical Transformation* – Invitrogen) de acordo com as instruções do fabricante.

Os recombinantes foram analisados através da clivagem com enzima de restrição e a junção (inserto:vetor) e a integridade do gene foram confirmadas pelo seqüenciamento (seção 3.6).

A análise da indução da expressão da proteína nas células de *E. coli* LMG (194) em fusão com HIS-Tag foi testada em diversas concentrações de L-arabinose de acordo com as instruções do fabricante e visualizada em SDS-PAGE 15% (LAEMMLI, 1970).

3.3.8 Construção do vetor pCX TOPO-DCRA

O vetor pCX-TOPO (Invitrogen) possibilita a clonagem de produtos de PCR para produzir uma fusão com o N-terminal da proteína RsaA. A RsaA participa da formação da parede celular da bactéria *Caulobacter crescentus* e direciona a secreção e agregação da proteína alvo para o meio de cultura da bactéria.

O pCX TOPO é um plasmídeo linearizado que contém uma desoxitimidina (T) na extremidade 3', por isso todos os passos de preparação do inserto para a posterior subclonagem nesse vetor foram realizados como descrito no item 3.3.7. Além disso, contém o promotor *lac* que permite a expressão constitutiva do gene de interesse e em altos níveis.

O inserto foi ligado no vetor pCX-TOPO de acordo com as especificações do fabricante.

A indução da expressão protéica na linhagem *Caulobacter crescentus* foi realizada de acordo com as especificações da Invitrogen. A análise da indução foi feita em SDS-PAGE 10%.

3.3.9 Construção dos vetores pCRT7/CT TOPO-DCRA e pCRT7/NT TOPO-DCRA

Os vetores pCRT7/CT TOPO e pCRT7/NT TOPO (Invitrogen) possibilitam a clonagem de produtos de PCR, respectivamente, produzindo a proteína de interesse em fusão em fase com o C ou N-terminal de seis resíduos de histidina.

O pCRT7-TOPO é um vetor linearizado que contém uma desoxitimidina (T) na extremidade 3' por isso todos os passos de preparação do inserto para a posterior subclonagem nesse vetor foram realizados como descrito no item 3.3.7.

O inserto foi ligado nos vetores de acordo com as especificações do fabricante e a mistura de ligação foi utilizada para a transformação das células de *E. coli* TOP10F' (*Kit One Shot Chemical Transformation* – Invitrogen) de acordo com as instruções do fabricante.

A indução da expressão da proteína em células de *E. coli* BL21 (DE3) foi feita com diferentes concentrações IPTG e visualizada em SDS-PAGE 12%.

3.4 Subclonagem da ORF DCRA no vetor de expressão para células de mamífero

3.4.1 Construção do vetor pcDNA3-DCRA

O vetor pcDNA3 (Invitrogen) é um vetor para expressão em células de mamífero, mas não apresenta nenhum tipo de fusão que possibilita a purificação em cromatografia de afinidade. Assim, optou-se por subclonar no pcDNA3 o gene DCRA ligado a poli-histidina. Para isso o DCRA foi subclonado a partir do vetor pET28a com as enzimas *Xba* I preenchido com dNTPs por atividade da *Klenow* polimerase e posteriormente clivado com *Not* I, o que permitiu a retirada do DCRA em fusão com uma cauda de histidina no N-terminal.

O vetor pcDNA3 foi linearizado com as enzimas *Eco* RV e *Not* I, recuperado do gel e ligado ao His-tag-DCRA.

Em seguida plasmídeo e inserto foram ligados durante a noite a 16⁰C. A mistura de ligação foi utilizada para transformar *E. coli* DH5 α cálcio competente.

Após a validação da integridade das junções por seqüenciamento, o DNA plasmidial purificado foi utilizado para transfectar células de mamífero.

3.5 Hospedeiros usados

Epicurian coli BL21(DE3) – Codon PlusTM (Stratagene) engenheirada para conter cópias extras de genes que codificam tRNAs (que reconhecem códons para arginina, isoleucina e leucina) que mais freqüentemente limitam a tradução de proteínas heterólogas em *E. coli*. A disponibilidade desses tRNAs permite alto nível de expressão de muitos genes heterólogos recombinantes que são fracamente expressos em linhagens convencionais de BL21. Essa linhagem foi cultivada em meio LB (Luria-Bertani) contendo 25 μ g/mL de cloranfenicol a 37⁰C.

Rosetta (DE3) (Novagen) linhagem derivada de TurnerTM (*lac*Y1) permite baixos níveis de expressão em todas as células possibilitando o aumento da solubilidade e da estabilidade. Permite a expressão de proteínas que contém códons raramente usados em *E. coli* (AGG, AGA, AUA, CUA, CCC, GGA). Essa bactéria foi gentilmente cedida pela Prof^a. Dra. Ana Paula U. Araújo (IFSC) e mantida em LB contendo 25 μ g/mL de cloranfenicol a 37⁰C.

BL21-SI (*Salt inducible*) linhagem derivada de GJ1158 e induzível por cloreto de sódio. A bactéria BL21-SI contém o gene da RNA polimerase T7 sob o controle do promotor *proU*. Tem sido proposto que a indução por choque osmótico aumenta a solubilidade da proteína expressa (BHANDARI & GOWRISHANKAR, 1997). Essa linhagem foi gentilmente cedida pelo Prof. Dr. Otávio H. Thiemann (IFSC) e mantida em meio LB sem NaCl e sem antibiótico, a 37⁰C.

BL21 pTgroE essa linhagem possui o vetor de expressão com o fragmento codante da groESL sob o controle do promotor T7 (YASUKAWA et al., 1995). GroESL codifica uma chaperonina de choque térmico cuja função é catalisar o correto enovelamento de um polipeptídeo recém sintetizado. Essa bactéria foi gentilmente cedida pela Prof^a. Dra. Ana Paula U. Araújo (IFSC) e cultivada em meio LB, contendo 25 µg/mL de cloranfenicol, a 37⁰C.

AD494(DE3) (Novagen) é um mutante tiorredoxina redutase negativa (*trxB*⁻) e permite que ligações dissulfeto sejam formadas no citoplasma (pET Manual System, Novagen, 2000). Portanto, proteína alvo que necessita da formação de ligação dissulfeto poderia ser mais facilmente expressa nesse hospedeiro. Essa linhagem foi cultivada em meio LB, contendo 25 µg/mL de canamicina, a 37⁰C.

Caulobacter crescentus (linhagem B5 BAC) (Invitrogen) é derivada da linhagem *Caulobacter* JS4000 que contém uma mutação que impede a expressão da proteína RsaA nativa. É uma bactéria gram-negativa, não patogênica comum em ambiente aquático. Essa linhagem foi cultivada em meio PYE (Seção Anexos), sem antibiótico, a 30⁰C para propagação e manutenção e em meio M11 (Seção Anexos) para expressão.

3.6 Análise dos recombinantes

As colônias transformantes obtidas em cada uma das construções foram agrupadas respectivamente de 10 em 10 e submetidas ao PCR de colônia com *primers* que flanqueavam o vetor e o inserto. Os grupos com resultados de amplificação positivos foram selecionados para extração do DNA plasmidial e posteriormente confirmados através da análise com enzimas de restrição.

Após a confirmação dos recombinantes os clones selecionados foram submetidos ao seqüenciamento para verificação da integridade do inserto e da junção vetor-inserto.

O seqüenciamento de ambas as fitas de DNA foi realizado pelo método didesoxi (SANGER, NICLEN & COULSON, 1977) em um seqüenciador automático (modelo ABI Prism 377 DNA Sequencer) utilizando-se primers que flanqueiam o inserto e/ou vetor. Todos os produtos do seqüenciamento foram analisados pelo alinhamento no MultiAlin (<http://prodes.toulouse.inra.fr/multalin/multalin.html>) (CORPET,1988).

3.7 Expressão da proteína recombinante

A expressão da proteína em pequena escala foi induzida na fase logarítmica (DO_{600nm} entre 0,5-0,6) do crescimento bacteriano através da adição de concentrações variáveis (0,4 a 1,0 mM) de isopropil- β -D-tiogalactopiranosídeo (IPTG). Esse procedimento foi adotado para todas as construções realizadas nesse trabalho. As construções que resultaram em expressão detectável em SDS-PAGE corado com azul de Comassie foram submetidas a testes para otimização da expressão e aumento da solubilidade, como por exemplo, variação da temperatura (37°C e 20°C) e expressão em *E. coli* de diferentes linhagens.

Uma vez estabelecidas às condições mais favoráveis, a expressão passou a ser realizada em maior escala. Deste passo em diante os testes foram realizados somente para as construções pET32aTiorredoxina-DCRA_{Nco/Bam} e pET 32aTiorredoxina-DCRA-GFP.

As células de AD494 pET32aTiorredoxina-DCRA_{Nco/Bam} foram crescidas em LB líquido (100mL) acrescido de canamicina (25 μ g/mL) e ampicilina (100 μ g/mL), induzidas com 0,4 mM de IPTG e mantidas a 37°C sob agitação de 200 rpm por 4 horas. Após esse período a cultura foi centrifugada a 4°C, 10.000g por 10 minutos. O sobrenadante foi descartado e o precipitado ressuspendido em 12 mL do tampão de sonicação (Tris-HCl 10mM pH 8,0; NaCl 100mM; NaH₂PO₄ 50mM) e submetido a sonicação, em banho de gelo, no sonicador Vibra cellTM (Sonic & Material Inc.) por 8 ciclos de 1 minuto com intervalo de 30 segundos a uma amplitude de 20mA.

As células após a lise foram centrifugadas a 4°C, 10.000g por 10 minutos para a separação da fração solúvel (sobrenadante) e insolúvel (precipitado). As alíquotas (fração solúvel, insolúvel e o controle não induzido) foram analisadas em SDS-PAGE 15%.

As bactérias BL21(DE3) pET32aTiorredoxina-DCRA.GFP foram crescidas em LB líquido (1000mL) acrescido de ampicilina (100µg/mL) e induzidas com 0,1 mM de IPTG e imediatamente após a indução a temperatura foi diminuída para 20°C, mantendo a agitação de 200 rpm por um período de 8 horas. Após esse período a suspensão bacteriana foi centrifugada a 4°C, 10.000g por 10 minutos. O sobrenadante foi descartado e o precipitado ressuspenso em 50 mL do tampão de sonicação (Tris-HCl 50mM, pH 8,0; NaCl 30mM). As células bacterianas foram lisadas por sonicação como descrito acima e após a lise foram centrifugadas a 4°C, 10.000g por 10 minutos. As alíquotas (fração solúvel, insolúvel e o controle não induzido) foram analisadas em SDS-PAGE 15%.

3.8 Purificação da proteína

Uma proteína deve ser purificada para que uma caracterização significativa de sua composição química, estrutura e função possa ser realizada. A purificação de uma dada proteína só é conseguida, em geral, pela combinação de vários processos que também podem ser parcialmente utilizados para a caracterização da proteína. A primeira tarefa na purificação de uma proteína consiste em isolar, concentrar e estabilizar a proteína alvo. O próximo passo consiste em remover a maioria das impurezas, tais como outras proteínas e ácidos nucleicos. A purificação final (polimento) incluirá uma combinação de técnicas que separem, com base em carga líquida, tamanho ou afinidade a proteína alvo das demais.

O ponto mais importante da purificação é a preservação da integridade da estrutura nativa da proteína, principalmente quando esta será utilizada para estudo estrutural ou funcional. Assim, Tiorredoxina-DCRA_{Nco/Bam} foi lavada e solubilizada respectivamente nos tampões (Tabela 1) descrito por LINDWALL et al. (2000). A seguir as células foram lisadas por sonicação e as alíquotas correspondentes às frações solúveis e insolúveis foram analisadas em SDS-PAGE 15%.

Tabela 1. Composição dos tampões utilizados para solubilização de proteínas superexpressas de acordo com LINDWALL et al. (2000).

Tampões	Composição
1	Tris 100mM, glicerol 10%, pH 7,6
2	Tris 100mM, LiCl 50mM, pH 7,6
3	HEPES 10mM, (NH ₄) ₂ SO ₄ 50Mm, glicerol 10%, pH 7,0
4	HEPES 10mM, KCl 100 mM, pH 7,0
5	Tris 100mM, NaCl 50mM, isopropanol 10%, pH 8,2
6	K ₂ HPO ₄ /KH ₂ PO ₄ 100Mm, (NH ₄) ₂ SO ₄ 50Mm, Triton X-100 1%, pH 6,0
7	HEPES 10mM, MgSO ₄ 1M, pH 7,0
8	K ₂ HPO ₄ /KH ₂ PO ₄ 100Mm, ZnCl ₂ 2,5 mM, pH 4,3
9	Tris 100mM, KCl 100 mM, EDTA 2mM, Triton X-100 1%, pH 8,2
10	Acetato de sódio 100mM, MgSO ₄ 1M, pH 5,5
11	Tris 100mM, (NH ₄) ₂ SO ₄ 1M, DTT 10 mM, pH 8,2
12	Tris 100mM, NaCl 50mM, uréia 100mM, pH 8,2
13	HEPES 10mM, LiCl 50 mM, deoxicolato 0,1%, pH 7,0
14	Tris 100mM, KCl 100 mM, deoxicolato 0,1%, glicerol 25%, pH 7,6

Os procedimentos utilizados para a purificação da proteína Trx-DCRA-GFP consistiram em fracionamento com sulfato de amônio, cromatografia de afinidade em resina de níquel, cromatografia de troca iônica (Mono-Q HR5/5) e cromatografia de filtração em gel em coluna Superose 12 HR 10/30.

3.8.1 Fracionamento com sulfato de amônio

O fracionamento com sulfato de amônio foi o primeiro passo utilizado para a eliminação de proteínas contaminantes. Quando sais são adicionados em grandes quantidades às soluções protéicas, as proteínas tendem a agregar-se, um efeito conhecido como *salting-out*. Embora a base dessa precipitação não seja bem entendida a precipitação com sulfato de amônio é um dos procedimentos mais utilizados para purificar proteínas,

pois permite que os precipitados protéicos mantenham sua conformação nativa e possam ser dissolvidos novamente, em geral, sem ocorrer desnaturação.

O sobrenadante obtido, após a lise das células, em tampão de sonicação (Tris-HCl 50mM pH 8,0; NaCl 30mM) foi fracionado pela precipitação, inicial de 10% de sulfato de amônio sólido macerado sob agitação suave por 20 minutos a 4°C. As proteínas precipitadas foram coletadas por centrifugação por 15 minutos a 20.000g.

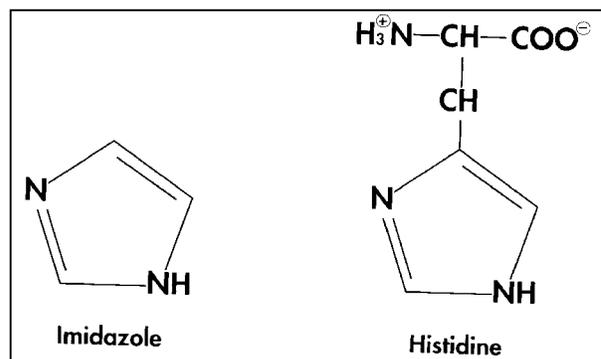
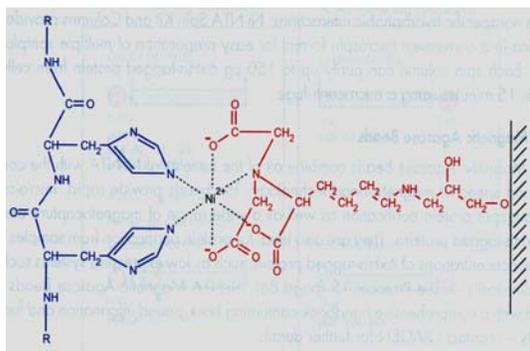
O precipitado obtido foi ressuspendido em 1mL do tampão de sonicação e reservado. Ao sobrenadante remanescente foi adicionado vagarosamente sulfato de amônio sólido até que a concentração final desse sal fosse 20%, sob agitação suave por 20 minutos a 4°C, centrifugado por 15 minutos a 20.000g para coleta do material precipitado e assim sucessivamente até que a concentração final de saturação com sulfato de amônio ficasse em 90%.

Todas as frações obtidas em cada precipitação com sulfato de amônio foram desalinizadas pela diálise contra o tampão de sonicação e analisadas em SDS-PAGE 15%.

A fração correspondente a 30% sulfato de amônio foi submetida à purificação por técnicas cromatográficas.

3.8.2 Cromatografia por Afinidade

A clonagem do DCRA nos vetores da linha pET e TOPO possibilitou a inserção de um resíduo de 6 histidinas. Essa “cauda” de seis histidinas liga-se de forma específica à matriz Ni-NTA (Ácido Nitriloacético carregado com íons níquel) facilitando a purificação.



Esquema mostrando a interação entre os resíduos da cauda da histidina e a matriz Ni-NTA (Qiagen/2001).

Estrutura química do imidazol e da histidina

Os contaminantes (proteínas indesejadas) são retirados quando a coluna é lavada com o tampão e a proteína desejada será eluída alterando-se as condições de pH, ou adicionando-se competidores aos ligantes (ex. imidazol) de forma a liberá-la da matriz. Como o anel imidazólico é parte integrante da estrutura da histidina, o imidazol competirá com a histidina na ligação aos íons níquel imobilizado pelo grupo NTA da matriz permitindo a eluição da proteína.

A proteína tiorredoxina-DCRA-GFP, proveniente de 1L de cultura de bactérias, precipitada pelo sulfato de amônio a 30% foi ressuspendida em 25 mL do tampão de sonicação e dialisada contra o mesmo tampão por 12 horas a 4⁰C em membrana de 14 MW de 28mm (Viskasec), massa de exclusão de 12 a 16 kDa.

A coluna de afinidade contendo a resina Ni-NTA Superflow (Qiagen) previamente empacotada (5 mL de resina), foi lavada com 5 volumes de água e equilibrada com 5 volumes do tampão S2 (Tris-HCl 50mM pH 8,0; NaCl 100mM) após, 25 mL da amostra, previamente filtrada em filtro de 0,22µm, foi aplicada na coluna. Em seguida fez-se uma lavagem com 3 volumes do tampão utilizado para equilibrar a coluna visando à retirada das proteínas que não se ligaram à resina. Para a eluição da proteína o tampão de eluição (Tris-HCl 50mM pH 8,0; NaCl 100mM; imidazol 250 mM) foi diluído no tampão S2 de modo que resultasse em um tampão com concentrações finais de 5, 10, 15, 20, 25, 50, 75, 100, 200 e 250 mM de imidazol. Este novo tampão aplicado na coluna (2 volumes) em ordem crescente de concentração de imidazol possibilitou, portanto, que a proteína fosse eluída em um gradiente linear de imidazol (5 a 250 mM). Todas as frações foram analisadas em SDS-PAGE e as que correspondiam às frações mais puras foram quantificadas pelo método de Bradford (BRADFORD, 1976) e concentradas por centrifugação em Centricon YM-50 (Millipore) até uma concentração final de 7mg/mL.

3.8.3 Cromatografia de Troca iônica

Em pHs superiores ao ponto isoelétrico a carga líquida de uma proteína será negativa, podendo ser adsorvida em grupos carregados positivamente na matriz. Assim, uma matriz insolúvel que tenha ligado grupos carregados positivamente será uma matriz trocadora de ânions, enquanto que uma que tiver grupos carregados negativamente será uma matriz trocadora de cátions. As proteínas que se ligam fortemente ao trocador de íons

são eluídas pela adição de um tampão com elevada concentração de sal ou um valor de pH que reduz a afinidade da proteína pela matriz.

A proteína tiorredoxina-DCRA-GFP com um pI predito igual a 6,0, continuava com contaminantes após a purificação na cromatografia de afinidade optou-se, então como próximo passo cromatográfico, pela coluna de troca aniônica Mono-Q HR5/5 (Amersham Biosciences), operada em sistema HPLC.

A coluna foi previamente lavada com água e equilibrada com 5 volumes do tampão Tris-HCl 25mM, pH 7,5. As frações mais puras (eluição nos gradientes 50, 75 e 100 mM de imidazol), obtidas na cromatografia de afinidade, foram reunidas em uma única amostra. Esta foi dialisada contra o tampão utilizado para equilibrar a Mono-Q, concentrada em centricon 50 para um volume de 500µL e aplicada na coluna. O fluxo utilizado foi de 1 mL/minuto e a eluição foi feita em um gradiente linear de 0 a 500 mM de NaCl em tampão Tris-HCl 25mM, pH 7,5.

3.8.4 Cromatografia de filtração em gel

Na cromatografia de filtração em gel as moléculas são separadas de acordo com o tamanho e a forma. A fase estacionária consiste de malhas contendo poros cujos intervalos de tamanho são relativamente estreitos. Portanto, quando uma solução protéica for passada através dessa “peneira molecular” as moléculas grandes serão eluídas da coluna mais rapidamente que as moléculas pequenas, que passam através dos poros.

A Superose 12 HR 10/30 (Amersham Biosciences) é uma coluna para a realização de cromatografia por exclusão molecular de proteínas e foi escolhida para essa etapa de purificação.

A coluna Superose 12 foi previamente lavada com água e equilibrada com 2 volumes do tampão Tris-HCl 50mM, pH 8,1; NaCl 100 mM e imidazol 75 mM. As frações mais puras (eluição nos gradientes 50, 75 e 100 mM de imidazol) obtidas na cromatografia de afinidade foram concentradas em centricon para 250µL e aplicadas na Superose 12 operada em sistema HPLC. O fluxo utilizado foi de 1,0 mL/minuto e a eluição foi feita com o mesmo tampão. Geralmente a concentração NaCl utilizada no tampão de equilíbrio da Superose é maior do que 100mM, mas neste caso foi necessário manter a força iônica em 100mM para evitar a precipitação da Trx-DCRA-GFP.

3.9 Clivagem com a enteroquinase

Após a purificação foi necessário clivar a proteína quimérica para a retirada das fusões, uma vez que a presença dos parceiros poderia interferir na cristalização. A DCRA poderia ser liberada da fusão tiorredoxina (N-terminal) e GFP (C-terminal) pela clivagem com a protease enteroquinase.

A enteroquinase é uma serino protease que reconhece a seqüência (Asp)₄ Lys e cliva após o resíduo de lisina.

A amostra foi dialisada contra o tampão da Enteroquinase (Novagen) (Tris-HCl 20mM pH 7,4; NaCl 50mM e CaCl₂ 2 mM) e submetida a uma série de reações piloto para determinar a melhor condição de digestão.

A digestão em pequena escala foi realizada com 10 µg e 35µg da proteína alvo, 0,1; 0,2; 0,5; 07 e 1,0 U de enzima enteroquinase em um volume final de 50µL de reação.

As preparações foram incubadas as temperaturas de 20⁰C e 4⁰C e alíquotas de 10 µL foram retiradas após 2, 4, 8 e 16 horas. A reação de digestão foi interrompida pela adição de tampão da amostra contendo β-mercaptoetanol e a extensão da clivagem das amostras foi analisada em SDS-PAGE 15%.

3.10 Espalhamento Dinâmico de Luz (DLS)

A cristalização de macromoléculas biológicas freqüentemente representa uma etapa limitante em análises cristalográficas com Raio X. O sucesso da cristalização depende de vários fatores entre eles uma solução protéica homogênea. A técnica do Espalhamento Dinâmico de Luz pode ser empregada para a seleção da solução protéica monodispersa candidata à cristalização.

A medida de DLS realizada no equipamento “ProteinSolutions DynaPro” forneceu uma série de medidas que foram selecionadas e representadas a partir de histogramas. Os histogramas representam o raio hidrodinâmico médio (R_H) da partícula. A partir dos valores do desvio médio do raio hidrodinâmico pode-se então estimar a massa molecular e o grau de homogeneidade da amostra. Assim, o espalhamento dinâmico de luz foi utilizado para estimar a massa molecular e a homogeneidade da fusão tiorredoxina-DCRA-GFP.

As amostras purificadas na cromatografia de afinidade e na cromatografia de troca iônica foram concentradas em centricon 50 até 3,8 mg/mL. De cada amostra retirou-se uma alíquota de 30µL centrifugou-se por 10 minutos a 10.000g e utilizou-se 20µL do sobrenadante para a leitura de DLS. Esse experimento foi feito pela Dra. Raquel Kelly Bortoleto Rugs do Instituto de Física da USP-São Carlos.

3.11 Ensaio de cristalização

A cristalografia de proteínas tem sido uma poderosa ferramenta para aplicações biotecnológicas, tais como o desenho de drogas, pois permite a determinação molecular da estrutura tridimensional de uma proteína através da realização de experimentos de difração de raios X. A determinação da estrutura tridimensional de uma proteína por difração de raios X tem como etapa crítica a obtenção de um mono cristal de proteína. Um cristal de proteína consiste de moléculas idênticas que se repetem em um arranjo ordenado e periódico. Para a obtenção de um cristal de proteína existem vários métodos, entre eles o método da Difusão de Vapor com gotas suspensas (OLLIS & WHITE 1990), no qual duas soluções tendem ao equilíbrio pela difusão do vapor de água. Na técnica da Gota Suspensa a gota da solução de proteína mais o agente precipitante são gradualmente expostos à saturação pela perda de água da gota para o reservatório que contém a solução precipitante, possibilitando assim o arranjo das moléculas de forma ordenada originando o cristal (BRADEN & TOOZE, 1991) (Figura 3).

Figura 3. Representação esquemática da técnica de cristalização conhecida como gota suspensa. A) Gota sobre a lamínula. B) Modelo de um poço de cristalização.

O método de cristalização adotado foi o de difusão de Vapor com gotas suspensas e os ensaios de cristalização foram realizados utilizando-se os fatoriais *Crystal Screen Reagent Formulation I* (50 condições) e *Crystal Screen Reagent Formulation II* (48 condições) da *Hampton Research* a temperatura de 4^oC.

A gota de cristalização foi montada com 2 µL da solução do poço e 2 µL da solução de proteína (~ 8mg/mL) em tampão de eluição.

Os Ensaios de Cristalização foram realizados em colaboração com o grupo de Cristalografia da IFSC da USP-São Carlos.

3.12 Análise da seqüência de aminoácidos *in silico* usando o pacote *ExpPASy* para similaridade, modificações pós-traducionais, topologia, estrutura primária e secundária.

O ExpPASy (Expert Protein Analysis System) é um servidor online de biologia molecular que possui *links* rápidos para acesso a vários *sites* dedicados as análises de identificação e caracterização da proteína, estudos de similaridade, predição de estrutura primária, secundária, topologia, etc.

A seqüência de aminoácidos prevista a partir do produto do gene DCRA foi analisada *in silico* no *ExpPASy site* (<http://ca.expasy.org/>) para:

1- Similaridade. A seqüência de aminoácidos da DCRA foi alinhada contra seqüências não redundantes depositadas no banco de dados do Centro Nacional de Informação para Biotecnologia <http://www.ncbi.nlm.nih.gov/BLAST/>

2- Modificações pós-traducionais no programa NetPhos que faz a predição de resíduos de serina, treonina e tirosina fosforilados. NetPhos é um método de rede neural que prediz sítios de fosforilação em seqüências independentes com uma sensibilidade de 69 a 96% (BLOM et al. 1999).

Disponível no servidor <http://www.cbs.dtu.dk/services/NetPhos/>

3- Topologia nos programas:

1- Psort II que faz a predição de sinalização e sítios de localização baseada na seqüência de aminoácidos do N-terminal da proteína (NAKAI & HORTON 1999).

Disponível no servidor <http://psort.nibb.ac.jp/form2.html>

2- TargetP que faz a predição de localização subcelular. TargetP é um método baseado em rede neural para localização subcelular de proteínas recém identificadas. Utilizando apenas informações da seqüência N-terminal discrimina proteínas endereçadas para mitocôndria, cloroplasto, vias secretórias e “outras” localizações com 90% de confiança para proteínas que não pertencem às plantas (OLOF et al., 2000).

Disponível no servidor www.cbs.dtu.dk/services/TargetP/

4- Estrutura primária nos programas:

1- ProtParam que é uma ferramenta que permite a análise computacional de vários parâmetros físicos e químicos para uma dada proteína depositada no [Swiss-Prot or TrEMBL](#) ou para uma seqüência de aminoácidos enviada ao ProtParam. <http://us.expasy.org/tools/protparam.html>

2- PESTfind é um programa de computador desenvolvido por RECHSTEINER & ROGER (1996) para identificação de regiões PEST. Sequências polipeptídicas ricas em prolina (P), ácido glutâmico (E), serina (S) e treonina (T) são sinais proteolíticos presentes em proteínas eucariontes.

Disponível no servidor <http://www.at.embnet.org/embnet/tools/bio/PESTfind/>

5- Estrutura secundária nos programas

1- *Predict Protein EMBL* é um método de predição de estrutura secundária desenvolvido por ROST & SANDER (1993) que combina informação de seqüências homólogas e rede neural. Esse procedimento permite a identificação de α -hélice e folhas β com 70% de acuracidade. A predição de estrutura secundária foi baseada na seqüência primária de aminoácidos da proteína DCRA enviada ao *Predict Potein EMBL*. http://www.public.iastate.edu/%7Epedro/pprotein_query.html

2- PSIPRED que é um método de predição de estrutura secundária baseado em rede neural e em análises obtidas de PSI-BLAST (Position Specific Iterated – BLAST) McGUFFIN; BRYSON & JONES (2000). A predição de estrutura secundária foi baseada na seqüência primária de aminoácidos da proteína DCRA enviada ao PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred/psiform.html>)

3.13 Produção de anticorpos policlonais contra a DCRA humana em camundongos.

Para confirmar a localização subcelular da DCRA e detectá-la nos mais diversos tecidos é necessário um ensaio que seja específico à proteína de interesse e altamente

sensível mesmo que a expressão seja baixa frente ao conteúdo protéico total. Assim optou-se pela produção de anticorpos policlonais contra a referida proteína.

3.13.1 Preparação das amostras para imunização

Aproximadamente 500 µL das frações da proteína recombinante tiorredoxina-DCRA-GFP semi purificada (frações saída da Mono-Q) foram aplicadas em gel preparativo SDS-PAGE 15% (LAEMMLI, 1970) e submetido a eletroforese por 2 horas a 120V. Após, uma pequena tira vertical foi retirada do gel corada com azul de Coomassie e utilizada como referência para o isolamento da banda, contendo a proteína, presente no gel não corado. Este fragmento de gel contendo a maior quantidade de proteína foi homogenizado em 0,5 mL de água em um homogenizador tipo Potter e incubado a 4^oC durante uma noite para difusão da amostra. A seguir, o sobrenadante foi coletado, o conteúdo protéico foi quantificado pelo método de Bradford (BRADFORD, 1976) e utilizado nas imunizações.

3.13.2 Imunização dos animais para produção do antissoro

Para a primeira imunização foram utilizados 50 µg da proteína recombinante tiorredoxina-DCRA-GFP recuperada do gel SDS-PAGE como descrito anteriormente. A proteína foi diluída em 1 volume do adjuvante Freud completo (Sigma) e administrada de forma subcutânea em oito camundongos machos.

Após quatro semanas da primeira imunização os animais receberam a segunda dose do antígeno (10 µg), isolada de um gel SDS-PAGE e misturada a igual volume de adjuvante Freud incompleto (Sigma). Após dez dias retirou-se uma pequena alíquota de sangue desses animais para testar o antissoro através de *Western Blot*. Posteriormente a verificação sinal/ruído, os animais foram sangrados e o sangue coletado foi centrifugado a 4000 rpm, por 15 minutos a 4^oC para a obtenção do antissoro.

Para diminuir interferências de ligações inespecíficas com proteínas de *E. coli* uma alíquota do antissoro foi incubada com lisado de *E. coli* de acordo com SAMBROOK & RUSSEL (2001) e armazenado a 8^oC. O antissoro remanescente foi armazenado a -20^oC.

3.13.3 Imunodeteção

Para a extração da proteína total das células de mamífero, as células foram centrifugadas para a retirada do meio de cultura e ressuspendidas em 1 mL do tampão Triton 1X (Triton X-100 1%, Tris-HCl (pH 8,0) 100mM, NaCl 100 mM e EDTA 10mM acrescidos de PMSF 2 mM, pepstatina 1µM, leupeptina 20 µM e DTT 5 mM.). As células foram lisadas no sonicador Vibra Cell™ (Sonic & Material Inc.) por 30 segundos (20 mA) e centrifugadas a 16.000g por 15 minutos. O sobrenadante foi transferido para um tubo novo e o conteúdo protéico foi quantificado pelo método de Bradford.

Para testar a produção do anticorpo policlonal anti-tiorredoxina-DCRA-GFP as amostras de proteína total de célula de mamífero (aproximadamente 50 µg) e proteínas GFP, tiorredoxina e tiorredoxina-DCRA recombinantes, foram submetidas à SDS-PAGE 15%.

Após a eletroforese, as proteínas foram transferidas para uma membrana de nitrocelulose por aproximadamente duas horas e trinta minutos a 150 V, como sugerido por AUSUBEL et al. (1994).

Terminada a transferência, a membrana foi corada com solução do *Ponceau* (Ponceau a 0,5% em ácido acético 0,1%), por no máximo 5 minutos e descorada com ácido acético a 0,1%. A membrana foi posteriormente incubada com agente bloqueador (2,5g leite em pó desnatado dissolvido em 50 mL de TBS 1X) a 8°C durante 12 horas e lavada 4 vezes com TBS 1X (Tris 70mM e NaCl 100 mM) com duração de 5 minutos cada lavagem.

A membrana foi incubada com o primeiro antissoro de camundongo anti-tiorredoxina-DCRA-GFP (1:250 em TBS [1X] acrescido de Tween-20 0,1%) por duas horas e trinta minutos e lavada novamente com TBS 1X. O soro pré-imune foi utilizado como controle negativo.

Após a membrana foi incubada com o segundo anticorpo anti-IgG de camundongo, obtido em coelho, marcado com fosfatase alcalina (Sigma), diluído 1:10.000 em TBS [1X], Tween-20 0,1% por duas horas e trinta minutos e enxaguadas várias vezes em TBS 1X.

A seguir a membrana foi revelada com substrato cromogênico (PA conjugate substrate – Bio-Rad) até o aparecimento das bandas. A reação de revelação foi bloqueada através da lavagem da membrana com água.

3.14 Localização subcelular em células de mamífero

3.14.1 Construção do plasmídeo pEGFPC1-DCRA

O vetor pEGFPC-1 (Clontech) permite a expressão da proteína alvo em fusão com a extremidade C-terminal da GFP (*Green Fluorescent Protein*).

O vetor pEGFPC-1 foi tratado com as enzimas de restrição *Bgl* II, preenchido com dNTPs, por atividade da *Klenow* polimerase e clivado com *Bam* HI. Após esses tratamentos o plasmídeo foi isolado por eletroforese em gel de agarose e posteriormente purificado.

A ORF DCRA foi amplificada a partir do vetor pET 28a-DCRA utilizando-se os *primers* sense 5'tc gaa ggt cgt cat atg ggg acc gcc ctg gac at3' (sítio de restrição para *Nde* I sublinhado) e o reverso 5'cgg gat cct ata tcc tgc aga get tca g3' (sítio de restrição para *Bam* HI sublinhado).

A reação em cadeia da polimerase foi realizada num termociclador (PTC-100TM - MJ Research). Utilizou-se *Deep Vent* polimerase (BioLabs), para evitar erros de amplificação. As condições utilizadas para as reações de amplificação seguem as recomendadas pelo fornecedor da enzima e foram adaptadas para as temperaturas de anelamento dos oligonucleotídeos. O ciclo de amplificação foi: 1 ciclo inicial de 94⁰C por 2 minutos (desnaturação); 10 ciclos de 94⁰C por 15 segundos (desnaturação), 53⁰C por 15 segundos (anelamento), 72⁰C por 45 segundos (extensão); 20 ciclos de 94⁰C por 15 segundos (desnaturação), 53⁰C por 15 segundos (anelamento), 72⁰C por 45 segundos (extensão), com adição de 20s/ ciclo e um ciclo final de 72⁰C por 7 minutos (extensão).

Posteriormente à confirmação da amplificação, o produto de PCR foi clivado com *Nde* I preenchido com dNTPs, por atividade da *Klenow* polimerase e clivado com *Bam* HI. Após esse tratamento o fragmento (com uma extremidade cega e outra coesiva) foi isolado em gel de agarose e posteriormente purificado. Em seguida o inserto DCRA e o plasmídeo pEGFPC-1 foram ligados durante a noite a 16⁰C para gerar GFP-DCRA.

A mistura de ligação foi utilizada para transformar *E. coli* DH5 α quimicamente competente por tratamento com cloreto de cálcio.

Os recombinantes foram detectados por PCR de colônia e posteriormente analisados através da clivagem com enzima de restrição. Após a validação da integridade das junções por seqüenciamento o DNA plasmidial purificado foi utilizado para transfecção de células de mamífero usando Lipofectamina 2000.

3.14.2 Construção do plasmídeo pEGFPC1-DCRA-formas truncadas

Para a obtenção das formas truncadas da DCRA quatro plasmídeos pEGFP-DCRA foram linearizados respectivamente com as enzimas *Nco* I, *Hind* III, *Acc* I e *Pvu* II em sítios presentes no gene DCRA, preenchidos com dNTPs, por atividade da *Klenow* polimerase e clivados com *Bgl* II. Após esses tratamentos cada um dos fragmentos, contendo respectivamente, 126, 180, 356 e 638 pb a partir da porção N-terminal foram purificados em gel de agarose.

Em seguida cada fragmento (com uma extremidade cega e outra coesiva) foi clonado num novo plasmídeo pEGFP-DCRA linearizado com *Bam* HI, preenchido com dNTPs por atividade da *Klenow* polimerase e posteriormente clivado com *Bgl* II. A clivagem com *Bam* HI e *Bgl* II remove o inserto DCRA e gera extremidades cega e coesiva que permite a clonagem dos fragmentos da DCRA em fase com a GFP.

As fusões da GFP com os fragmentos de DCRA foram denominadas de: pEGFP-DCRA¹⁻⁴², pEGFP-DCRA¹⁻⁶⁰, pEGFP-DCRA¹⁻¹¹⁸, pEGFP-DCRA¹⁻²¹², pEGFP-DCRA¹⁸⁷⁻²⁹⁷ e pEGFP-DCRA²⁵¹⁻²⁹⁷ (Figura 4 e 5).

DCRA *Nco* I 126 pb (42 aa) denominado de pEGFP-DCRA¹⁻⁴²

pEGFPC1MGTALDIKIKRANKVYHAGEVLSGVVVISSKDSVQHQGVSLT

DCRA *Hind* III 180 pb (60 aa) denominado de pEGFP-DCRA¹⁻⁶⁰

pEGFPC1MGTALDIKIKRANKVYHAGEVLSGVVVISSKDSVQHQGVSLTMEGTVN
LQLSAKSVGVFE

DCRA *Acc* I 356 pb (118 aa) denominado de pEGFP-DCRA¹⁻¹¹⁸

pEGFPC1MGTALDIKIKRANKVYHAGEVLSGVVVISSKDSVQHQGVSLTMEGTVN
LQLSAKSVGVFEAFYNSVKPIQIINSTIEMVKPGKFPSPGKTEIPFEFPLHLKGNKVL
ETYHGVFVNIQY

DCRA *Pvu* II 638 pb (212 aa) denominado de pEGFP-DCRA¹⁻²¹²

pEGFPC1MGTALDIKIKRANKVYHAGEVLSGVVVISSKDSVQHQGVSLTMEGTVN
LQLSAKSVGVFEAFYNSVKPIQIINSTIEMVKPGKFPSPGKTEIPFEFPLHLKGNKVL
ETYHGVFVNIQYTLRCMDKRSLLAKDLTKTCEFIVHSAPQKKGKFTSPVDFVTITPET
LQNVKERALLPKFLLRGHLNSTNC**VITQPLTGELVVESEAAIRSVELQ**

DCRA final (1) 334pb (110 aa) denominado de pEGFP-DCRA¹⁸⁷⁻²⁹⁷

pEGFPCI VITQPLTGELV**VESSEAAIRSV**ELQLV**RVETCGCAEGY**ARDATEIQNIQ
IADGDVCRGLSVPIY**MV**FPRLFTCPTLETTNFKVEFEVNIVVLLHPDHLITENFPLKL
CRI

DCRA final (2) 138pb (45aa) denominado de pEGFP-DCRA²⁵¹⁻²⁹⁷

pEGFPCI FPRLFTCPTLETTNFKVEFEVNIVVLLHPDHLITENFPLKLCRI

Figura 4. Sequência de aminoácidos presente nas construções DCRA-formas truncadas. Y em vermelho destaca um possível sítio de fosforilação. Os aminoácidos destacados em azul mostram a região comum entre as construções pEGFP-DCRA¹⁻²¹² e pEGFP-DCRA¹⁸⁷⁻²⁹⁷.

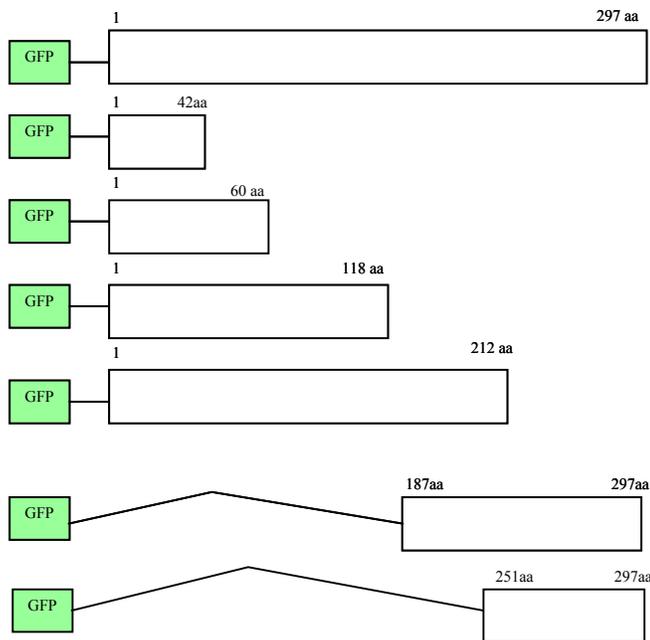


Figura 5. Representação esquemática da clonagem dos fragmentos do DCRA no plasmídeo pEGFP. Os números sobrescritos indicam o número de aminoácidos codificado por cada fragmento.

A construção contendo a parte final (1) da DCRA (pEGFP-DCRA¹⁸⁷⁻²⁹⁷) foi obtida por PCR, utilizando-se o *primer* sense 5'tgt cat cac gca gcc act aac gg3' que flanqueia o gene DCRA na posição 561pb e o *primer* reverso 5'cgg gat cct ata tcc tgc aga gct tca g3'

(sítio de *Bam* HI sublinhado). O produto amplificado foi purificado e clonado no pEGFP linearizado com *Bgl* II, preenchido com *Klenow* e posteriormente clivado com *Bam* HI.

A construção contendo a parte final (2) da DCRA (pEGFP-DCRA²⁵¹⁻²⁹⁷) foi obtida por PCR, utilizando-se o *primer* sense 5'tgt ctt ccc tag gct gtt cac c 3' (acrescido de um t que permite a junção em fase e flanqueia o gene na posição 757pb) e o *primer* reverso 5'cgg gat cct ata tcc tgc aga gct tca g 3' (sítio de *Bam* HI sublinhado). O produto amplificado foi purificado e clonado no pEGFP linearizado com *Bgl* II, preenchido com *Klenow* e posteriormente clivado com *Bam* HI.

Em seguida o gene DCRA e o plasmídeo pEGFPC-1 foram ligados durante a noite a 16°C para gerar as formas truncadas da GFP-DCRA.

A mistura de ligação foi utilizada para transformar *E. coli* DH5α quimicamente competente por tratamento com cloreto de cálcio (AUSUBEL et al., 1994).

Os recombinantes foram detectados por PCR e após a validação da integridade das junções por seqüenciamento, o DNA plasmidial purificado foi utilizado para transfecção de células de mamífero usando Lipofectamina 2000.

3.14.3 Construção do vetor p3XFLAG-*myc*-CMV25-DCRA

O vetor p3XFLAG-*myc*-CMV25 (Sigma) permite a expressão da proteína de interesse em fusão com o epítipo *myc* no C-terminal e epítipo FLAG no N-terminal. Esses epítipos podem ser detectados por anticorpos anti-*myc* conjugado com Cy3 e anti-FLAG conjugado com FITC, respectivamente.

O vetor p3XFLAG-*myc*-CMV25 foi primeiramente clivado com *Xba* I, preenchido com dNTPs por atividade da *Klenow* polimerase e posteriormente clivado com *Bgl* II.

O gene DCRA foi amplificado a partir do pET28a-DCRA com a enzima polimerase *Platinum High Fidelity* (GIBCO) utilizando oligonucleotídeo sense 5'gaa gat ctg atg ggg acc gcc3' (sítio *Bgl* II sublinhado) e reverso 5'cta gtc tag ata tcc tgc aga g 3' (permite a retirada do códon finalizador e insere sítio para *Xba* I sublinhado).

Em seguida o inserto e vetor foram ligados durante a noite a 16°C para gerar FLAG-DCRA-*myc*.

A mistura de ligação foi utilizada para transformar *E. coli* DH5α quimicamente competente.

Após a validação da integridade das junções por seqüenciamento o DNA plasmidial purificado foi transfectado em células de mamífero usando Lipofectamina 2000.

3.14.4 Construção do vetor pEGFPN1-DCRA

O vetor pEGFPN-1 (Clontech) permite a expressão da proteína de interesse em fusão com o N-terminal da GFP. O vetor foi clivado com *Eco* RI preenchido com dNTPs por atividade da *Klenow* polimerase e posteriormente clivado com *Bgl* II.

A ORF DCRA foi isolada do vetor p3XFLAG-*myc*-CMV25 com a enzima *Xba* I, preenchida com dNTPs por atividade da *Klenow* polimerase e posteriormente clivada com *Bgl* II.

Em seguida DCRA e o plasmídeo pEGFPN-1 foram ligados durante a noite a 16⁰C para gerar DCRA-GFP.

A mistura de ligação foi utilizada para transformar *E. coli* DH5 α quimicamente competente.

Após a validação da integridade das junções por seqüenciamento o DNA plasmidial purificado e utilizado para transfectar as células de mamífero.

3.14.5 Construção do plasmídeo pEGFPC1-DSCR8

O vetor pEGFPC-1 foi clivado com as enzimas de restrição *Bgl* II e *Eco* RI.

A ORF DSCR8 foi subclonada a partir do pGEM-T-DSCR8 com as enzimas *Bam* HI e *Eco* RI. Fragmentos de DNA clivados com *Bam* HI e *Bgl* II podem ser ligados, pois as seqüências de bases que codificam os sítios dessas enzimas são compatíveis. A seguir vetor e inserto foram recuperados do gel de agarose como descrito na seção (3.2.1) e ligados durante a noite a 16⁰C para gerar GFP-DSCR8.

Os recombinantes foram confirmados pela análise com enzimas de restrição e o DNA plasmidial purificado foi transfectado em células de mamífero usando Lipofectamina 2000.

3.14.6 Cultura de células e Transfecção Celular

3.14.6.1 Linhagens de células utilizadas:

- Células CHO-K1 (“Chinese Hamster Ovary cells”)
- Células COS-7 (“Monkey Kidney Cells”)
- Células HEK 293 (“Human Embryonic Kidney”)

3.14.6.2 Transfecção Celular

As células inicialmente foram descongeladas e semeadas em garrafas (Corning-Costar) de 75 cm², numa densidade de aproximadamente 10⁴ células/cm² e mantidas em meio de cultura DMEM, contendo soro fetal bovino 10%, glutamina 2mM, penicilina e estreptomicina 100U/μl, até atingir semiconfluência (24 horas).

Quando as células atingiram a semiconfluência desejada foram transfectadas com DNA purificado usando lipofectamina 2000 (Invitrogen).

O procedimento apresentou as seguintes etapas:

1^o Dia:

As células foram desprendidas com uma solução de tripsina a 0,25% e semeadas em placas de 35 mm (Corning-Costar), numa concentração de 10⁵ células/placa e incubadas a 37^oC, com 5% de CO₂ junto a uma lamínula de vidro previamente esterilizada.

2^o Dia:

- a- Após 24 horas aproximadamente 5 μg de DNA foram misturados com meio sem soro (MSS), num volume total de 100μL;
- b- Aproximadamente 5 μL de lipofectamina foram misturadas em MSS, num volume total de 100μL;
- c- As duas soluções (descritas nos itens a e b) foram misturadas suave e cuidadosamente e incubadas à temperatura ambiente por 15 minutos;

- d- O meio das placas foi removido por aspiração e as células foram lavadas com MSS. A mistura das duas soluções acrescida de 800µL de MSS foi adicionada às placas. As placas foram agitadas com movimentos suaves e circulares, incubadas por 6 horas a 37⁰C e 5% de CO₂.
- e- Após seis horas o meio das placas foi suplementado com soro (concentração final 10%) e as células foram incubadas por 24 horas na estufa de CO₂.

3⁰ Dia:

O meio de cultura foi trocado e as células foram incubadas na estufa a 37⁰C com 5% de CO₂ por mais 48 horas.

3.14.7 Análise da localização subcelular da DCRA, DCRA-formas truncadas e DSCR8.

As células transfectadas com pEFGPC1-DSCR8 e com as diferentes formas da DCRA clonada nos plasmídeos pEGFP foram enxaguadas em PBS [1X] e fixadas em paraformaldeído (pH 7,3) a 3,7% por 10 minutos.

Um exemplar de cada representante destas construções foi permeabilizado com Triton X-100 0,2% por 5 minutos e corado com solução de DAPI (4',6-diamidino-2'-phenylindole dihydrochloride) 10mg/mL diluída 1:5000 por 5 minutos a temperatura ambiente para confirmação da integridade do núcleo. A seguir a solução de DAPI foi retirada por aspiração e as células foram enxaguadas em PBS [1X].

As células de ambos os tratamentos foram analisadas em microscopia de epifluorescência convencional (Olympus BX50) e em confocal (*confocal laser scanning microscope*).

As imagens da microscopia de epifluorescência convencional foram capturadas pelo *software* CoolSNAP-pro (Media Cybernetic). Estas análises foram realizadas no laboratório de Genética e Citogenética de Peixes (DGE-UFSCar).

Enquanto que as imagens do “*confocal laser scanning microscope*” (Zeiss Axiovert LSM510) foram obtidas usando a excitação do laser a 488 nm, sendo detectadas no filtro “band pass” a 515-540 nm. Estas análises foram realizadas no Instituto de Ciências Biomédicas (ICB–USP, SP) em colaboração com a prof^a. Dra Gláucia Maria Machado Santeli.

Para o ensaio imuno-citoquímico as células transfectadas com o plasmídeo FLAG-DCRA-*myc* foram permeabilizadas e incubadas por 1 hora com o anticorpo anti-*myc* conjugado com Cy3 (1:50) (Sigma-Aldrich) contra o epítopo *myc* ou anti-FLAG conjugado com FITC (1:100) contra o epítopo FLAG. A seguir as células foram enxaguadas em PBS [1X] e analisadas em microscopia confocal (Zeiss axiovert LSM510).

4 Resultados e Discussão

4.1 Amplificação da ORF DCRA a partir da biblioteca de cDNA de cérebro fetal humano

O produto de amplificação a partir da biblioteca de cDNA de cérebro fetal, delimitado pelos oligonucleotídeos específicos e conservando os códons de iniciação e terminação originais do gene, resultou em um fragmento único de 914 pb, correspondente a ORF do DCRA (894 pb), acrescido de 20 pb que correspondem aos sítios de clivagem para as enzimas *Nde* I na extremidade 5' e *Bam* HI na extremidade 3' do produto amplificado (Figura 6).

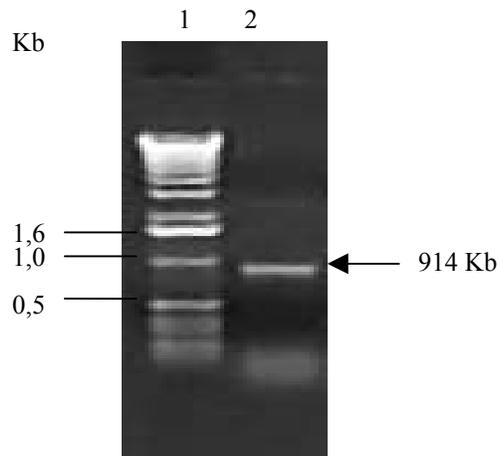


Figura 6. Amplificação da ORF DCRA a partir da biblioteca de cDNA de cérebro fetal humano. Eletroforese em gel de agarose 1% mostrando em (1) marcador de massa molecular e (2) produto de amplificação de 914 pb correspondente ao DCRA.

4.2 Subclonagem e expressão da DCRA no vetor pET28a

Os recombinantes mostraram a liberação de um fragmento de 893pb, quando digeridos com enzimas que flanqueiam o inserto confirmando o sucesso da subclonagem.

A análise da expressão da proteína 6His-tag-DCRA em SDS-PAGE não evidenciou a expressão de uma banda de aproximadamente 35kDa quando comparada à bactéria não induzida (Figura 7).

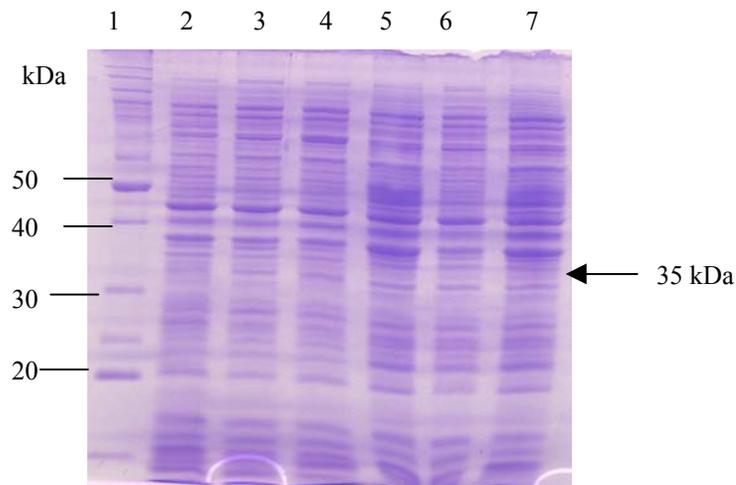


Figura 7. Análise da expressão de recombinantes pET28aDCRA. SDS-PAGE 15% mostrando em (1) Marcador de massa molecular; (2) controle não induzido em BL21 (DE3); (3,4) pET28aDCRA-BL21(DE3) induzido por 3 horas; (5) controle não induzido em BL21 Codon Plus; (6, 7) pET28aDCRA-BL21 Codon Plus induzido por 3 horas. Notar a ausência da banda de indução *His tag*-DCRA de aproximadamente 35 kDa.

A análise do seqüenciamento do plasmídeo pET28a-DCRA mostrou que a junção gene-vetor estava correta, em fase de leitura com a *His tag* e que havia integridade do inserto descartando, portanto, a possibilidade de uma eventual mutação sem sentido ter inviabilizado a expressão.

O baixo nível de expressão de proteínas recombinantes pode ocorrer devido a vários fatores (comentados na seção 4.9). Um dos motivos mais comuns que contribuem para o baixo nível de expressão é a toxicidade ou instabilidade da proteína recombinante ou a perda do plasmídeo durante o crescimento da bactéria. Uma maneira fácil e rápida de verificar se a proteína está sendo tóxica para o hospedeiro é comparar a eficiência de transformação com plasmídeos que possuem o gene cuja expressão é tóxica para o hospedeiro com o plasmídeo controle. Comumente a eficiência observada no primeiro caso é bem menor do que na segunda opção e geralmente os transformantes possuem plasmídeos com mutações.

A ausência de uma banda de indução neste caso muito provavelmente se deu por outras razões, pois não se observou baixa eficiência de transformação com o plasmídeo

pET28aDCRA em relação ao controle, presença de mutações ou morte das bactérias logo após a expressão.

Outro fato que contribui para o baixo nível de expressão é que o RNAm de bactérias não possui qualquer tipo de proteção e é rapidamente degradado durante a transcrição e tradução. Além disso, a estabilidade do RNAm é dependente do tamanho, os mais longos são mais estáveis, e quanto maior a estabilidade maior é a disponibilidade de molde para a tradução e mais eficiente é a expressão protéica (DABROWSKI, BRILLOWSKA & KUR 1999). Assim, uma opção para contornar uma provável instabilidade do transcrito seria selecionar um vetor que possibilitasse a clonagem da DCRA em fusão com um parceiro que pudesse dar maior estabilidade à fusão e que também fosse altamente solúvel.

No entanto, a escolha da melhor proteína parceira ainda é empírica. HAMMARSTRÖM et al. (2002) expressaram em *E. coli* 32 proteínas humanas associadas com desordens genéticas e encontraram profundas diferenças com relação ao nível de expressão e solubilidade dependendo da proteína de fusão usada. Estes autores concluíram que de acordo com seu efeito sobre a solubilidade as fusões poderiam ser classificadas em ordem decrescente: Trx, Gb1/MBP, ZZ (duplo domínio Z da proteína A), NusA, GST e 6 His tag. A His tag não é normalmente usada para melhorar a solubilidade, mas é comumente usada para fins de purificação e por isso HAMMARSTRÖM et al. (op cit) a incluíram nesse estudo. Já com relação ao efeito das fusões sobre a expressão a classificação das fusões em ordem decrescente seria Trx, ZZ, NusA/MBP, GST/Gb1 e 6 His tag. No entanto, estas 32 proteínas possuem o tamanho de 6 a 20 kDa, não apresentam regiões transmembranas e homologia com outras proteínas com estrutura resolvida, deste modo o resultado de HAMMARSTRÖM et al. (op cit.) fica restrito a um grupo de proteínas com características particulares.

Assim os sistemas de expressão pET32a, pET37b, pET41a, pMALc2, pTYB2, pBAD-TOPO, pCX-TOPO que permitem a clonagem em fase com diversos tipos de proteínas de fusão e possibilitam algumas vantagens que facilitam a purificação foram escolhidos uma vez que DCRA em fusão com His tag (pET28a) não trouxe resultado positivo.

4.3 Subclonagem e expressão da DCRA no vetor pET32a

O vetor pET 32a permite a expressão da proteína alvo em fusão com o C-terminal de uma seqüência codificadora de 109 aminoácidos da proteína tiorredoxina (Trx.TagTM). A tiorredoxina (Trx) é um pequeno monômero que facilita a expressão solúvel de fatores de crescimento e de citocinas (LAVALLIE et al., 1993), então essa fusão poderia ser interessante não somente para melhorar a estabilidade do RNAm como também a solubilidade da proteína recombinante, pois muitas proteínas que normalmente são produzidas na forma insolúvel em *E. coli* tendem a se tornarem mais solúveis quando expressas em fusão com a tiorredoxina.

A ORF do gene DCRA de 894 pb introduzido no vetor de expressão pET32a nos sítios *Nco* I e *Bam* HI conservou os códons de iniciação e terminação originais do gene e permitiu a clonagem da DCRA em fusão com a tiorredoxina, como descrito anteriormente.

O DNA plasmidial das colônias submetidos a PCR resultaram na amplificação de um fragmento de aproximadamente 894 pb, confirmando o sucesso da subclonagem. A análise do seqüenciamento confirmou a integridade do inserto, junção do inserto-vetor e descartou a possibilidade de mutações, no entanto, não foi possível detectar em SDS-PAGE a presença de uma banda de indução no tamanho correspondente a 50 kDa (Figura 8). De acordo com WANG, et al. (2003) alguns genes são preditos codificar proteínas que são tóxicas para a *E. coli*, como as proteínas que possuem domínios transmembrana outros não expressam por alguma razão que não é óbvia.

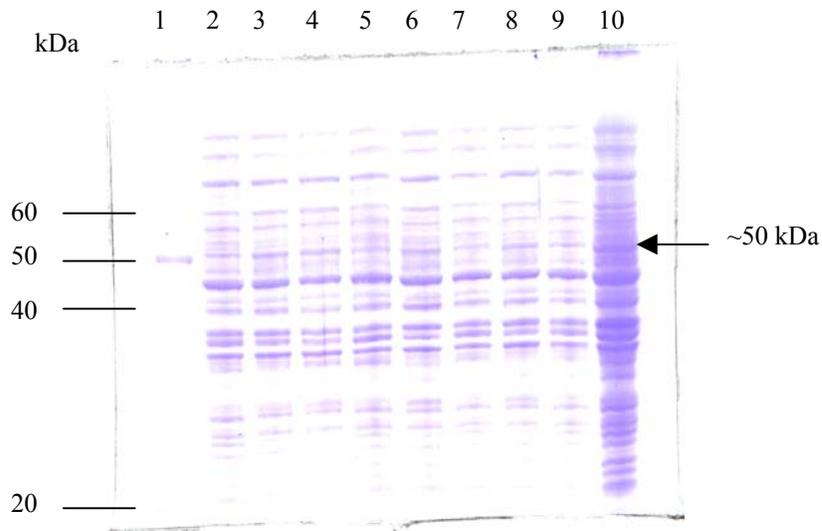


Figura 8. Análise da expressão de recombinantes pET32aTrx-DCRA. SDS-PAGE 15% mostrando em (1) Marcador de massa molecular; (2) controle não induzido em AD494(DE3); (3,4,5,6,7,8,9,10) pET32aTrx-DCRA induzido (0,4 mM IPTG, DO_{600nm} 0,5) respectivamente por 1,2,3,4,5,6,7 e 8 horas. Notar a ausência da banda de indução Trx-DCRA de aproximadamente 50 kDa.

4.3.1 Subclonagem e expressão da Tiorredoxina-DCRA_{Nco/Bam} no vetor pET32a

Após inúmeras clonagens na tentativa de expressar a DCRA íntegra, acidentalmente foi gerada a construção denominada pET32aTiorredoxina-DCRA_{Nco/Bam}. Essa construção permitiu a retirada de 126 pb iniciais do gene DCRA devido à presença de um sítio de *Nco* I no gene. Embora tenham sido retiradas 126 pb distantes do ATG inicial (sentido 5'-3') a fase de leitura foi preservada, pois a Trx ficou em fusão com outro códon de iniciação presente no DCRA resultando na expressão de uma proteína Trx-DCRA_{Nco/Bam} de aproximadamente 46 kDa (Figura 9).

A análise da seqüência dos 42 aminoácidos do N-terminal da DCRA codificado por esse fragmento de 126 pb revelou um putativo motivo de fosforilação de caseína quinase II e caseína quinase C, mas nenhum outro tipo de mudança pós traducional foi predito para esse polipeptídeo. O significado desses motivos ainda não é conhecido, mas se eles fossem funcionais poderiam explicar o provável motivo pelo qual a ausência dos 126 pb resultou em expressão detectável em SDS-PAGE. Sabe-se que procariotos não possuem a maquinaria enzimática necessária para a maioria dos processamentos pós-traducionais e a

ausência de tais modificações em *E. coli* poderia desestabilizar algumas proteínas eucariontes, particularmente as proteínas grandes (BRAUN et al. 2002).

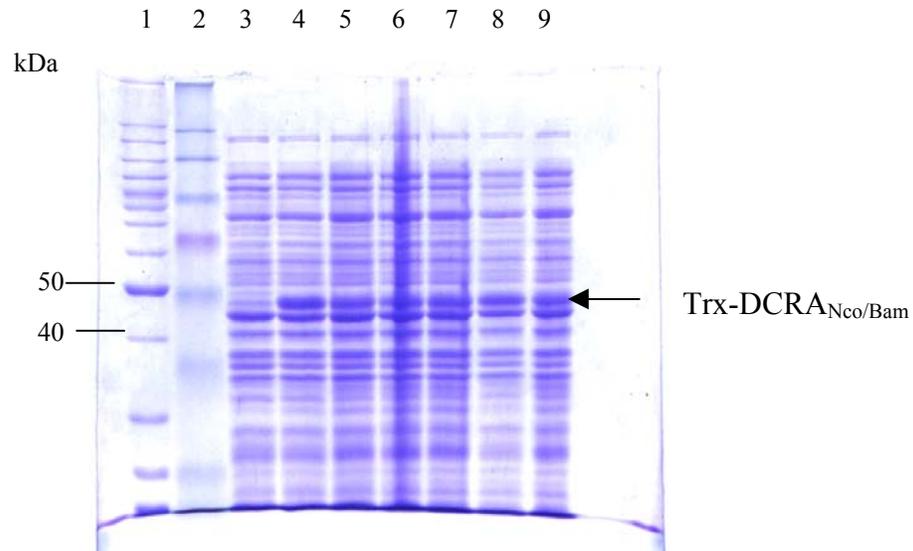


Figura 9. Análise da expressão de recombinantes pET32aTrx-DCRA_{Nco/Bam}. SDS-PAGE 15% mostrando em (1 e 2) Marcador de massa molecular; (3) controle não induzido em AD494; (4, 5 e 6) AD494 induzida respectivamente por 1, 2 e 3 horas; (7, 8 e 9) BL21(SI) pET32aTrx-DCRA_{Nco/Bam} induzida respectivamente por 1, 2 e 3 horas. Notar a presença da banda de indução Trx-DCRA_{Nco/Bam} de aproximadamente 46 kDa.

Após a confirmação da expressão, as células bacterianas foram lisadas por sonicação, centrifugadas e as frações solúvel e insolúvel analisadas em SDS-PAGE, como descrito na metodologia. Os resultados indicaram que a fusão tiorredoxina-DCRA estava no precipitado, o que significava que a fusão estava sendo expressa na forma insolúvel.

A formação de corpos de inclusão (agregado protéico insolúvel) por um lado não é tão desvantajosa, pois, o produto expresso é menos tóxico para o hospedeiro, está em maior parte protegida da digestão proteolítica e até poderia facilitar a purificação, já que os corpos de inclusão são facilmente isolados pela centrifugação e geralmente consistem de agregados parcialmente puros da proteína superexpressa (SAMBROOK & RUSSEL, 2001). Porém, o aspecto negativo é que nem sempre é possível fazer o renovelamento da proteína para sua conformação nativa a partir dos corpos de inclusão e quando os experimentos subsequentes envolvem estudos estruturais ou de função catalítica a melhor opção é tentar obter a expressão na forma solúvel.

Muito freqüentemente proteína heteróloga expressa em *E. coli* acumula-se como corpo de inclusão (YASUKAWA, et al., 1995), por causa disso várias explicações para a formação de corpos de inclusão têm sido relatadas.

Os estudos de RINAS & BAYNES (1993) sugerem que a formação dos corpos de inclusão é (1) devido à limitada quantidade de chaperoninas frente a superexpressão gênica aliada a taxa da síntese protéica que excede a capacidade da célula de enovelar a proteína recém sintetizada (KIEFHABER et al., 1991); (2) à diferença do estado redox entre *E. coli* e a célula eucariota (YASUKAWA et al., 1995); (3) à presença de uma seqüência particular de aminoácidos; e (4) ao fato do enovelamento de proteínas em eucarioto ocorrer simultaneamente com a tradução e o domínio ser formado seqüencialmente enquanto, que em procarioto o enovelamento é pós-traducional (NETZER & HARTL, 1997).

Estas hipóteses poderiam explicar o motivo pelo qual a maioria das proteínas de eucariotos agrega-se quando superexpressa em procariotos. Contudo, o problema do enovelamento protéico não pode ser completamente explicado sem referência ao contexto biológico do enovelamento, especialmente para proteínas grandes com multidomínios e multisubunidades (FEDOROV & BALDWIN, 1997).

Diante de tantas variáveis que contribuem para a formação dos corpos de inclusão, aliadas à dificuldade de renovar a proteína após a purificação em condições desnaturantes, a meta principal na expressão de proteínas recombinantes tem sido o delineamento experimental para a obtenção da solubilidade.

Os quatro principais caminhos seguidos na obtenção de proteínas solúveis são a redução da temperatura de incubação após a indução, a co-expressão de catalisadores de enovelamento como as DPI (*Protein Disulfide Isomerase*) e as chaperoninas, a fusão com um parceiro altamente solúvel e a mutação sítio-dirigida de resíduos específicos para alterar as propriedades de estabilidade e solubilidade da proteína recombinante (MURBY, UHLÉN & STAHL, 1996). Infelizmente essas abordagens nem sempre trazem resultados satisfatórios.

A mutação sítio-dirigida em proteínas tem aumentado a solubilidade principalmente quando afeta seqüências hidrofóbicas, no entanto, a modificação da proteína não é uma estratégia universal para evitar a formação dos corpos de inclusão, já que outras variáveis tais como condições ambientais como temperatura e acidificação do meio (SCHEIN, MATHIEN & NOTEBORN, 1988), característica do vetor e do hospedeiro e a concentração do indutor também estimulam a formação dos corpos de inclusão.

Assim uma das abordagens mais utilizadas e talvez a primeira opção testada para aumentar a solubilidade é diminuir a temperatura de indução visando reduzir a taxa de tradução.

Como os testes iniciais indicavam que a fusão (Trx-DCRA_{Nco/Bam}) estava insolúvel, após a indução da expressão a temperatura foi reduzida para 20⁰C, o meio de cultura LB foi substituído pelo *Circle Grow* e a concentração do IPTG foi diminuída para 0,1mM.

No entanto, as alterações nas condições de cultivo e indução não foram suficientes para a obtenção de proteína solúvel. Como bem lembrado por SAMBROOK & RUSSEL (2001) a ausência de um motivo único ou de um guia prático que resolva todos os problemas torna cada caso de expressão um caso ímpar que deve ser resolvido empiricamente.

A outra abordagem seguida para aumentar a solubilização de proteína superexpressa foi sugerida em LINDWALL et al. (2000). Essa proposta consiste em combinar diferentes valores de pH, sais e aditivos como isopropanol e glicerol em uma matriz semelhante à usada em cristalografia visando encontrar a melhor condição para a solubilização da proteína alvo.

Os tampões sugeridos por Lindwall et al. foram escolhidos de acordo com as disponibilidades de reagentes em nosso laboratório e estão descritos na Tabela 1. Contudo, os tampões testados não foram capazes de solubilizar a fusão Tiorredoxina-DCRA_{Nco/Bam}. Este resultado se deu muito provavelmente porque (1) foi possível testar apenas 14 dos 30 tampões propostos e/ou (2) esses tampões são capazes de desfazerem agregados que resultam da força iônica do tampão usado na extração da proteína ou aqueles que se formam durante a lise da célula, mas não conseguem solubilizar os “verdadeiros” agregados formados durante a síntese protéica.

A co-expressão em BL21 pTgroE que codifica uma chaperonina de choque térmico cuja função é catalisar o enovelamento “correto” de um polipeptídeo recém sintetizado, também não melhorou a solubilidade da Trx-DCRA_{Nco/Bam}. A co-expressão de chaperoninas tem sido um caminho promissor para aumentar a solubilidade e o enovelamento da proteína recombinante, no entanto, o sucesso dessa estratégia parece ser proteína-específica (WALL & PLUCKTHUN 1995).

Para DAVIS et al. (1999) a solubilidade de uma proteína é altamente dependente da sua seqüência de aminoácidos. Assim, a versão modificada do modelo estatístico de Wilkinson e Harrison (1991) para predição de solubilidade da proteína recombinante e

modificada por DAVIS et al. (1999) foi utilizada para determinar o potencial de solubilidade da DCRA.

Esse modelo analisa dois parâmetros, considerados críticos, para a formação do corpo de inclusão que são a carga média dada pela diferença da frequência dos resíduos de Asp mais Glu *versus* Lys mais Arg e a formação de *turn* que é calculada pelo número total dos resíduos de Asn, Gly, Pro e Ser. Segundo esse modelo DCRA possui 83% de probabilidade de ser insolúvel. Fato curioso é que a expressão de GST-Vps26, também resultou em proteína insolúvel (REDDY & SEAMAN, 2001). DCRA possui 21% de identidade e um provável domínio de Vps26 (discutido com mais vagar na seção 9.2).

Segundo o modelo de DAVIS et al. (1999) a tiorredoxina possui uma solubilidade intermediária quando comparada a NusA e MBP, talvez por isso a fusão com a Trx não conseguiu aumentar a solubilidade da proteína alvo.

4.3.2 Subclonagem e expressão da Tiorredoxina-DCRA_{Hind III} no vetor pET32a

A clonagem da DCRA em fusão com a Trx (pET32aTrx-DCRA) não resultou em expressão detectável em SDS-PAGE, no entanto quando a DCRA foi clonada sem os primeiros 126 pb (construção Trx-DCRA_{Nco/Bam}) foi possível detectar que a proteína estava sendo expressa. Com o objetivo de verificar se o fragmento de 126 pb retirado do início do DCRA contribuía de alguma forma para a não expressão gênica, o fragmento contendo os primeiros 180 pb do DCRA foi clonado em fase com a seqüência codificadora do C-terminal da tiorredoxina.

Essa construção denominada Trx-DCRA_{HindIII} possibilitaria que aproximadamente 60 aminoácidos do N-terminal da DCRA fossem expressos em fusão com o C-terminal da tiorredoxina.

O DNA plasmidial das colônias submetidos a PCR resultaram na amplificação de um fragmento de aproximadamente 180 pb, confirmando o sucesso da subclonagem.

A análise do seqüenciamento do plasmídeo pET32a tiorredoxina-DCRA_{HindIII} mostrou que a junção gene-vetor estava correta, em fase de leitura e que havia integridade do inserto descartando, portanto, a possibilidade de uma eventual mutação ter inviabilizado a expressão. No entanto, não foi evidenciada em SDS-PAGE a expressão da Trx-

DCRA_{HindIII} de aproximadamente 26 kDa resultante de 60 aminoácidos do N-terminal DCRA, 122 aminoácidos da Trx e 17 aminoácidos codificados pelo DNA do vetor de expressão, quando comparada a uma bactéria não induzida (Figura 10).

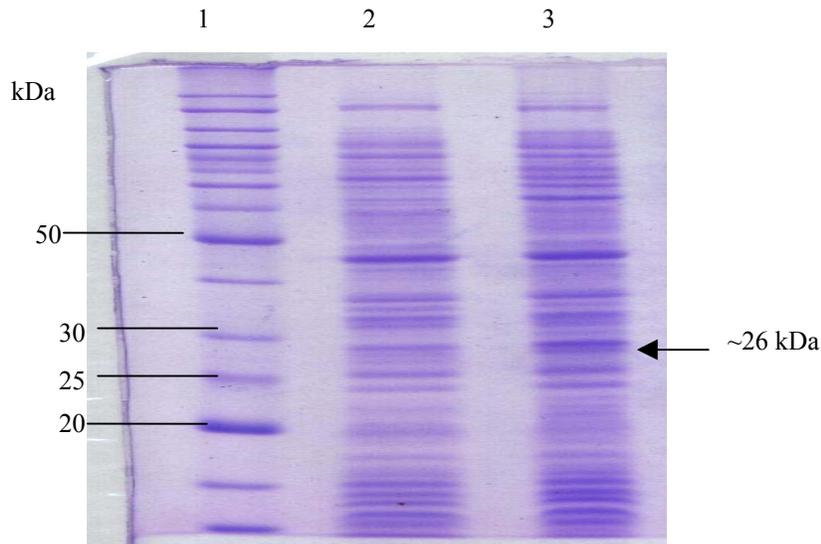


Figura 10. Análise da expressão de recombinantes pET32aTrx-DCRA_{HindIII}. SDS-PAGE 15% mostrando em (1) Marcador de massa molecular; (2) controle não induzido em AD494; (3) colônia pET32aTrx-DCRA_{HindIII} induzida por 4 horas. Notar a ausência da banda de indução Trx-DCRA_{HindIII} de aproximadamente 26 kDa.

4.3.3 Subclonagem e expressão da Tiorredoxina-DCRA-GFP no vetor pET32a

A clonagem da DCRA em fusão com a Trx representada pelas construções pET32aTrx-DCRA e Trx-DCRA_{HindIII} não resultou em expressão detectável em SDS-PAGE, porém foi possível detectar a presença da fusão na forma de corpos de inclusão quando expressa sem o N-terminal (construção Trx-DCRA_{Nco/Bam}).

Como a tiorredoxina possui uma solubilidade intermediária quando comparada a NusA e MBP (DAVIS, et al., 1999) a fusão da DCRA com outra parceira tal como a GFP (*Green fluorescent protein*) poderia melhorar a expressão da DCRA. A GFP tem sido amplamente usada como proteína parceira para monitoramento de expressão, solubilidade e estabilidade da proteína alvo (WU et al., 2002).

O DCRA foi então clonado em fase com a sequência codificadora do C-terminal da tiorredoxina e do N-terminal da GFP. Essa construção denominada Tiorredoxina-DCRA-GFP foi realizada em colaboração com o aluno Fernando Corrêa como descrito em Métodos.

A GFP é uma proteína, codificada por um gene isolado de *Aequorea victoria*, possui um pico de absorção em 395 nm, um pico de emissão de 508 nm e emite fluorescência verde quando irradiada com luz ultravioleta (SCHWALBACH, et al., 2000).

É uma proteína de 27 kDa composta por 238 aminoácidos e contém um cromóforo fluorescente formado por uma ciclização intramolecular de um segmento de um tripeptídeo dentro de uma cadeia monomérica (PRASHER et al., 1992). Dispensa co-fatores, substrato ou produto de outros genes para fluorescer e pode ser expressa como fusão no C ou no N-terminal.

Os recombinantes foram analisados com enzimas de restrição, seqüenciados e submetidos a experimentos de expressão gênica.

A expressão da fusão tiorredoxina-DCRA-GFP a 37°C resultou na formação de agregados na forma de corpos de inclusão, porém a indução com 0,1 mM de IPTG a 20⁰C possibilitou a obtenção de proteína solúvel que pode ser comprovada pela inspeção visual (fração solúvel com coloração esverdeada) e pela análise em SDS-PAGE (Figura 11B). A GFP tem demonstrado ser um sensível indicador de enovelamento, já que o correto enovelamento da GFP e conseqüentemente a formação de seu cromóforo são diretamente proporcionais ao enovelamento da proteína *upstream* a GFP (WALDO, et al., 1999). A banda de indução de aproximadamente 77 kDa (Trx 17 kDa, DCRA 33 kDa e GFP 27 kDa) pode ser visualizada na Figura 11A.

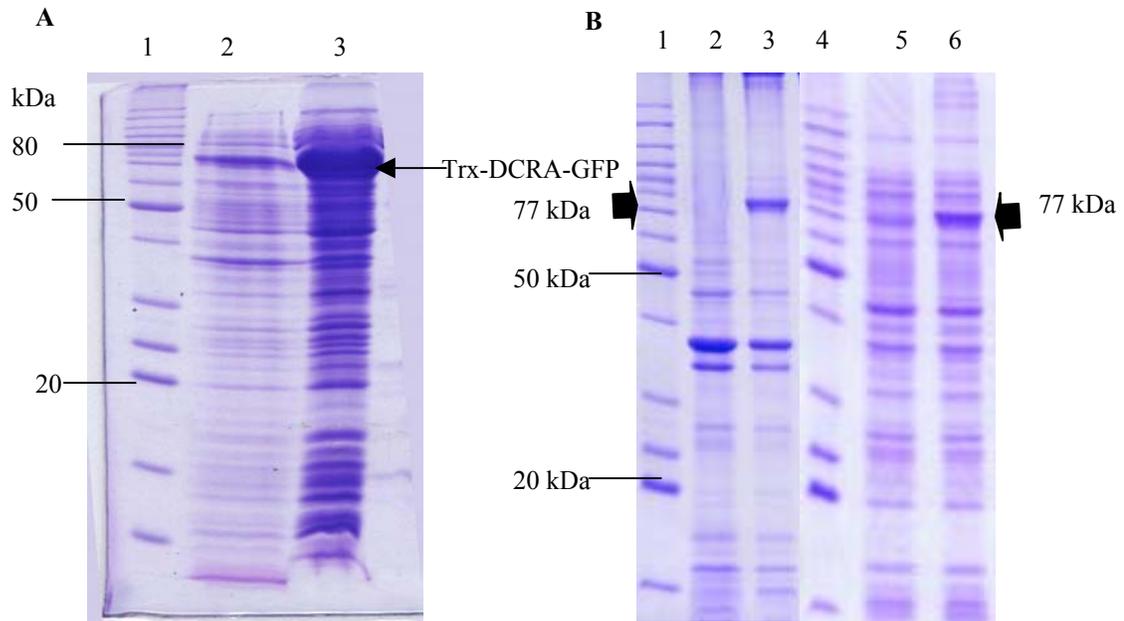


Figura 11. A Análise da expressão de recombinantes pET32aTRX-DCRA-GFP e das frações solúvel e insolúvel. SDS-PAGE 15% mostrando em (1) Marcador de massa molecular; (2) controle não induzido em BL21 (DE3); (3) colônia BL21pET32aTRX-DCRA-GFP induzida por 4 horas. Notar a da banda de indução Trx-DCRA-GFP de aproximadamente 77 kDa. **B.** SDS-PAGE 15% mostrando em (1) Marcador de massa molecular; (2) fração solúvel (37°C); (3) fração insolúvel (37°C); (4) Marcador de massa molecular; (5) fração insolúvel (20°C) e (6) fração solúvel (20°C). As setas indicam a proteína Trx-DCRA-GFP de aproximadamente 77 kDa.

Contudo, apesar das fusões terem aumentado a estabilidade e a solubilidade da DCRA observou-se que após a sonicação havia a formação de agregados protéicos que podiam ser detectados pela inspeção visual. Sabe-se que a estabilidade conformacional da proteína depende fortemente da força iônica e do pH, assim concluiu-se que a composição salina ou o pH do tampão de sonicação (Tris-HCl 50 mM pH 8,0; NaCl 100 mM; NaH₂PO₄ 50mM) precisaria ser modificado.

Após vários testes para definir a melhor condição do tampão chegou-se a conclusão de que a precipitação era causada pela força iônica do tampão usado, assim o tampão passou então a ser Tris-HCl 50 mM pH 8,0; NaCl 30 mM e Imidazol 10mM. O imidazol foi acrescentado somente com o objetivo de melhorar a purificação.

A proteína tiorredoxina-DCRA-GFP foi então purificada e utilizada em todos os ensaios posteriores.

Nossos resultados estão de acordo com a observação de (1) DAVIS et al. (1999), de que o tamanho da proteína parceira poderia aumentar a solubilidade, pois DCRA em fusão somente com a tiorredoxina expressou na forma de corpos de inclusão, mas quando ligada

a tiorredoxina e a GFP foi expressa na forma solúvel e (2) HAMMARBERG et al. (1989) que observaram que apesar das proteínas parceiras trazerem resultados satisfatórios quando fusionadas no N, no C-terminal ou em ambos o resultado mais efetivo tem sido obtido com a estratégia da dupla fusão, isto é, quando o gene que codifica a proteína de interesse é fusionado entre dois genes que codificam dois diferentes domínios de afinidade, pois além da estabilidade *in vivo* a dupla fusão permite a purificação através de dois passos subseqüentes de purificação por afinidade.

4.4 Subclonagem e expressão da DCRA nos vetores pET37b, pET41 e pET43c

A ORF do gene DCRA foi isolada do vetor pET28a com as enzimas *Nde* I e *Bam* HI e subclonado nos vetores pET37b e 41a nos sítios *Nco* I e *Bam* HI. A escolha do sítio *Nco* I se deu porque o gene seria subclonado o mais próximo possível do sítio de clivagem reconhecido pelas proteases Fator Xa e enteroquinase nos vetores pET37b e 41a respectivamente, o que permitiria a presença do menor número de resíduos possível na extremidade N-terminal da DCRA após a retirada da proteína carreadora codificada por esses vetores.

Para tal construção foi necessário criar uma extremidade 5' cega no sítio de *Nco* I dos vetores pET, para isso havia duas possibilidades: (1) tratar com uma nuclease ou (2) preencher com dNTPs utilizando a atividade da *Klenow* polimerase, escolheu-se a primeira, pois era a única possibilidade que permitiria a clonagem da DCRA em fase de leitura com as proteínas de fusão após seu isolamento do pET 28a.

A nuclease escolhida foi a Mung bean. Essa enzima catalisa a degradação específica de nucleotídeos de fita simples possibilitando a formação de extremidade cega gerando um mono ou polinucleotídeo com fosfato 5' terminal (SAMBROOK & RUSSEL, 2001).

O vetor pET37b foi particularmente escolhido porque possui uma seqüência sinalizadora que envia a proteína híbrida para o espaço periplasmático. Segundo JONASSON, et al. (2002) a vantagem da secreção para o espaço periplasmático está na possibilidade da formação de ligações de dissulfeto em um ambiente oxidativo, o que poderia melhorar não somente o enovelamento, mas também a redução dos níveis de contaminação.

Os recombinantes pET37b-DCRA e pET43cNusA-DCRA mostraram a liberação de um fragmento de 894 pb, quando digeridos com enzimas que flanqueiam o inserto confirmando, portanto, a subclonagem.

O recombinante pET41a-DCRA, quando digerido com *Hind* III, mostrou a liberação de um fragmento de aproximadamente 760 pb, devido à presença de um sítio interno de reconhecimento para *Hind* III a 179 pb distantes do ATG do DCRA e outro presente no sítio de múltipla clonagem do pET41a, confirmando o sucesso da subclonagem (Figura 12).

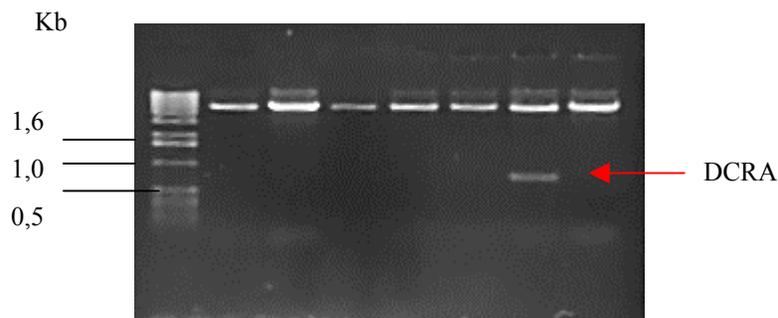


Figura 12. Análise de restrição do pET41aDCRA. Eletroforese em gel de agarose 1% representativa de uma das várias análises de restrição realizadas nesse trabalho. Colunas (1) marcador de massa molecular 1Kb (Gibco); (2,3,4,5,6,7,8) pET41aDCRA *Hind* III. O plasmídeo recombinante apresenta além da banda de 6,9 Kb uma banda de aproximadamente 760 pb.

As análises do seqüenciamento das construções pET37b-DCRA e pET41a-DCRA mostraram que a junção inserto-vetor estava correta e em fase de leitura com as proteínas de fusão desses vetores e que havia integridade do inserto descartando, portanto a possibilidade de uma mutação inviabilizar a expressão.

Entretanto como mostrado respectivamente nas Figuras 13, 14 e 15 não foi possível detectar em SDS-PAGE a presença das fusões CBD-DCRA (46 kDa), GST-DCRA (59 kDa) e NusA-DCRA (88kDa).

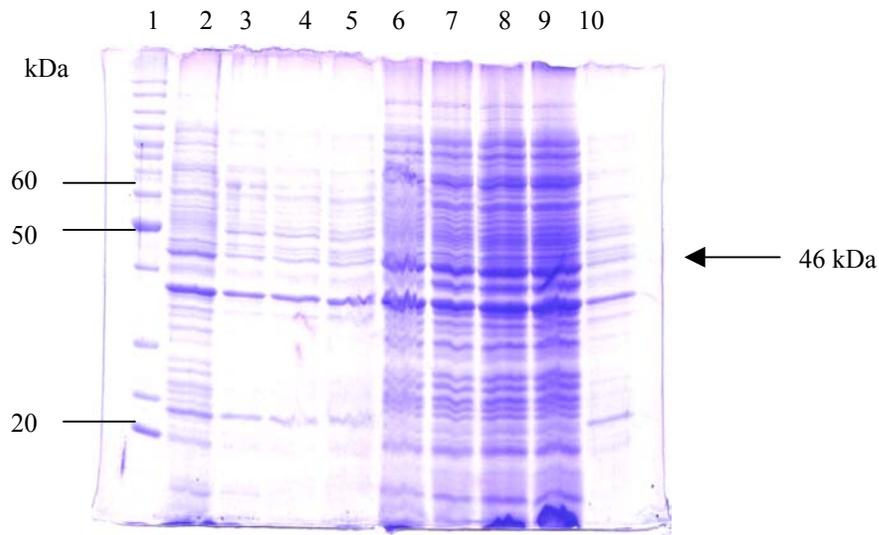


Figura 13. Análise da expressão de recombinantes pET37bCBD-DCRA. SDS-PAGE 15% mostrando em (1) Marcador de massa molecular; (2) controle não induzido (3,4 e 5) colônias pET37bCBD-DCRA em BL21 Codon Plus induzidas por respectivamente 1, 2, 4 horas; (6) controle não induzido (7,8,9 e 10) colônias pET37bCBD-DCRA em BL21(DE3) induzidas por respectivamente por 1, 2, 4 horas. Notar ausência de uma banda de indução CDB-DCRA de aproximadamente 46 kDa.

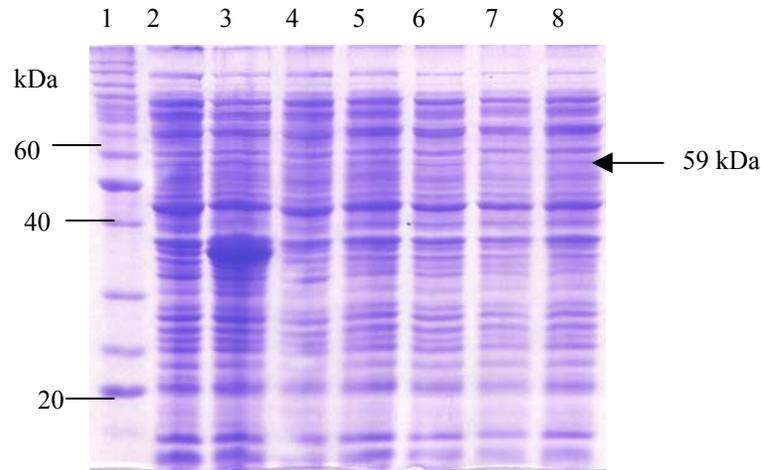


Figura 14. Análise da expressão de recombinantes pET41aGST-DCRA. SDS-PAGE 15% mostrando em (1) Marcador de massa molecular; (2) controle não induzido em BL21(DE3); (3) pET41aGST induzido por 3 horas; (4,5) colônias pET41aGST-DCRA em BL21 Codon Plus induzidas por 3 horas; (6,7 e 8) colônias pET41aGST-DCRA em BL21(DE3) induzidas por 3 horas. Notar ausência de uma banda de indução GST-DCRA de aproximadamente 59 kDa.

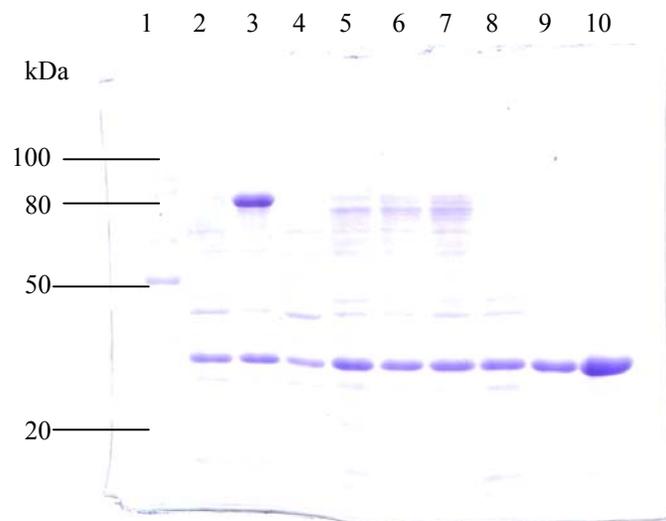


Figura 15. Análise da expressão de recombinantes pET43cNusA-DCRA. SDS-PAGE 12% mostrando em (1) Marcador de massa molecular; (2) controle não induzido (3) pET43cNusA induzido por 3 horas; (4 a 10) colônias pET43cNusA-DCRA em BL21(DE3) induzidas por 1,2,3,4,5,6 e 7 horas respectivamente. Notar ausência de uma banda de indução NusA-DCRA de aproximadamente 88 kDa.

4.5 Subclonagem e expressão da DCRA no vetor pMALc2

A proteína parceira geralmente é escolhida por sua habilidade em se ligar especificamente a um ligante imobilizado facilitando, portanto a purificação da fusão.

A MBP (*Maltose Binding Protein*) é freqüentemente usada como parceira não somente por causa de sua afinidade por maltodextrinas, mas também por aumentar a estabilidade e a solubilidade da proteína alvo (FOX, KAPUST & WAUGH, 2001). Pode ser usada para expressão intracelular bem como em fusão com um sinal para importe para o espaço periplasmático, não contém cisteínas e, portanto a estabilidade da proteína independe de agentes redutores e a superexpressão da proteína em fusão com a MBP pode ser obtida em quantidades que excedem 2% da proteína total de *E. coli* (PRYOR & LEITING, 1997).

Além disso, a MBP apresenta vantagem quando comparada a GST, pois o gene *malE* que codifica a MBP é um gene de *E. coli* e, portanto deve conter códons apropriados e elementos de estrutura secundária de mRNA que são necessários para a eficiente expressão em bactéria (SAMBROOK & RUSSEL, 2001).

Portanto, as vantagens da MBP como proteína parceira fez com que o vetor pMALc fosse particularmente escolhido para a expressão da DCRA.

Assim DCRA foi clonado no pMALc nos sítios de *Eco* RI e *Sal* I para gerar pMalcMBP-DCRA (como descrito na Metodologia).

Os recombinantes, quando digeridos com *Hind* III, mostraram a liberação de um fragmento de aproximadamente 725 pb, devido à presença de um sítio interno de reconhecimento para *Hind* III a 179 pb distantes do ATG do DCRA e outro presente no sítio de múltipla clonagem do pMALc, confirmando o sucesso da subclonagem.

A análise do seqüenciamento do recombinante pMALc-DCRA mostrou que a junção gene-vetor estava correta e em fase de leitura com a proteína MBP, descartando a possibilidade de uma mutação sem sentido inviabilizar a expressão.

Contudo, tentativas de expressar essa construção em *E. coli* BL21(DE3) e em *E. coli* BL21-Codon Plus não foram satisfatórias e conseqüentemente não foi possível detectar em SDS-PAGE uma banda de indução de 78 kDa (Figura 16).

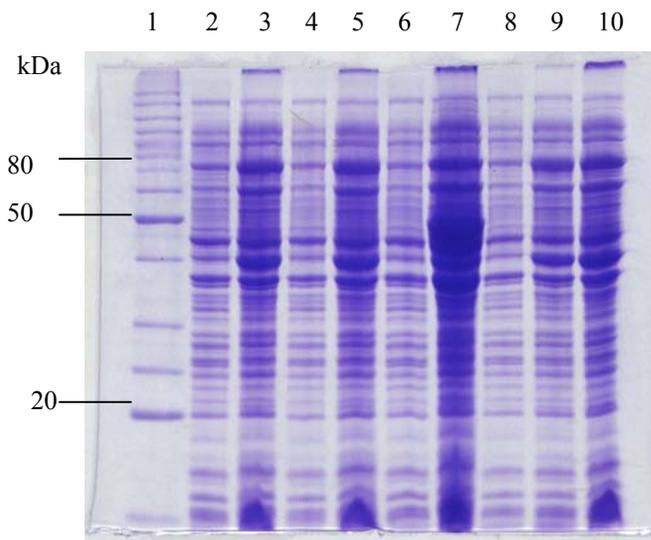


Figura 16. Análise da expressão de recombinantes pMalcMBP-DCRA. SDS-PAGE 15% mostrando em (1) Marcador de massa molecular; (2) controle não induzido em BL21(DE3); em BL21 Codon Plus; (3 a 6) colônias pMalcMBP-DCRA em BL21(DE3) induzidas por 1, 4 e 6 horas; (7 a 10) colônias pMalcMBP-DCRA em BL21Codon Plus induzidas por 1, 4 e 6 horas. Notar ausência de uma banda de indução MBP-DCRA de aproximadamente 78 kDa.

4.6 Subclonagem e expressão da DCRA no vetor pTYB2

O DNA plasmidial de 10 colônias resultantes da transformação *E. coli* DH5 α com o plasmídeo pTYB2-DCRA-inteína foi analisado com a enzima de restrição *Nco* I. Destes, quatro foram linearizados após a clivagem, confirmando o sucesso da subclonagem uma vez que o sítio de reconhecimento desta enzima é codificado somente pelo DCRA.

A análise do seqüenciamento dos recombinantes pTYB2-DCRA-inteína mostrou que a junção inserto-vetor estava correta e em fase de leitura com a inteína descartando, portanto a possibilidade de uma mutação inviabilizar a expressão. Esse vetor foi escolhido porque DCRA seria clonada em fusão com o N-Terminal da proteína parceira e esse fato de algum modo poderia facilitar o enovelamento.

SACHDEV & CHIRGWIN (1998) relatam que a ordem das fusões entre proteínas de bactéria e de mamíferos pode determinar a solubilidade quando expressa em *E. coli* e sugerem que o enovelamento normal da primeira seqüência que emerge do ribossomo determina a via de enovelamento em *E. coli*, contrariando o modelo proposto por NETZER & HARTL (1997) de que o enovelamento em *E. coli* seria pós-traducional. No entanto, a fusão da DCRA no N-terminal da inteína não trouxe resultado positivo e não foi possível detectar em SDS-PAGE a banda de indução (DCRA-inteína) de 88 kDa (Figura 17).

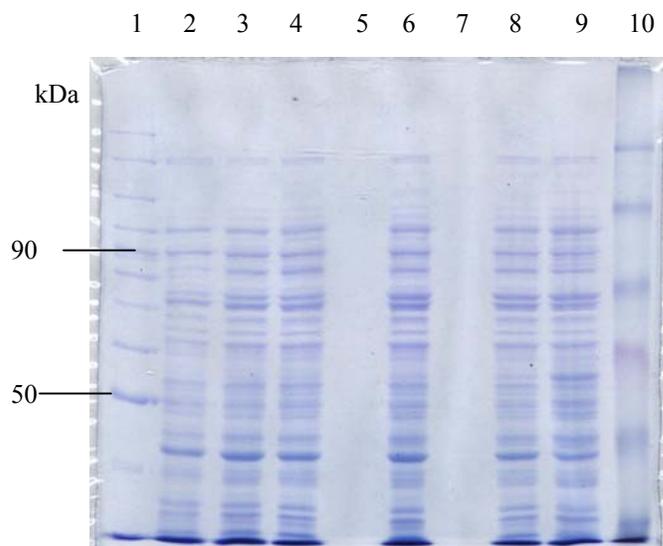


Figura 17. Análise da expressão de recombinantes pTYB2-DCRA-inteína. SDS-PAGE 15% mostrando em (1 e 10) Marcador de massa molecular; (2) controle não induzido em ER2566; (3,4,6,8 e 9) colônias pTYB2-DCRA-inteína induzidas por 1, 2, 3, 4 e 5 horas respectivamente. Notar ausência de uma banda de indução DCRA-inteína de aproximadamente 88 kDa.

4.7 Subclonagem e expressão da DCRA nos vetores PCRT7/CN-TOPO, pBAD-TOPO e pCX-TOPO

Todos os recombinantes de cada construção foram caracterizados por análises com enzimas de restrição, seqüenciados e submetidos a experimentos de expressão gênica. Cada um desses vetores possui uma particularidade como já mencionada na seção 3, no entanto nenhum deles proporcionou a expressão da DCRA (aproximadamente 35 kDa) como observado em SDS-PAGE (Figura 18 e 19).

Além disso, o vetor pCX-TOPO possui uma seqüência sinal que permite secretar a proteína para o meio; assim, se tivesse ocorrido a expressão de DCRA-RsaA (53 kDa) seria detectada pela inspeção visual a presença de aglomerados protéicos de coloração avermelhada no meio de cultura, fato que não ocorreu.

Sabe-se que diferentes tipos de promotores podem afetar o nível de expressão em *E. coli* e a habilidade de promover a alta expressão gênica são governadas por vários critérios (HANNING & MAKRIDES, 1998) por isso optou-se pelo vetor pBad que possui o promotor P_{BAD} induzível por L-arabinose, no entanto a mudança do indutor IPTG para L-arabinose, bem como a utilização dos hospedeiros LMG 194 e *Caulobacter* não influenciaram na expressão da proteína estudada.

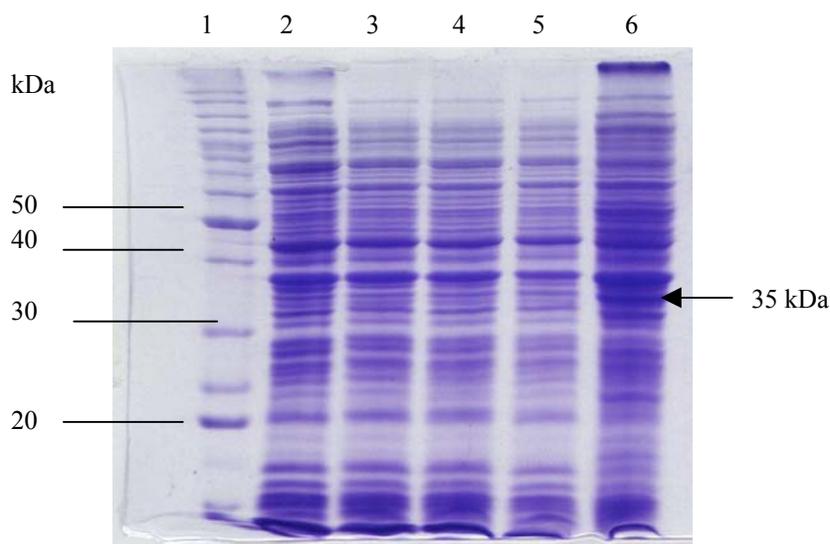


Figura 18. Análise da expressão de recombinantes pCRT7C/N-DCRA. SDS-PAGE 15% mostrando em (1) Marcador de massa molecular; (2) controle não induzido em BL21(DE3); (3 e 4) pCRT7CT-DCRA induzidas por 2 e 4 horas respectivamente; (5 e 6) pCRT7NT-DCRA induzido por 2 e 4 horas respectivamente. Notar ausência de uma banda de indução DCRA de aproximadamente 35 kDa.

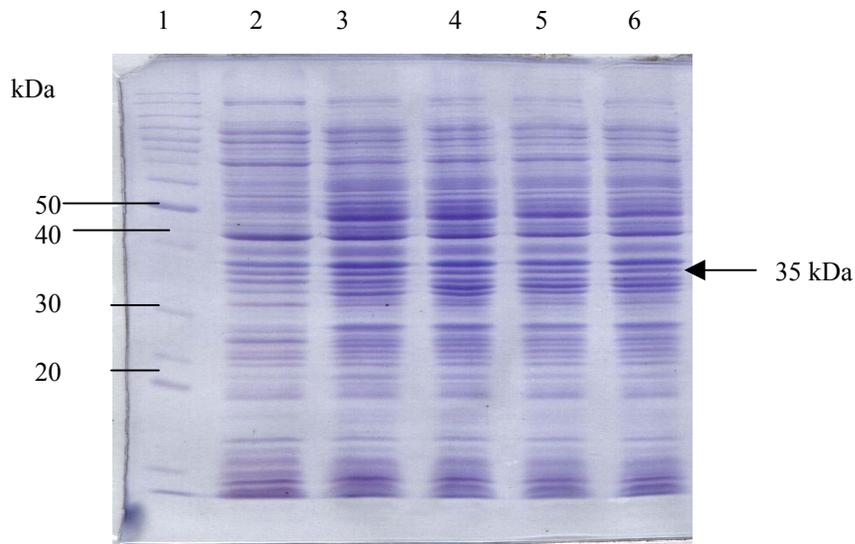


Figura 19. Análise da expressão de recombinantes pBAD-DCRA. SDS-PAGE 15% mostrando em (1) Marcador de massa molecular; (2) controle não induzido em LMG194; (3 a 6) colônias induzidas com diferentes concentrações de L-arabinose. Notar ausência de uma banda de indução DCRA de aproximadamente 35 kDa.

4.8 Subclonagem e expressão da DSCR8 no vetor pET28a

Os recombinantes pET28aDSCR8, quando digeridos com *Nde* I e *Xho* I, mostraram a liberação de um fragmento de aproximadamente 300 pb, confirmando o sucesso da subclonagem.

A expressão da proteína 6His-*tag*-DSCR8 em BL21 (DE3) foi confirmada pela visualização em SDS-PAGE (15%) de uma banda de aproximadamente 12,2 kDa quando comparada à proteínas totais da bactéria não induzida, no entanto os testes de solubilidade indicaram que a proteína recombinante estava na fração insolúvel (Figura 20 A).

A redução da temperatura de indução para 30⁰C, da concentração de IPTG para 0,2mM e a mudança da célula hospedeira para Rosetta (DE3) (seção 3.5) não contribuíram para o aumento da solubilidade da His-*tag*-DSCR8 (Figura 20 B).

Os prováveis motivos que contribuem para a formação dos corpos de inclusão foram comentados na seção 4.3.1.

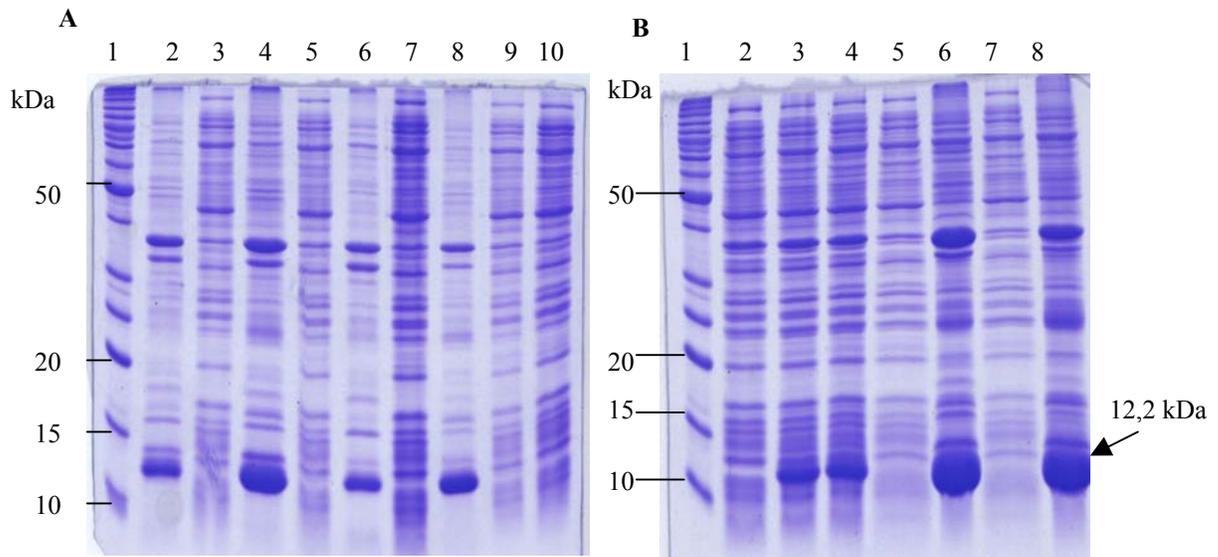


Figura 20. A Análise da expressão de recombinantes pET28aDSCR8 e dos ensaios de solubilidade. A. SDS-PAGE 15% mostrando em (1) Marcador de massa molecular; (2,4,6,8) fração insolúvel; (3,5) fração solúvel; (7, 9, 10) controle não induzido em BL21 (DE3); **B.** SDS-PAGE 15% mostrando em (1) Marcador de massa molecular; (2) controle não induzido; (3) Rosetta (DE3) pET28aDSCR8 induzida 1 hora; (4) Rosetta (DE3) pET28aDSCR8 induzida 3 horas (5) fração solúvel referente a 5 horas de indução (6) fração insolúvel (5 horas), (7) fração solúvel (15 horas) e (8) fração insolúvel (15 horas). A seta indica a proteína His-DSCR8 de aproximadamente 12 kDa.

4.8.1 Subclonagem, expressão e ensaios iniciais de purificação da DSCR8 no vetor pET32a

Os recombinantes pET32aDSCR8, quando digeridos com *Bam* HI e *Eco* RI, mostraram a liberação de um fragmento de aproximadamente 300 pb, confirmando o sucesso da subclonagem.

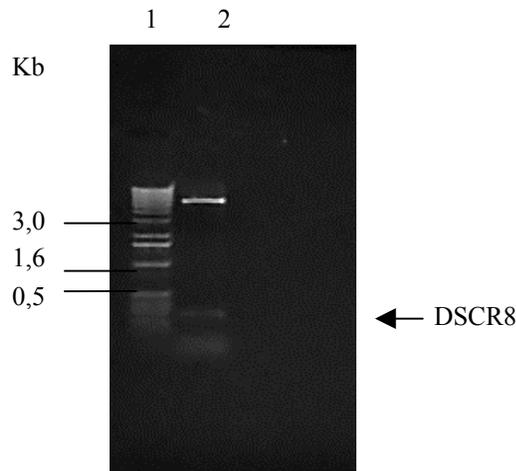


Figura 21. Análise de restrição do pET32aDSCR8. Eletroforese em gel de agarose 1% representativa de uma das várias análises de restrição realizadas nesse trabalho. Colunas (1) marcador de massa molecular 1Kb (Gibco); (2) pET32aDSCR8 clivado com *Bam* HI e *Eco* RI. O plasmídeo recombinante apresenta além da banda de 6,0 Kb uma banda de aproximadamente 300 pb.

A expressão da proteína DSCR8 como uma fusão com o C-terminal da tiorredoxina em Rosetta (DE3) foi confirmada pela visualização em SDS-PAGE (15%) de uma banda de aproximadamente 27 kDa quando comparada à bactéria não induzida (Figura 22).

O teste de solubilidade indicou que a proteína recombinante estava parcialmente solúvel (Figura 22). Esse resultado foi atribuído à fusão da DSCR8 a tiorredoxina, às características da linhagem Rosetta e as condições de incubação logo após a indução, isto é temperatura de indução (30⁰C) e concentração de IPTG (0,4 mM).

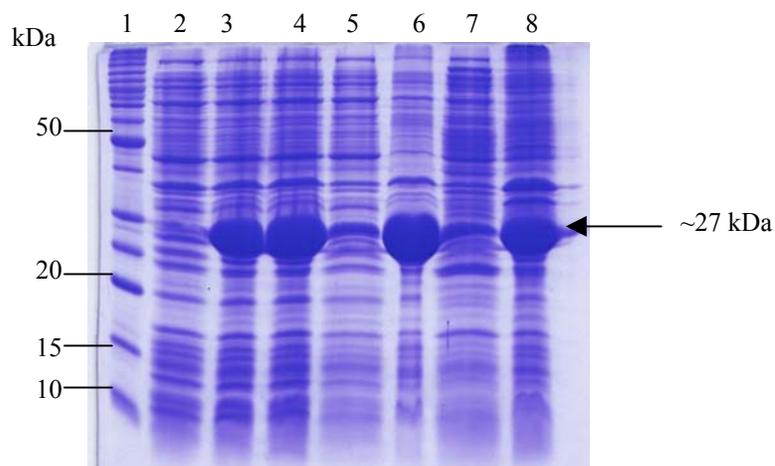


Figura 22. A Análise da expressão de recombinantes pET32aDSCR8 e das frações solúvel e insolúvel. SDS-PAGE 15% mostrando em (1) Marcador de massa molecular; (2) controle não induzido; (3) Rosetta (DE3) pET32aDSCR8 induzida 1 hora; (4) Rosetta (DE3) pET32aDSCR8 induzida 3 horas (5) fração solúvel referente a 5 horas de indução; (6) fração insolúvel (5 horas), (7) fração solúvel (15 horas) e (8) fração insolúvel (15 horas). A seta indica a proteína Trx-DSCR8 de aproximadamente 27 kDa.

Estas condições de expressão possibilitaram que na fração solúvel houvesse Trx-DSCR8 em quantidade suficiente para os ensaios iniciais de purificação por cromatografia de afinidade em coluna de níquel (Ni-NTA). Os resultados preliminares de purificação foram animadores, pois demonstraram que a Trx-DSCR8 tinha afinidade pela resina de níquel e embora estivesse sendo eluída com contaminantes a purificação poderá ser melhorada mudando-se algumas condições. A banda referente a Trx-DSCR8 (~27 kDa) eluída com 250 mM de imidazol pode ser visualizada na Figura 23.

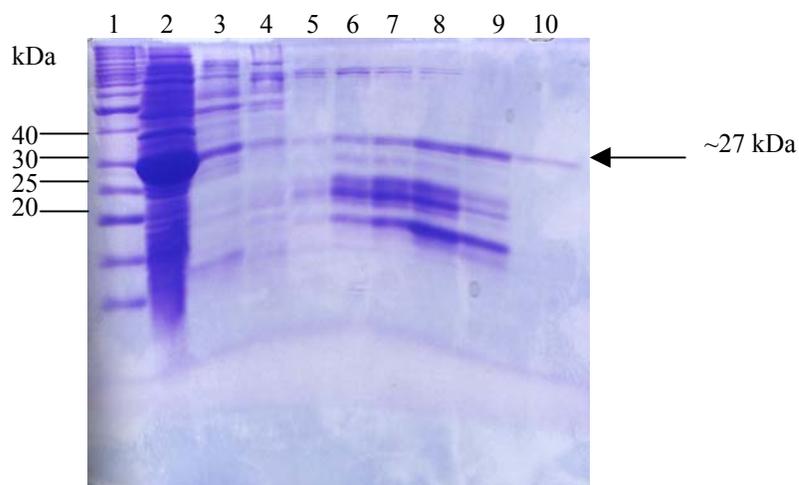


Figura 23. Análise da purificação da DSCR8 por cromatografia de afinidade em coluna de níquel Ni-NTA *Superflow*. Proteínas visualizadas em SDS-PAGE 15%, coradas com azul de *Coomassie*. (1) Marcador de massa molecular; (2) Rosetta (DE3) pET32aDSCR8 induzida por 15 horas; (3, 4) fração eluída em tampão contendo 50 mM de imidazol; (5,6) 75 mM de imidazol; (7,8) 100 mM de imidazol; (9,10) 250 mM de imidazol.

4.8.2 Análises *in silico* e de localização subcelular da DSCR8

O principal objetivo da pesquisa sobre Síndrome de Down é correlacionar o desequilíbrio de dose de genes específicos do cromossomo 21 com os vários fenótipos clínicos da síndrome. Essa tarefa ficaria teoricamente mais fácil se a função molecular e os processos biológicos dos genes do 21 fossem conhecidos. Assim, com o objetivo de dar maiores esclarecimentos sobre a função da DSCR8, a estrutura primária dessa proteína foi analisada *in silico* e o N-terminal desta proteína foi fusionado a GFP para estudos de localização subcelular em células de mamífero, uma vez que até o presente momento nada foi descrito sobre sua localização subcelular, função, bem como sua implicação com a SD.

A proteína DSCR8 é composta por 91 aminoácidos (com uma massa molecular de 10,2 kDa), tem um pI predito de 7,85, apresenta ausência do aminoácido triptofano, tem duas cisteínas e predomínio do aminoácido leucina (13,2%) (Figura 24).

Além disso, o alinhamento da seqüência de aminoácidos da DSCR8 contra seqüências não redundantes depositadas no NCBI (Centro Nacional de Informação para Biotecnologia) mostrou que esta proteína apresenta 100% de identidade com a proteína

MMA-1a (*malignant melanoma associated protein 1a*) e nenhum putativo domínio conservado.

MKEPGPNFVTVRKGLHSFKMAFVKHLLLECSGSITDHC SLHLPVQEILMSQP
PEQLGLQTNLGNQESSGMMKLFMPRPKVL AQYESIQFMP

Figura 24. Sequência de aminoácidos da proteína DSCR8. Os aminoácidos cisteína (C) e leucina (L) estão em destaque.

As células CHO-K1 transfectadas com pEGFPC1-DSCR8 exibem um padrão de localização bastante peculiar, pois ela não está homoganeamente distribuída no citoplasma (padrão C) ou no núcleo (padrão N) e sim como grânulos distribuídos pelo citoplasma celular com ausência total no núcleo (Figura 25).

Os compartimentos onde uma proteína pode ser encontrada são cito-esqueleto, retículo endoplasmático, complexo de Golgi, lisossomo, mitocôndria, entre outros, assim, esse resultado deverá ser investigado em experimentos futuros, uma vez que as análises de predição *in silico* da DSCR8 não revelaram sinal algum de endereçamento para um destes compartimentos.

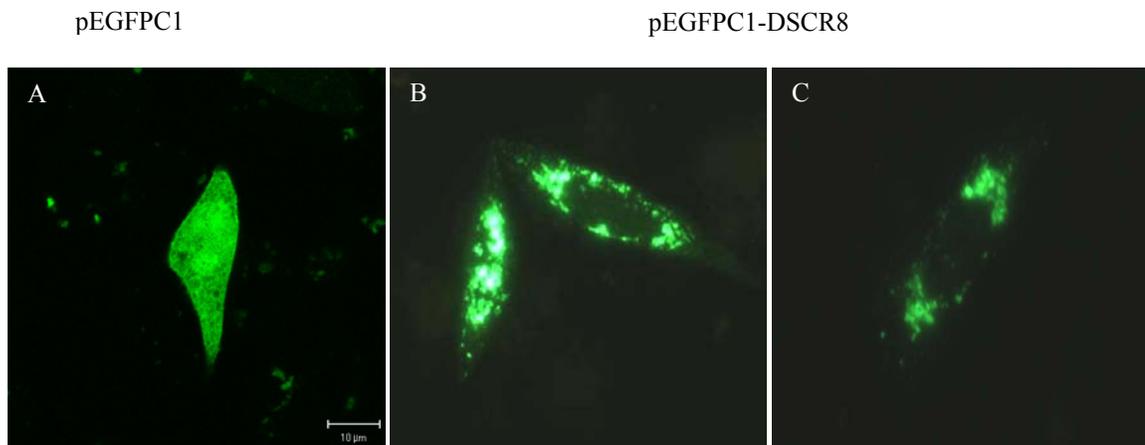


Figura 25. Distribuição subcelular da GFP-DSCR8 em célula CHO-1 transfectada transientemente. Em (A) pEGFPC1 exibe o padrão N+C, isto é distribuição homogênea na célula. (B, C) pEGFPC1-DSCR8 está distribuída como grânulos pelo citoplasma celular com ausência total no núcleo.

4.9 Expressão da proteína DCRA em *E. coli*.

A produção de proteínas em *E. coli* geneticamente engenheiradas tem permitido a obtenção de proteínas para caracterização bioquímica, produção de anticorpos, determinação da estrutura tridimensional, entre outros (RÜCKER et al., 2001). Em adição, a expressão de proteínas heterólogas em fusão com cauda de histidina (*His Tag*) tem possibilitado a obtenção de proteínas com grau de pureza adequado para ensaios cristalográficos, necessários para a compreensão da estrutura e função da proteína alvo (SAMBROOK & RUSSEL, 2001).

Embora a clonagem do cDNA seja relativamente fácil, a expressão dos genes clonados ainda é uma tarefa árdua. As explicações do porque disto, infelizmente não são totalmente compreendidas. Geralmente as proteínas recombinantes são expressas em *E. coli* como agregados insolúveis, denominados de *corpos de inclusão*, ou são prematuramente degradadas e/ou desnaturadas (SAMBROOK & RUSSEL, 2001, op cit.).

Segundo STEVENS (2000) 15 a 20% de genes humanos expressos em *E. coli* geram proteínas solúveis, 20 a 40% formam corpos de inclusão e os 60% remanescentes não são significativamente expressos ou são degradados.

O baixo rendimento de polipeptídeos expressos *in vivo* é geralmente devido à síntese intracelular ineficiente e/ou degradação de tais polipeptídeos (SKOSYREV et al., 2003). Tem sido sugerido que proteínas bacterianas são mais facilmente expressas em bactéria enquanto que proteínas tipicamente de eucariotos são incorretamente enoveladas e endereçadas para degradação (BANEYX & GEORGIU, 1992, apud, CHA et al., 2000).

Proteínas incorretamente enoveladas (*misfolded*) freqüentemente agregam-se quando superexpressas em bactérias e a despeito dos inúmeros relatos de delineamento experimental sobre enovelamento de proteínas e tentativas de relacionar solubilidade com a seqüência de aminoácidos, ainda não existe nenhum método para prever o *misfold in vivo*.

Apesar dessas limitações, *E. coli* é o sistema procarionte mais freqüentemente usado para expressão de proteínas heterólogas por ser um método barato, mais rápido de ser obtido do que em sistemas eucariontes e não gerar heterogeneidade do produto devido à ausência de modificação pós-traducional, fator fundamental caso a proteína superexpressa seja utilizada em estudos estruturais. Os sistemas eucariontes também apresentam várias limitações para a expressão de proteínas heterólogas, por exemplo, células de mamífero (necessidade de ambiente estéril e manipulação em condições estritamente assépticas,

instabilidade das células geneticamente manipuladas, baixo rendimento quando comparado com microorganismos); fungos (dificuldade de controle da expressão, mecanismo de glicosilação diferente do mamífero); células de insetos (o produto nem sempre é funcional).

A expressão de proteínas recombinantes pode ser obtida utilizando uma variedade de sistemas de expressão já disponíveis comercialmente. A escolha do sistema adequado depende de vários fatores, dentre eles estão: a necessidade de grandes quantidades de proteína, a facilidade posterior de purificação e a toxicidade da proteína alvo para a célula hospedeira.

Na literatura há várias informações referentes à otimização das condições de expressão em *E. coli* para proteínas específicas. Contudo, as proteínas freqüentemente apresentam propriedades físicas e químicas diferentes e, portanto as condições ideais de expressão para uma dada proteína não são iguais para todas as proteínas (BRAUN et al., 2002). Assim sendo, triagens empíricas devem ser realizadas para prever as construções que expressarão proteínas solúveis e insolúveis.

Em adição outros fatores, incluindo a localização da proteína, custos da resina de afinidade e tampões e as possibilidades de remoção da proteína de fusão por clivagens enzimáticas também são fatores que devem ser considerados na escolha do sistema de expressão (NILSSON et al., 1997).

Assim, optamos pelo sistema de expressão em *E. coli* e escolhemos vários vetores disponíveis comercialmente na tentativa de estabelecer o melhor sistema para a expressão da DCRA em procarioto. Os critérios usados para a escolha dos vetores de expressão foram baseados nas vantagens que as proteínas parceiras poderiam apresentar no que se refere à capacidade de aumentar a solubilidade (Trx, NusA, MBP) e facilitar a purificação tais como: sinal para exportar a proteína para o espaço periplasmático (pET 37) ou para o meio de cultura (pCX TOPO), seqüências His-Tag, S-Tag, GST, CBD que possibilitariam a purificação por afinidade e sítios de clivagem para proteases que permitiriam a retirada das respectivas fusões.

Contudo como pode ser visto na tabela 2 a maior parte dos vetores utilizados para clonagem da DCRA não resultou em expressão.

Tabela 2. Vetores de expressão construídos para a expressão da DCRA e da DSCR8 em *E. coli*.

Vetor expressão	Seleção	Proteína expressa	Massa Molecular esperada (kDa)	Teste solubili-dade	Presença de expressão
pET 28a	Kan	His tag-DCRA	~ 35	--	Não
pET 32a	Amp	Trx-DCRA	50	--	Não
pET 32a	Amp	Trx-DCRA _{Nco/Bam HI}	~ 46	insolúvel	Sim
pET 32a	Amp	Trx-DCRA _{HindIII}	~ 26	--	Não
pET 32a	Amp	Trx-DCRA-GFP	77	solúvel	Sim
pET 37b	Kan	CDB-DCRA	46	--	Não
pET 41a	Kan	GST-DCRA	59	--	Não
pET 43a	Amp	NusA-DCRA	88	--	Não
pMALc2	Amp	MBP-DCRA	78	--	Não
pTYB2	Amp	DCRA-Inteína	88	--	Não
pBAD-TOPO	Amp	DCRA-His tag	~ 35	--	Não
pCX-TOPO	Cloranfenicol	DCRA-RsaA	53	--	Não
PCRT7/CN-TOPO	Amp	His tag-DCRA e DCRA-His tag	~ 35	--	Não
pCDNA3 para eucarioto	Amp/Neo	His tag-DCRA	~ 35	?	Não concluído
pET 28a	Kan	His tag-DSCR8	~ 12,2	insolúvel	Sim
pET 32a	Amp	Trx-DSCR8	~ 27,2	solúvel	Sim

Por que algumas proteínas de eucariotos são expressas em *E. coli* em quantidades muito pequenas? Algumas razões são óbvias em alguns casos e desconhecidas em outros.

BRAUN et al. (2002) sugerem que o sucesso da expressão da proteína está relacionado com a presença de certos domínios e que a ausência de modificações pós-traducionais em *E. coli* poderia desestabilizar algumas proteínas eucariontes particularmente as proteínas grandes.

Os manuais da Qiagen (2001) e do Sistema pET sugerem alguns motivos para a ausência ou baixa expressão tais como: (1) a proteína alvo poderia ser tóxica ou instável

para a célula hospedeira, ou o sistema de expressão poderia ser perdido durante o crescimento da bactéria; (2) a expressão de proteínas heterólogas poderia diminuir o crescimento do hospedeiro (produto tóxico) inviabilizando a tradução da proteína alvo ou toda a síntese proteica poderia parar antes da proteína alvo ser detectada e nesses casos observa-se à lise da célula; (3) proteínas recombinantes com regiões hidrofóbicas freqüentemente possuem um efeito tóxico sobre o hospedeiro devido à associação ou interação dessa proteína com o sistema de membrana da bactéria; (4) poderia ocorrer a degradação da proteína por proteases bacterianas, embora pouco provável uma vez que a linhagem BL21(DE3) normalmente usada para expressão é deficiente para as proteases *lon* e *ompT*; (5) a proteína alvo poderia ter uma meia vida de somente 2 minutos quando os aminoácidos Arg, Lys, Phe, Leu, Trp e Tyr estão presentes no amino-terminal. Por alguma razão esses aminoácidos determinariam uma relação entre os aminoácidos presentes no amino-terminal da proteína e sua instabilidade na bactéria; (6) excesso de códons raros na proteína alvo e (7) mutação sem sentido.

As causas enumeradas acima sugerem razões que são evidentes em alguns casos, mas muito provavelmente não explicam a ausência de expressão da DCRA. Segundo WANG, et al. (2003) alguns genes não se expressam por alguma razão que não é óbvia.

Dentre todas as construções realizadas (Tabela 2) a denominada Trx-DCRA_{Nco/BamHI} possibilitou a expressão da DCRA, porém na forma de corpos de inclusão. Deste modo foi necessário otimizar as condições de cultivo para a expressão da referida proteína na fração solúvel.

Para isso os experimentos de expressão foram feitos com (1) a linhagem BL21(DE3) (STUDIER et al., 1990) uma linhagem que cresce bem em meio LB e codifica a T7 RNA polimerase sob o controle do promotor T7 e (2) linhagens hospedeiras com diferentes características genotípicas, como por exemplo, à linhagem AD494 para expressão do pET 32a, a ER2566 para o pTYB2 (ver material e métodos) e outras como a BL21 SI, BL21 pT-GROE visando especificamente o aumento da solubilidade da proteína alvo. No entanto, como já citado, tais cuidados não trouxeram resultados positivos.

Ao mesmo tempo as construções realizadas nesse trabalho foram induzidas com 0,1; 0,4 e 1,0 mM de IPTG, induzidas a 37⁰ C e 20⁰ C por 1, 2, 4, 6, 15 horas nos meios nutritivos LB e *Circle Grow*. Essas variáveis às vezes trazem resultados satisfatórios, mas não necessariamente.

PRYOR & LEITING (1997) compararam o uso de meio mínimo (MM/CA) com suplementação *versus* meio rico (LB) e concluíram que o meio mínimo traz algumas

vantagens, pois reduz a expressão do promotor *tac*, possibilita que proteínas tóxicas possam ser induzidas em uma densidade celular maior do que no meio LB e geralmente facilitam a purificação devido ao menor número de proteínas do hospedeiro frente à proteína superexpressa. A redução da fonte de carbono e nitrogênio no meio simplifica as vias de biossíntese de aminoácidos e, portanto a quantidade de proteína do hospedeiro. Além disso, a substituição do meio de cultura por um meio fresco antes da indução possibilita uma redução no acúmulo de metabólitos no final do período de expressão, novamente facilitando a purificação. Porém segundo esses autores a composição do meio não tem nenhuma relação significativa com nível de expressão ou crescimento celular.

Nesse trabalho a fusão dupla de genes (Trx-DCRA-GFP) e a temperatura de 20°C após a indução foram estratégias que trouxeram resultado satisfatório na expressão da DCRA (Tabela 2).

Pela literatura observa-se que as proteínas parceiras amplamente usadas são a tiorredoxina (Trx), *maltose-binding protein* (MPB), *glutathione S-transferase* (GST), *calmodulin binding protein* (CBP), NusA e *cellulose-associate protein* (CAP). Em adição o aumento da solubilidade frequentemente tem sido associado com o tamanho da proteína parceira, por exemplo, NusA (54 kDa), MBP (42 kDa) e GST (24 kDa) contribuíram para aumentar a solubilidade respectivamente em 60%, 60% e 38% dos casos testados por SHIH et al. (2002).

Embora o uso dessas proteínas tenha trazido resultados positivos na expressão de muitas proteínas heterólogas ainda não há um estudo comparativo dos efeitos de diferentes proteínas de fusões e possibilidades de *tags* na expressão e na solubilidade (HAMMARSTRÖM, et al., 2002) corroborando novamente para que cada caso de expressão seja único.

Além disso, a diversidade de situações testadas para aumentar a solubilidade é tão vasta que nos parece que não basta somente encontrar a melhor parceira é preciso também saber combiná-la, já que às vezes uma fusão no C-terminal pode ser mais efetiva do que no N-terminal. Em suporte a essa afirmação está o fato de DIECKMAN et al. (2002) observarem que vários clones reunidos no grupo denominado de grupo 2 exibiam solubilidade diferencial quando expressos somente em fusão com a His *tag* ou em fusão His-CBP-*tag*, já a combinação dos dois vetores his:pMCSG3 e his-CBP:pMCSG4 possibilitou um aumento de 10% na solubilidade das proteínas que compõem o grupo 2 quando comparado ao obtido com somente um dos vetores.

COULOMBE & MELOCHE (2002) observaram que a inserção da GST no C-terminal da proteína de interesse, ao invés da configuração usual N-terminal derivada do pGEX, foi mais efetiva na diminuição de forma de proteína truncada, isto é uma proteína menor devido há um códon de terminação prematuro (N-terminal) ou um possível início de tradução interno (C-terminal).

Finalizando, a fusão de genes é uma abordagem que tem trazido sucesso para a obtenção de proteínas solúveis em *E. coli*, no entanto a escolha da melhor parceira é empírica e pelos inúmeros exemplos citados na literatura e depois de todas as tentativas para expressar a DCRA com diferentes parceiras concluímos que a proteína de fusão é proteína-específica.

Assim a melhor abordagem para a expressão da DCRA foi a dupla fusão de genes, isto é, tiorredoxina do N-terminal e GFP no C-terminal da DCRA, pois permitiu não somente que DCRA fosse expressa em *E. coli* como também contribuiu para o aumento da solubilidade e da estabilidade.

5 Purificação da proteína Trx-DCRA-GFP

5.1 Cromatografia NI-NTA *Superflow*

O resultado da análise da purificação parcial através do fracionamento com sulfato de amônio em SDS-PAGE 15% indica que a fusão tiorredoxina-DCRA-GFP precipita a 30% de sulfato de amônio. Após a precipitação com sulfato de amônio a solução protéica foi devidamente dialisada e submetida à cromatografia de afinidade NI-NTA *Superflow*. Durante a purificação observou-se que ao passar o sobrenadante pela coluna, a resina de níquel de cor originalmente azul, passou a ter a coloração esverdeada, dada pela GFP, indicando que a fusão ficou retida na resina de afinidade.

A solução protéica contendo a fusão tiorredoxina-DCRA-GFP começou a ser eluída na concentração de 50mM de imidazol. O fato da concentração de imidazol ser relativamente baixa possibilita que a fração seja eluída com proteínas contaminantes, ou seja, proteínas que se ligaram fracamente à resina.

Essa afinidade de ligação relativamente fraca entre a fusão e a resina pode estar relacionada ao fato da cauda de histidina ficar “ensanduichada” na construção, isto é tiorredoxina-cauda histidina-DCRA-GFP.

Segundo o manual da Qiagen (2001) para evitar que contaminantes sejam eluídos com a proteína é aconselhável aumentar a força iônica ou a concentração de detergente, glicerol ou etanol para impedir interações inespecíficas com a matriz, ou adicionar no tampão de ligação baixas concentrações de imidazol ou β -mercaptoetanol. Todas essas sugestões foram realizadas de acordo com o Manual da Qiagen (2001), mas apesar disso, a proteína continuou sendo eluída com contaminantes (Figura 26). Optou-se então por mais uma etapa cromatográfica.

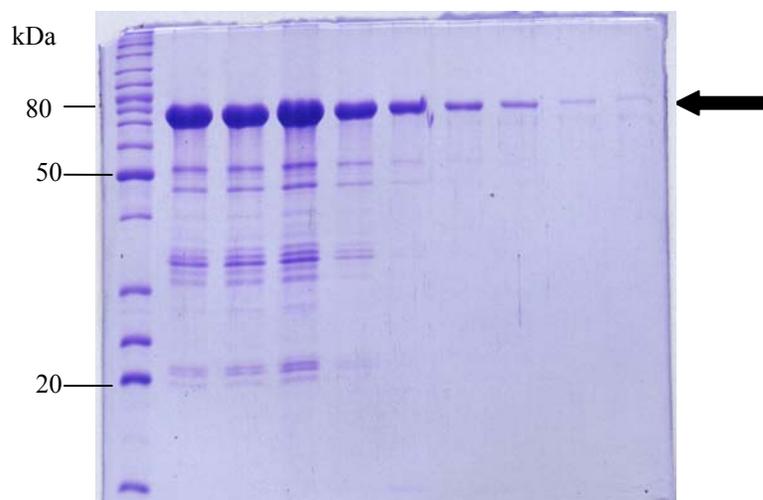


Figura 26. Análise da purificação da Trx-DCRA-GFP em cromatografia de afinidade Ni-NTA *Superflow*. Proteínas visualizadas em SDS-PAGE 15%, coradas com azul de *Coomassie*. (1) Marcador de massa molecular; (2) fração eluída em tampão contendo 50 mM de imidazol; (3 e 4) 75 mM de imidazol; (5 e 6) 100 mM de imidazol; (7 e 8) 200 mM de imidazol; (9 e 10) 250 mM de imidazol.

5.2 Cromatografia Mono-Q HR5/5

Para a retirada dos contaminantes a fração purificada em cromatografia de afinidade foi recromatografada em coluna de troca aniônica Mono-Q HR5/5. Apesar de essa recromatografia ter exibido um pico de eluição bastante promissor (Figura 27), a

análise em SDS-PAGE dessas frações mostrou que a solução protéica ainda continha contaminantes (Figura 28). Embora a análise de espalhamento dinâmico de luz das frações purificadas indicasse um grau de homogeneidade satisfatório para ensaios subseqüentes de cristalização o rendimento final da purificação era de 1 a 1,2 mg de proteína para cada 2 litros da suspensão de células de *E. coli*, o que limitava a realização dos referidos experimentos.

A presença de proteínas contaminantes na amostra sugere que (1) muito provavelmente esses contaminantes possuem afinidade pela DCRA e não pela resina, já que Trx e GFP independentemente purificadas não apresentam contaminantes (2) ou a proteína tiorredoxina-DCRA-GFP estava conformacionalmente heterogênea apresentando, portanto, populações com diferentes afinidades pela resina.

Concluindo, as cromatografias de afinidade e de troca iônica usada para purificar a fusão tiorredoxina-DCRA-GFP não foram resolutivas devido à presença constante de contaminantes.

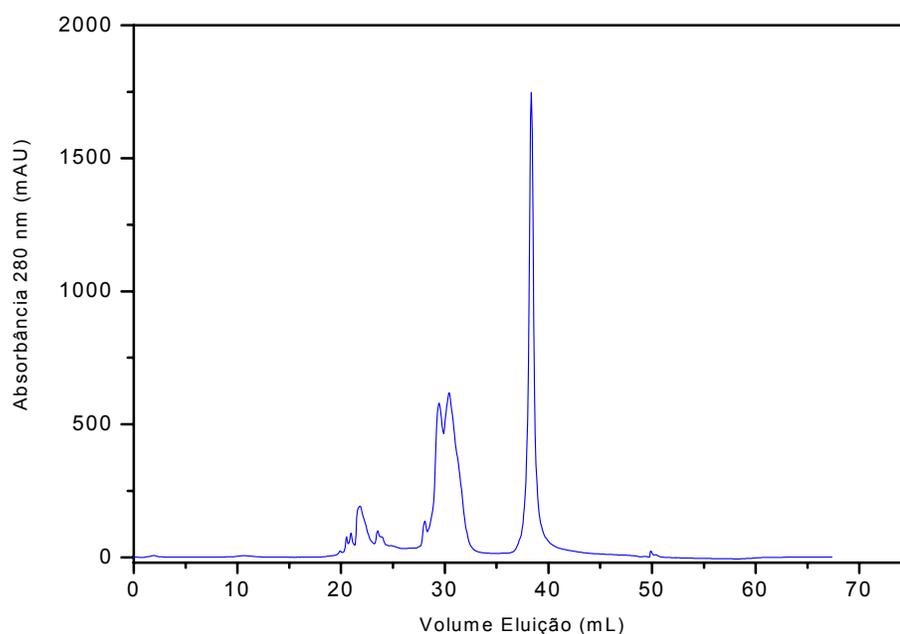


Figura 27. Perfil cromatográfico da eluição da tiorredoxina-DCRA-GFP na coluna de troca aniônica Mono-Q HR5/5. A amostra aplicada (frações mais puras obtidas na cromatografia de afinidade NTI-NTA) foi concentrada por centrifugação em Centricon 50. A eluição foi feita em um gradiente linear de 0 a 500 mM de NaCl em tampão Tris-HCl 25mM, pH 7,5 sob o fluxo de 1 mL/minuto e a eluição com 500 mM de NaCl.

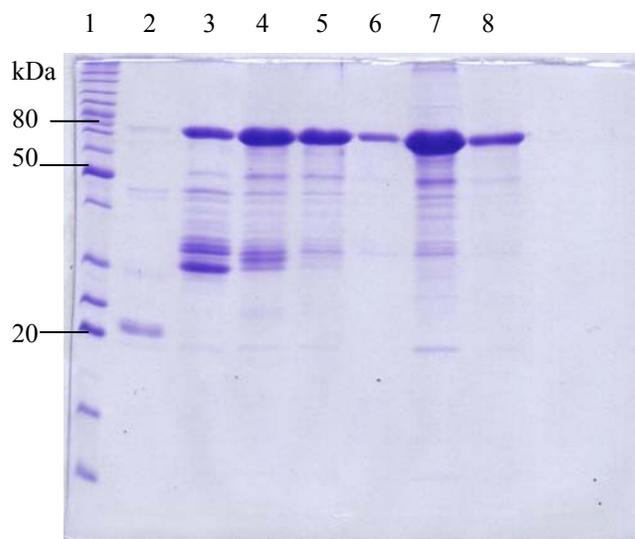


Figura 28. Análise da purificação da Trx-DCRA-GFP em cromatografia de troca aniônica Mono-Q HR5/5. Proteínas visualizadas em gel SDS-PAGE 15%, coradas com azul de *Coomassie*. (1) Marcador de Massa molecular; (2, 3, 4 e 5) amostras obtidas durante o início do gradiente salino; (6, 7 e 8) frações eluídas com 500 mM de NaCl. As amostras aplicadas nas colunas 6, 7 e 8 correspondem ao pico maior observado na figura 27.

5.3 Cromatografia em Superose 12 HR10/30

Para a retirada dos contaminantes, a fração purificada em cromatografia de afinidade foi recromatografada em coluna Superose 12 HR10/30. Embora essa recromatografia tenha exibido um pico de eluição promissor (Figura 29), a análise em SDS-PAGE dessas frações mostrou que a solução protéica ainda continha contaminantes (Figura 30) o que limitava a realização dos experimentos subseqüentes uma vez que se perde muita proteína após várias cromatografias.

Como a DCRA ainda estava em fusão com as proteínas parceiras tiorredoxina e GFP e foi eluída com vários contaminantes não foi possível a determinação da massa molecular aparente da DCRA.

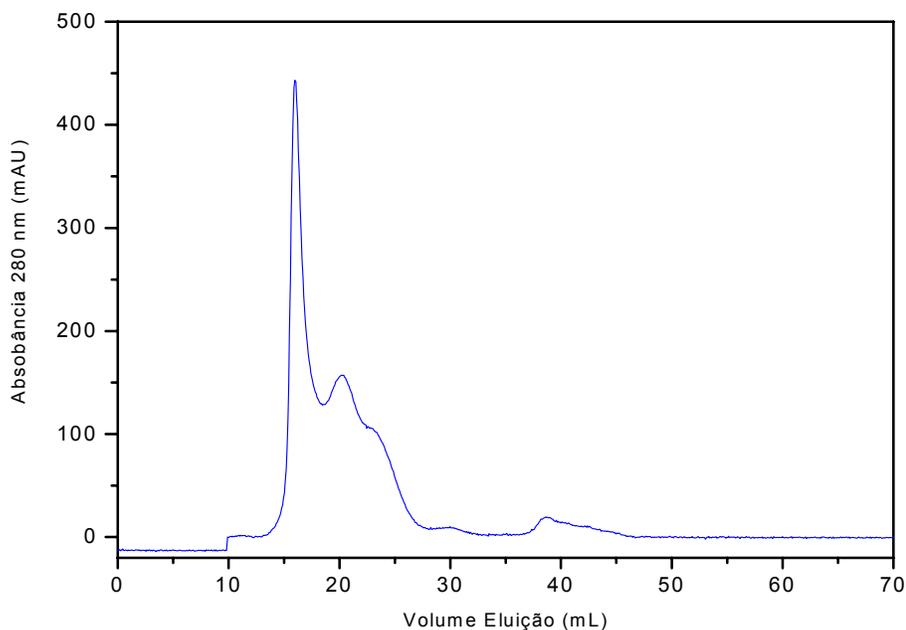


Figura 29. Perfil cromatográfico da eluição da tiorredoxina-DCRA-GFP na coluna Superose 12 HR 10/30. A amostra aplicada (frações mais puras obtidas na cromatografia de afinidade (NTI-NTA) foi concentrada por centrifugação em Centricon 50).

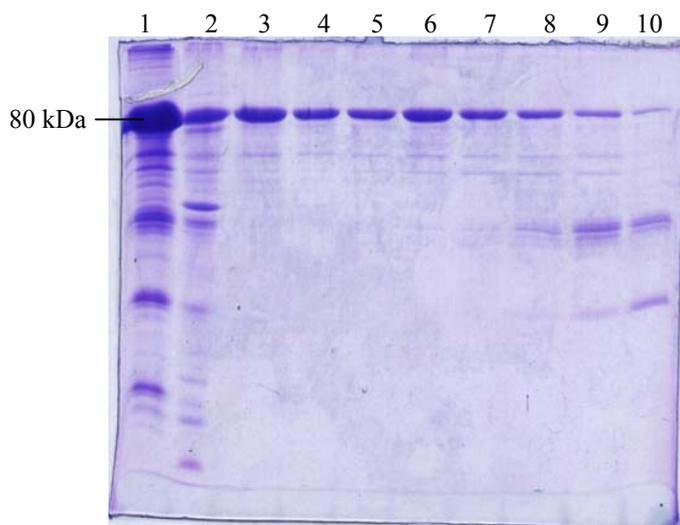


Figura 30. Análise da purificação da Trx-DCRA-GFP em cromatografia de exclusão molecular Superose 12 HR 10/30. Proteínas visualizadas em SDS-PAGE 15%, coradas com azul de *Coomassie*. (1) Amostra injetada na coluna; (2 a 10) frações eluídas (estas frações correspondem aos dois picos maiores observados na figura 29).

6 Clivagem da proteína Trx-DCRA-GFP com a enteroquinase

Tradicionalmente a purificação de proteínas recombinantes em cromatografia de afinidade frequentemente inclui passos extras para remover a fusão da proteína alvo através de clivagens enzimáticas.

A presença do sítio de clivagem reconhecido pela enteroquinase presente no vetor pET32a e entre DCRA e GFP permitiriam separar as fusões tiorredoxina e GFP respectivamente fundidas no N e no C-terminal da DCRA. Para isso fez-se à digestão proteolítica com a enzima enteroquinase, visto que, tomou-se o cuidado inicial de analisar se havia algum sítio interno de reconhecimento para proteases na seqüência primária de aminoácidos da DCRA.

A análise *in silico* revelou que na seqüência primária de aminoácidos da DCRA não há seqüências de reconhecimento para as enzimas enteroquinase e Fator Xa, entretanto, após a clivagem com a protease enteroquinase observou-se que ocorreram várias proteólises inespecíficas em todas as condições testadas (Figura 31).

Infelizmente as endopeptidases sofrem muitas limitações tais como presença de sítios de clivagem secundários, clivagem incompleta ou até mesmo inibição da clivagem devido ao enovelamento da proteína (STEVENS, 2000). A enteroquinase reconhece a seqüência Asp-Asp-Asp-Asp-Lys[↓], mas dependendo da conformação do substrato pode reconhecer outros resíduos básicos formando o então denominado sítio de clivagem secundário (STEVENS, 2000, op cit).

STOLL et al. (1998), relataram que somente foi possível a expressão da proteína VanH na forma solúvel quando expressa em fusão com a tiorredoxina, no entanto todas as tentativas de clivagem com enteroquinase para a retirada da fusão resultaram em proteólise inespecífica.

Inicialmente pensou-se que tais proteólises teriam acontecido porque as fusões tiorredoxina e GFP poderiam ter deixado o acesso ao sítio de protease estericamente desfavorecido. No entanto, o padrão de clivagens observado demonstrou que com o passar do tempo de incubação, a banda referente à fusão gradualmente ia diminuindo, confirmando que a fusão estava sendo clivada. Porém também relevou que além das bandas na altura 17 kDa e 27 kDa previstas respectivamente para a tiorredoxina e a GFP, havia inúmeras bandas inespecíficas. Além disso, a ausência da banda na altura de 33 kDa prevista para DCRA (Figura 31) sugere que esta proteína após ser liberada das fusões, poderia ter um comportamento muito instável. Em contraste GFP é altamente resistente à

proteólise *in vivo* e *in vitro* (SKOSYREV et al., 2003) e de alguma forma estabilizou a fusão, pois não foram observadas clivagens inespecíficas no controle.

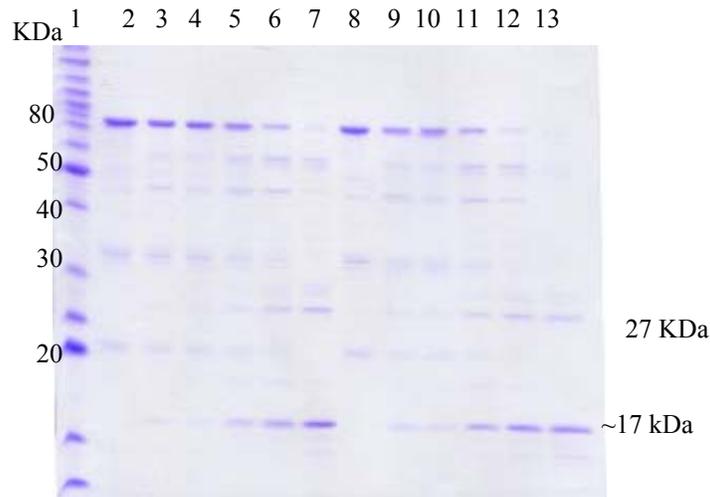


Figura 31. Análise da clivagem da proteína tiorredoxina-DCRA-GFP. SDS PAGE 15% corado com azul de Coomassie. (1) Marcador de massa molecular; (2) controle não clivado; (3 a 7) clivagem durante 1 hora com 0,1; 0,2; 0,5; 0,7 e 1,0 U de enteroquinase respectivamente; (8) controle não clivado (9 a 13) clivagem durante 2 horas com 0,1; 0,2; 0,5; 0,7 e 1,0 U de enteroquinase respectivamente. As bandas de aproximadamente 77, 27 e 14 kDa correspondem respectivamente à fusão, a GFP e tiorredoxina.

Segundo VARSHAVSKY (1996) muitas proteínas possuem um longo tempo de vida como componentes de grandes complexos tais como ribossomos e proteínas oligoméricas, mas, são metabolicamente instáveis como subunidades livres. Se considerarmos que DCRA possui um putativo domínio conservado de Vps26 e que esta proteína é componente de um retrômero (complexo multimérico) que compreende 5 proteínas (Vps35, Vps29, Vps5, Vps17, e Vps26), então DCRA poderia ser estável quando fusionada a outras proteínas e se tornar instável quando desagregada da associação com Trx e GFP. Além disso, REDDY & SEAMAN (2001) observaram que a remoção do C-terminal de Vps26p fez com que essa proteína tivesse uma meia-vida inferior a 10 minutos, sugerindo que a remoção do C-terminal afetou a estabilidade da proteína *in vivo*.

SATI et al. (2002) clonaram seis formas truncadas da proteína Pfg27 em fusão com a MBP (*maltose binding protein*) e em todas elas havia sucesso na retirada da fusão por clivagem proteolítica, no entanto, em cinco das seis construções a proteína Pfg27 truncada precipitava após a retirada da proteína de fusão indicando que Pfg27 estava incorretamente enovelada. Somente a construção denominada Pfg27^A e que continha um resíduo adicional

de 17 aminoácidos era estável e solúvel após a clivagem da fusão. Segundo SATI et al. (op cit) o tamanho e a seqüência específica desses 17 aminoácidos podem ter contribuído para o efeito observado. Fenômeno semelhante em que a presença de uma cauda extra no C-terminal aumenta a estabilidade de proteína superexpressa também foi relatado por SILBER, KEILER & SAUER (1992).

Para YOKOYAMA (2003) uma grande fração de proteínas identificadas como “bem expressa” no sistema *high-throughput* poderia ter um comportamento instável, formar corpos de inclusão e enovelamento parcial uma vez que as *tag* popularmente usadas algumas vezes melhoram o comportamento de tais proteínas, contudo freqüentemente a clivagem dessa cauda da fusão resulta em baixa solubilidade e estabilidade.

A meia vida de proteínas em qualquer célula viva varia de poucos segundos há muitos dias. Entre as funções intracelulares de proteólise celular está a eliminação de proteína anormal, manutenção do reservatório de aminoácidos e geração de fragmentos protéicos que atuam como hormônios, antígenos ou outros efetores. A instabilidade *in vivo* das proteínas está possivelmente condicionada a uma determinada ordem de certos aminoácidos na estrutura primária de uma dada proteína (GURUPRASAD, REDDY & PANDIT, 1990).

GURUPRASAD, REDDY & PANDIT (op cit.) analisaram a seqüência de aminoácidos de proteínas estáveis e instáveis e propuseram um índice para predição de instabilidade. De acordo com a predição desses autores proteínas com índice de instabilidade menor do que 40 é predita ser estável, enquanto que valores maiores de 40 predizem que a proteína possa ser instável. O índice de instabilidade, predito no servidor <http://us.expasy.org/tools/protpar-ref.html> para a DCRA é 35,26 sendo então, classificada como estável.

Vários estudos têm mostrado que resíduos polares no carboxi terminal da proteína são capazes de evitar o reconhecimento por proteases específicas e apontam para uma natureza complexa e multifatorial do enovelamento da proteína indicando que os fenômenos de estabilidade ou instabilidade de uma proteína ainda não são completamente entendidos (SATI et al., 2002).

7 Espalhamento Dinâmico de Luz

A cristalização de macromoléculas biológicas freqüentemente representa a etapa limitante em análises cristalográficas com Raio X. O sucesso da cristalização depende de vários fatores entre eles uma solução protéica homogênea. Infelizmente a pureza covalente normalmente analisada em SDS-PAGE ou espectrofotometria de massa aliada com exaustiva seleção de condições de cristalização não garantem o sucesso desejado (FERRE-D'AMARÉ & BURLEY, 1994).

A técnica do Espalhamento Dinâmico de Luz pode ser empregada para a seleção da solução protéica monodispersa candidata à cristalização restringindo as condições de cristalização somente para preparações monodispersas e conseqüentemente eliminando tempo e custos de se trabalhar com materiais inadequados (FERRE-D'AMARÉ & BURLEY, 1994 op cit).

Os resultados do espalhamento dinâmico de luz indicaram que a amostra protéica tiorredoxina-DCRA-GFP tinha um raio hidrodinâmico de 4,00 nm e uma massa molecular estimada de 86 kDa. A massa molecular predita era de aproximadamente 80 kDa, assim essa diferença de 6 kDa poderia ser atribuída ao fato da fusão não ser totalmente globular. Além disso, pelo histograma representativo do raio hidrodinâmico da proteína Trx-DCRA-GFP (Figura 32) pode-se dizer que a solução protéica apresentava-se com um grau de homogeneidade adequado para que fosse submetida aos ensaios de cristalização.

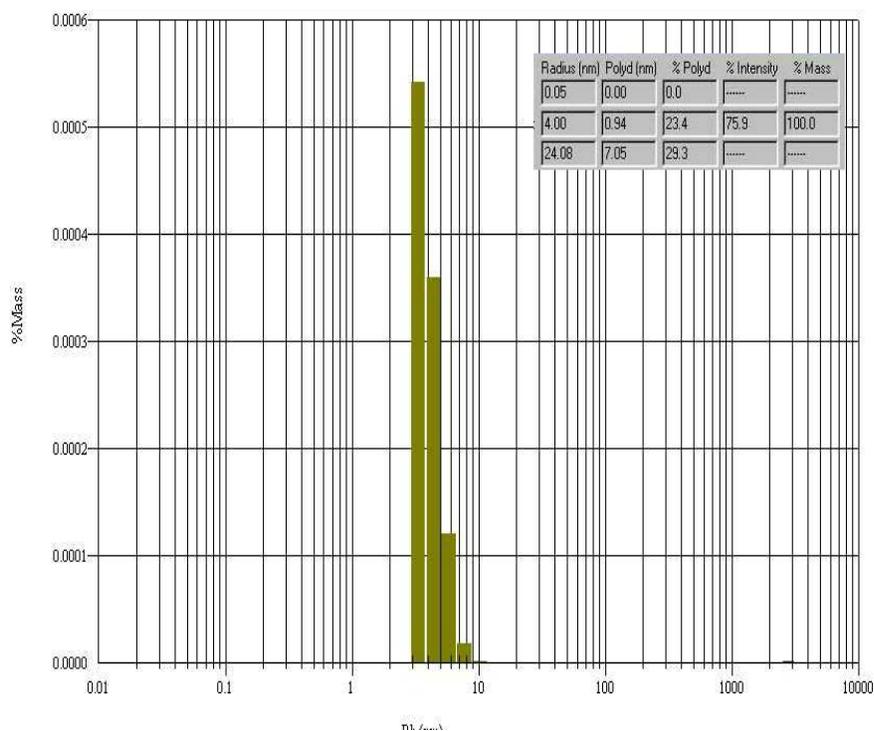


Figura 32. Histograma representativo do raio hidrodinâmico da proteína thiorredoxina-DCRA-GFP obtido pela técnica de espalhamento dinâmico de luz.

8 Ensaio de cristalização

A cristalização freqüentemente representa uma etapa crítica nos estudos estruturais, pois o sucesso do crescimento do cristal de proteínas é estritamente dependente da pureza da amostra, isto é, livre de contaminantes, monodispersa (todas do mesmo tamanho), homogênea, “conformacionalmente pura” e com ausência de outras microheterogeneidades estruturais que compreendem variações na estrutura primária, secundária, terciária bem como a presença de agregados (STEVENS, 2000). Além disso, cada proteína possui condições de cristalização específicas e algumas até cristalizam mesmo na presença de impurezas (STEVENS, 2000 op cit.). Por isso, o crescimento de um cristal pode ocorrer em 15 minutos ou em 1 ano (HAMPTON Research User Guide).

Até o momento, as condições de cristalização testadas (fatoriais I e II), não foram satisfatórias. Em algumas condições dos fatoriais I e II houve a formação de precipitados, uma semana após a montagem das caixinhas de cristalização e ausência de cristais nas demais. A formação de precipitado nas gotas indica que a supersaturação relativa da

amostra ou do reagente era alta demais, que houve desnaturação da amostra ou que esta estava heterogênea.

Devido à grande dificuldade na obtenção da solução protéica em quantidade e com grau de homogeneidade necessária bem como a retirada das fusões, acredita-se que a expressão da DCRA em hospedeiro eucarioto seria a melhor abordagem, assim a *Histag-DCRA* foi clonada no vetor pcDNA3.

9 Predições *in silico*

Através da análise da seqüência primária de aminoácidos é possível prever e sugerir uma possível função de acordo com a presença de motivos e domínios específicos, caracterização bioquímica e a localização subcelular. A função de uma proteína está estritamente correlacionada com sua localização subcelular. Assim, a identificação e a caracterização de componentes protéicos da maquinaria e da estrutura celular são passos essenciais não somente para definir a função gênica, mas também para o entendimento das funções celulares e sua regulação.

Os *Web Softwares* disponibilizam ferramentas que possibilitam a obtenção de pistas adicionais que auxiliam na caracterização funcional de novas proteínas e conseqüentemente essas informações auxiliam no delineamento experimental que visa à caracterização de proteínas de função ainda desconhecidas.

Para prever a localização subcelular de uma proteína, o primeiro método poderia ser à procura de seqüências sinal na estrutura primária da proteína. Outra possibilidade seria inferir a localização a partir da comparação da seqüência de aminoácidos da proteína alvo com outras proteínas de localização conhecida. Proteínas citoplasmáticas possuem proporções eqüitativas de aminoácidos básicos e ácidos, proteínas extracelulares possuem maior porcentagem de resíduos ácidos enquanto que em proteínas nucleares predominam os resíduos básicos (ANDRADE, O'DONOGHUE & ROST, 1998), poucos aminoácidos hidrofóbicos especialmente os aromáticos, e geralmente alta proporção de serina, treonina, prolina, asparagina e glutamina (CEDANO et al., 1997). As proteínas de membranas são ricas em resíduos hidrofóbicos já as intracelulares são pobres em cisteína e relativamente ricas em aminoácidos alifáticos e com carga (CEDANO et al., 1997 op cit).

Os compartimentos onde uma proteína pode ser encontrada são, citoplasma, citoesqueleto, retículo endoplasmático, complexo de Golgi, lisossomo, mitocôndria, núcleo, peroxissomo, membrana plasmática, cloroplasto (plantas) e vacúolo (fungos).

9.1 Estrutura primária da proteína DCRA

A proteína DCRA (**D**own **S**ndrome **C**ritical **R**egion **P**rotein **A**), também denominada DSCR3 (**D**own **S**ndrome **C**ritical **R**egion **P**rotein **3**) é composta por 297 aminoácidos (com uma massa molecular de 33 kDa), possui proporções equitativas de aminoácidos hidrofóbicos e hidrofílicos, tem um pI predito de 7,63, apresenta ausência do aminoácido triptofano, oito cisteínas e predomínio dos aminoácidos valina e leucina.

Analisando a seqüência de aminoácidos da DCRA observamos que ela possui 49,83% de aminoácidos hidrofóbicos e 50,17% de aminoácidos hidrofílicos.

```
1   MGTALDIKIK RANKVYHAGE VLSGVVVISS KDSVQHQGVS LTMEGTVNLQ LSAKSVGVFE
61  AFYNSVKPIQ IINSTIEMVK PGKFPSGKTE IPFEFPLHLK GNKVLVYETIH GVFVNIQYTL
121 RCDMKRSLLA KDLTKTCEFI VHSAPQKGKF TPSPVDFITIT PETLQNVKER ALLPKFLLRG
181 HLNSTNCVIT QPLTGELVVE SSEAAIRSVE LQLVRVETCG CAEGYARDAT EIQNIQIADG
241 DVCRGLSVPI YMVFPRLFTC PTLTETNFKV EFEVNIVVLL HPDHLITENF PLKLCRI
```

Figura 33. Seqüência de aminoácidos codificados pelo gene DCRA.

9.2 Similaridade

O alinhamento da seqüência de aminoácidos da DCRA mostrou que esta proteína apresenta 57% e 49 % de identidade respectivamente com uma proteína de *Anopheles gambiae* e de *Drosophila melanogaster* ambas de função desconhecida. Além de apresentar 91% de identidade com uma proteína conservada não caracterizada de *Mus musculus*; 41% de identidade com uma proteína desconhecida de *Arabidopsis thaliana*; 21% de identidade com Vps26 (*Vacuolar sorting protein*) e 79% de identidade com uma proteína hipotética humana sugerindo uma provável função conservada dessas proteínas.

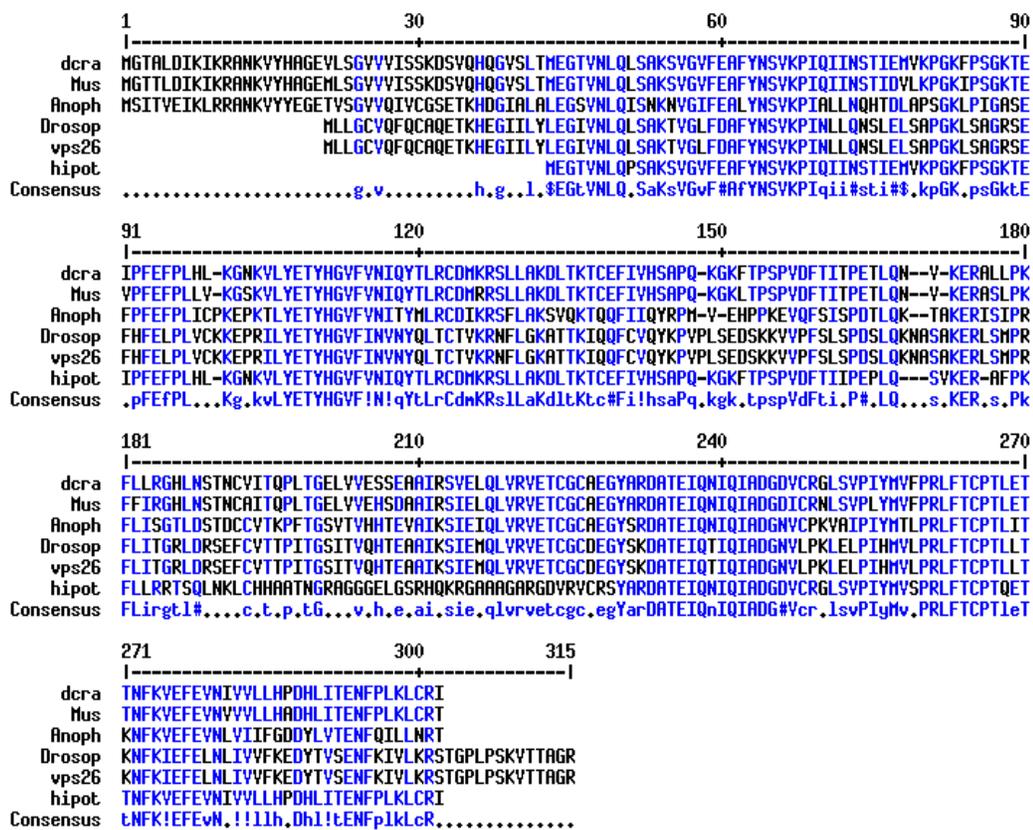


Figura 34. Alinhamento múltiplo de seqüências similares a DCRA realizado no programa *Multalin* usando a matriz *Blosum62-12-2*. A proteína DCRA foi alinhada com *dcrA* de *Mus musculus* (Mus), proteína de *Anopheles gambiae* (Anoph), *Drosophila melanogaster* (Drosop), *Vacuolar sorting protein* (Vsp26) e uma proteína hipotética humana, isto é codificada por um gene conservado, mas com função ainda desconhecida (hipot).

9.3 Predição de domínio

A predição de que DCRA possui um putativo domínio conservado de Vps 26 sugere uma similaridade de função entre essas duas proteínas. As Vps (*vacuolar sorting protein*) são proteínas envolvidas no transporte para o vacúolo.

O gene *Vps26* de levedura é homólogo a *Hβ58* de camundongo, um gene essencial para o completo desenvolvimento embrionário de camundongo (LEE et al., 1992). Segundo EDGAR & POLAK (2000) a expressão de *Hβ58* é particularmente essencial não no início da embriogênese, já que o desenvolvimento das anormalidades inicia-se no 7,5

dias após a fecundação, mas sim após esse período, pois mutações no gene H β 58 causam uma paralisação no desenvolvimento por volta do 9,5 dias.

EDGAR & POLAK (2000) observaram que as proteínas Vps26 e H β 58 são homólogas a uma grande família de transportadores ABC (*ATP Binding Cassete*) das adenosinas trifosfatases de tráfegos (ATPases) bacterianas responsáveis pelo transporte de pequenas moléculas através da membrana. Esses autores propuseram que a hidrólise de ATP mediada pela Vps26 e H β 58 seria utilizada para o revestimento interno do retrômero. Essa hipótese está baseada no fato de que essas proteínas são homólogas a HisP (uma ATPase bacteriana envolvida no transporte de histidina), além disso, essa homologia permitiu Edgar e Polak identificarem um domínio de ligação a ATP (motivo GX β XXG) localizado no N-terminal de H β 58 entre os resíduos 45-50.

Análises *in silico* da seqüência de aminoácidos da DCRA indicam um putativo domínio de ligação a ATP nos resíduos 82-89 (Tabela 3) reforçando a hipótese de que DCRA e Vps26 podem desempenhar função similar.

Apesar das Vps serem bem conhecidas em levedura, em mamíferos a função de proteínas homólogas a Vps ainda não está completamente elucidada.

KIM et al. (2001) identificaram recentemente duas homólogas as Vps em humanos (hVps1 e hVps18) e sugeriram que as hVps estão associadas com a via endossomo/lisossomo e localizam-se nos compartimentos citossólico e frações de membrana. As hVps são altamente expressas em coração, músculo esquelético, rim, fígado, placenta e em várias linhagens celulares sugerindo que estas proteínas exercem uma função fisiológica comum (KIM et al., 2001 op cit.).

Experimentos de purificação, realizados em cromatografia de exclusão molecular, mostram que hVps35, hVps29 e hVps26 tanto endógenas como recombinantes co-eluem como um grande complexo de aproximadamente 240-440 kDa, dando suporte a hipótese de similaridade estrutural entre as proteínas de levedura e as de humanos (HAFT et al., 2000).

Segundo HAFT et al. (2000 op cit.) na ausência de hVps35 as proteínas hVps29 e hVps26 não se associam ao complexo sugerindo que em mamíferos também há a formação de um putativo complexo multimérico semelhante ao de leveduras.

A verdadeira função de Vps26 ainda é desconhecida, porém os dados de REDDY & SEAMAN (2001) sugerem que Vps26 promove as interações entre o componente de seleção de carga de Vps35p e os componentes estruturais Vps5p/Vps17 (Figura 35).

As proteínas Vps5p e Vps17 possuem homologia com os genes SNX1 e SNX2 (*Sorting Nexins*) de mamíferos (EDGAR & POLAK 2000, op cit.). O SNX1 liga-se à cauda citoplasmática do fator de crescimento epidermal (EGFR) que provavelmente tem a função de direcionar o EGFR para o lisossomo (KURTEN, CADENA & GILL, 1996).

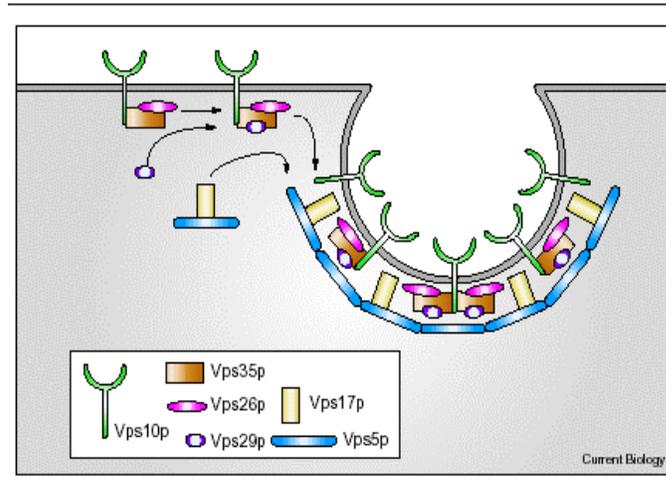


Figura 35. Representação esquemática da formação do retrômero. O domínio citoplasmático de Vps10 de levedura é reconhecido por um complexo de multisubunidades denominado de retrômero, o qual transporta Vps10 para o complexo de Golgi. O retrômero contém subunidades de Vps35p, Vps26p, Vps29p, Vps17p e Vps5p. Vps35 é uma subunidade do retrômero que seleciona a proteína cargo a ser transportada. Cada subunidade deste retrômero existe dentro de dois subcomplexos, e acredita-se que a auto-reunião de cada subcomplexo de Vps35p e Vps17p coordena os complexos que contém Vps35p, Vps29p, Vps26 e moléculas cargos e conseqüentemente a formação da vesícula (Fonte: PFEFFER, 2001).

9.4 Modificações pós-traducionais

A análise *in silico* da seqüência de aminoácidos da DCRA sugeriu 2 prováveis sítios de N-glicosilação, 5 prováveis sítios de fosforilação de caseína quinase C, 4 prováveis sítios fosforilação de caseína quinase II, 4 prováveis sítios N-miristoilação, 1 provável sítio microcorpo C-terminal, 1 provável sítio de ligação ATP/GTP (P-loop), 7

serinas (aminoácidos 29,33,55,74,86,153,201), 4 treoninas (aminoácidos 119,151,160,218) e 4 tirosinas (aminoácidos 63,106,225,251) potencialmente alvos de fosforilação.

Os resíduos de serina, treonina e tirosina quando fosforilados afetam a magnitude dos processos de sinalização celular, no entanto, o significado biológico dessas informações na seqüência da DCRA ainda não é conhecido.

Tabela 3: Predição *in silico* de motivos estruturais na proteína DCRA. A predição dos motivos foi realizada como descrito em métodos.

	<i>Motivos</i>	Aminoácidos
N-glicosilação	NSTI	73-76
	NSTN	183-186
Fosforilação de caseína quinase C	SSK	29-31
	SAK	52-54
	SVK	65-67
	SGK	86-88
	TLR	119-121
Fosforilação de caseína quinase II	TALD	3-6
	SSKD	29-32
	STIE	74-77
	SPVD	153-156
N-miristoilação	GVSLTM	38-43
	GVFEAF	57-62
	GVFVNI	111-116
	GCAEGY	220-225
Microcorpo C-terminal	CRI	295-297
Ligação ATP/GTP (P-loop)	GKFPSKT	82-89

9.5 Topologia

A célula eucariota é constituída por organelas, compartimentos, cada um com uma função específica que exige distintos conjuntos de proteínas e enzimas. A maioria das proteínas é sintetizada sobre ribossomos no citossol, algumas poucas que são sintetizadas pelos ribossomos de mitocôndrias e plastos, e endereçadas a locais específicos na célula devido à presença de uma curta seqüência de aminoácidos denominada de seqüência sinal.

A maioria das proteínas não possui esse sinal e conseqüentemente permanecem no citoplasma como residentes permanentes. Outras, contudo, têm sinais específicos que

direcionam seu transporte do citoplasma para o núcleo, retículo endoplasmático, mitocôndria, peroxissomos ou plastos (em plantas). Além disso, os peptídeos sinais também podem direcionar o transporte de proteínas do retículo endoplasmático para outros destinos na célula como, por exemplo, à secreção para o espaço extracelular.

9.5.1 PSORT II

A predição de sinais e a possível localização subcelular da DCRA foram feitas utilizando-se o programa PSORT II (<http://psort.nibb.ac.jp/form2.html>) e resultou numa predição de 60,9% de possibilidade de ser citoplasmática devido à ausência de sinais para localização da proteína em diferentes organelas.

9.5.2 TargetP

A análise da seqüência de aminoácidos da DCRA não revelou a presença de nenhum peptídeo sinal reforçando a predição de que sua localização seja citoplasmática. O peptídeo sinal (SPs) é responsável pelo endereçamento da proteína para o ER para subsequente transporte por vias secretórias.

9.6 Estrutura secundária

A predição de estrutura secundária, baseada na seqüência primária de aminoácidos da proteína DCRA, foi feita pelo método de ROST & SANDER (1993) que combina informações evolutivas e de rede neural. O resultado indica que DCRA pode ser classificada como uma proteína globular mista com ausência de um domínio compacto composta por 2,7 % de α -hélice; 41,8 % de folhas β e 55,6 % de *loop*. Os resultados da predição de estrutura secundária obtida pelo método McGUFFIN; BRYSON & JONES (2000) são semelhantes aos de ROST & SANDER e podem ser visualizados na Figura 36.

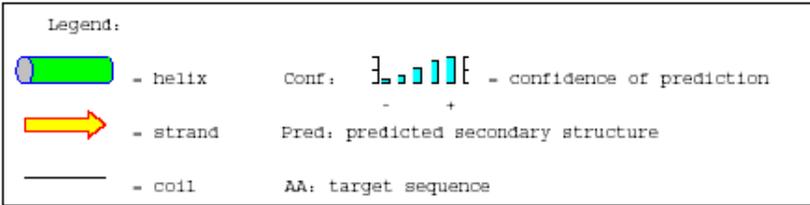
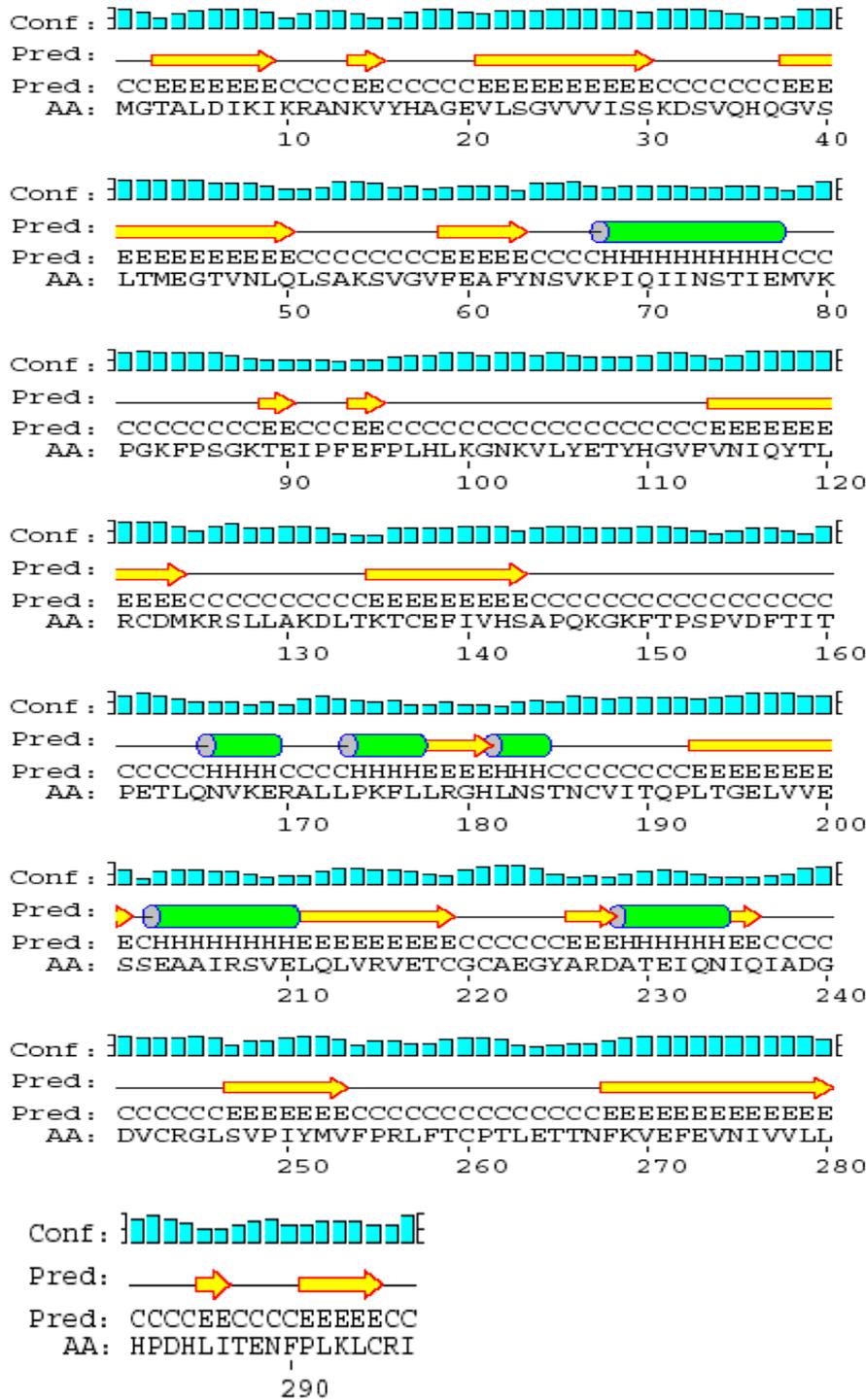


Figura 36. Representação esquemática da estrutura secundária da DCRA obtida pelo método de McGUFFIN; BRYSON & JONES (2000) como descrito em Métodos.

10 Anticorpos

Os anticorpos podem ser úteis para prover informações críticas sobre a localização da proteína dentro de uma célula, para a determinação dos níveis de expressão protéica, entre outros, daí a grande importância da obtenção de anticorpos contra a DCRA.

O antissoro contra a tiorredoxina-DCRA-GFP recombinante foi satisfatoriamente produzido em camundongo e a especificidade do antissoro tiorredoxina-DCRA-GFP foi avaliada usando *Western Blotting* contra GFP e tiorredoxina recombinantes, solução protéica de tiorredoxina-DCRA-GFP parcialmente purificada, extrato protéico de células CHO transfectadas com Flag-DCRA-myc e extrato protéico de células de mamífero completa ou parcialmente desnaturadas. No entanto, o soro imune reconheceu somente a tiorredoxina, tiorredoxina-DCRA, Trx-DCRA-GFP e a Flag-DCRA-myc (Figura 37).

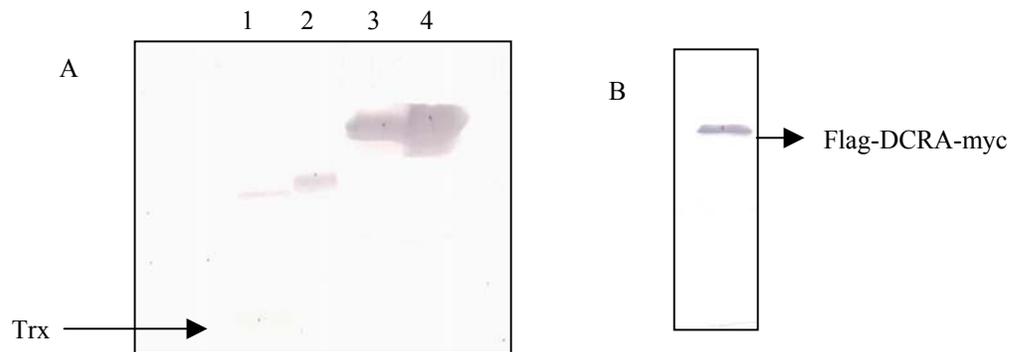


Figura 37. Análise do antissoro através de *Western-Blot*. Membrana de nitrocelulose mostrando em A: (1) o reconhecimento das proteínas Trx; (2) Trx-DCRA; (3,4) Tiorredoxina-DCRA-GFP parcialmente purificada e em (B) reconhecimento da proteína Flag-DCRA-myc.

O não reconhecimento da proteína DCRA endógena no extrato celular pelo antissoro sugere que (1) o teste usado poderia não ter tido a sensibilidade necessária para a detecção da DCRA, (2) a proteína DCRA endógena não é eficientemente sintetizada ou é rapidamente degradada ou (3) o anticorpo obtido poderia não ter especificidade contra a proteína parental. A DCRA humana possui 91% de identidade com a dcra de camundongo e essa seqüência altamente conservada poderia ter estimulado uma pequena resposta imune. Além disso, os anticorpos policlonais são produtos de muitas populações diferentes

de células produtoras de anticorpos e, portanto, apresentam diferente afinidade pelo antígeno.

Uma maneira de contornar a dificuldade em se obter anticorpos com titulação desejável é sintetizar peptídeos baseados na predição do potencial antigênico da proteína alvo. Para gerar anticorpos policlonais os cuidados que devem ser tomados na escolha de tais fragmentos peptídicos de acordo com KOLASKAR & TONGAONKAR (1990) são: (1) peptídeos antigênicos deveriam estar localizados em uma região acessível ao solvente e conter resíduos hidrofílicos e hidrofóbicos. Estes resíduos são raramente encontrados em estruturas ordenadas como α -hélice e folha β (SAMBROOK & RUSSEL, 2001), (2) preferencialmente deve-se selecionar peptídeos em regiões de *loop* que conectam motivos de estrutura secundária evitando-se peptídeos localizados em regiões de hélice, (3) quando possível escolher peptídeos que não pertencem ao N ou C-terminal da proteína, pois se a proteína alvo sofrer modificação pós traducional o anticorpo não mais reconhecerá a proteína alvo (VAN BEIJNUM et al., 2002) e (4) eliminar peptídeos que contenham sítios para N-glicosilação. Mesmo com todos esses cuidados VAN BEIJNUM et al. (2002) observaram que dos 20 anticorpos fago específicos contra as 5 proteínas superexpressas em câncer de cólon-retal apenas 2 reconheceram a putativa proteína parental no tecido primário.

A análise *in silico* da seqüência da DCRA, realizada no programa desenvolvido por KOLASKAR & TONGAONKAR (1990 op cit.) e disponível no servidor <http://mif.dfci.harvard.edu/Tools/antigenic.html>, indica a presença de 11 regiões antigênicas espalhadas ao longo da proteína. Se considerarmos essas 11 regiões mais a presença de *loops* preditos *in silico*, então teoricamente DCRA seria capaz de evocar uma resposta imune. Sabe-se que as regiões de antigenicidade para as proteínas globulares normalmente são atribuídas às regiões de *loop* (FERRIERES, et al., 1998).

Além disso, a presença da GFP na fusão também deveria ser fortemente antigênica uma vez que essa proteína pertence a um cnidário e, portanto, bastante distante do ponto de vista evolutivo.

Conclui-se, portanto, que o anticorpo obtido era capaz de reconhecer a DCRA superexpressa na bactéria, mas não a DCRA endógena por motivos ainda desconhecidos.

11 Localização subcelular da DCRA em células de mamíferos

A identificação e a caracterização da localização subcelular de proteínas são passos essenciais não somente para sugerir a função do gene, mas também para contribuir para o entendimento da função e da regulação celular.

A GFP (*Green fluorescent protein*) uma proteína de 27 kDa é comumente utilizada como proteína repórter para estudos de localização celular por não apresentar qualquer seqüência envolvida em endereçamento celular, dispensar tratamentos que poderiam prejudicar a arquitetura celular tais como a permeabilização e permitir a visualização mais detalhada dos eventos celulares muitas vezes prejudicados pela coloração com anticorpos (*background*). Estas propriedades intrínsecas da GFP somadas a sua fluorescência estável e independente de qualquer fator ou substrato, acrescido do fato dessa proteína não ter mostrado interferência na localização de proteínas (PINES, 1995) a tornam uma atrativa escolha para estudos de processos celulares. Devido ao seu pequeno tamanho, a GFP pode difundir passivamente através do complexo do poro nuclear, por isso é visualizada, em célula de mamífero, uniformemente distribuída entre núcleo e citoplasma (padrão N+C) (BULLIDO et al., 2000).

A proteína DCRA é composta por 297 aminoácidos com uma massa molecular de 33 kDa, possui proporção equitativa de aminoácidos hidrofílicos e hidrofóbicos, mas até o presente momento nada foi descrito sobre sua localização subcelular, função, bem como sua implicação com a SD.

A localização da DCRA em fusão com o C-terminal da GFP (GFP-DCRA) foi predominantemente nuclear em todas as linhagens celulares usadas (CHO-K1, COS-7 e HEK293) (Figura 38) apesar da análise *in silico* da seqüência de aminoácidos da DCRA não ter revelado nenhum sinal clássico de importe para o núcleo (como o NLS SV40 *T-antigen* ou *M9 Domain*), mitocôndriais, vias secretórias bem como sítios de clivagem. Enquanto que células transientemente transfectadas somente com pEFGPC1 ou pEGFPN1 exibiram o padrão N+C, isto é distribuição homogênea na célula (Figura 38).

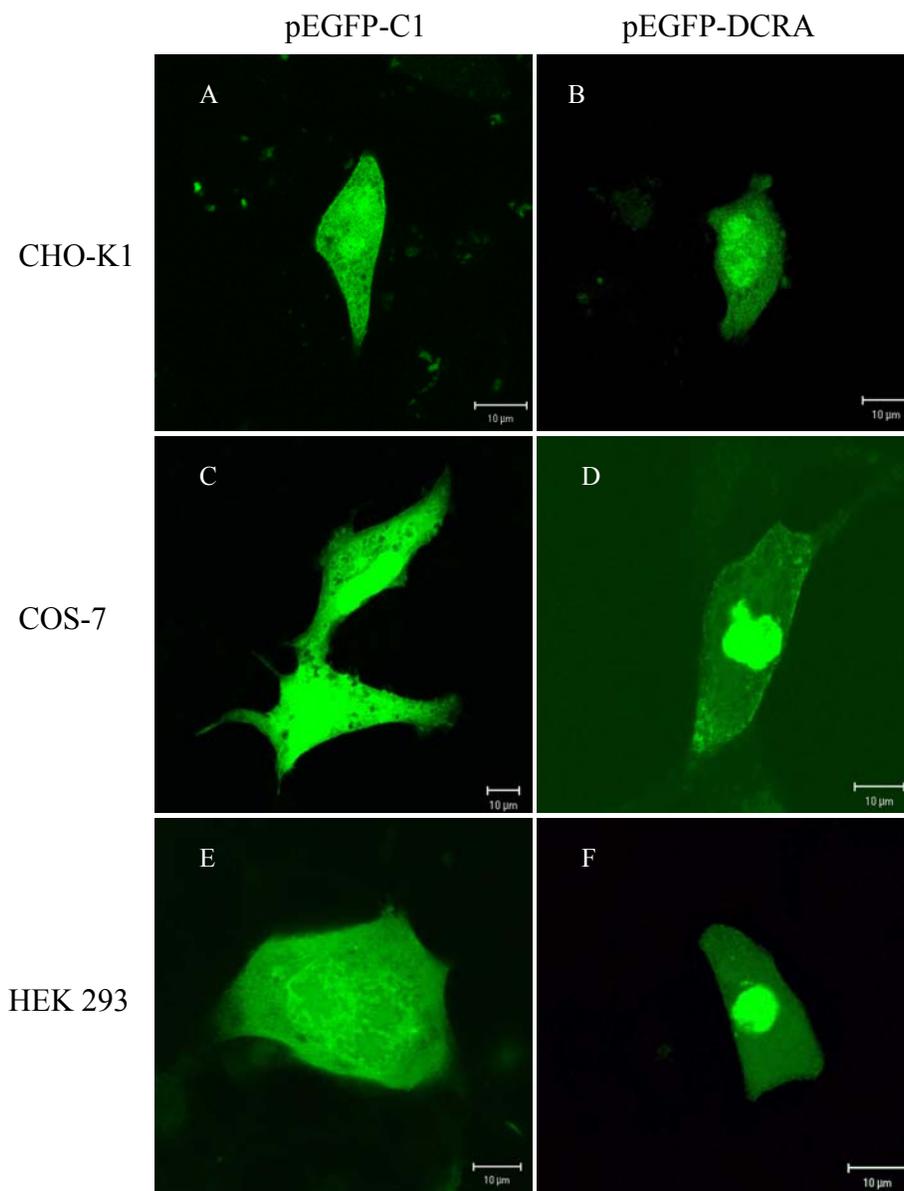


Figura 38. Distribuição subcelular de GFP e GFP-DCRA nas linhagens CHO-K1, COS-7, e HEK 293 transientemente transfectadas. (A,C, E) padrão de distribuição N+C (distribuição homogênea na célula); (B) padrão de distribuição N>C (intensidade mais forte no núcleo do que no citoplasma); (D e F) padrão de distribuição N (forte intensidade no núcleo).

A localização nuclear da GFP-DCRA foi confirmada pela comparação da fluorescência azul emitida pelo DNA nuclear corado com DAPI (Figura 39).

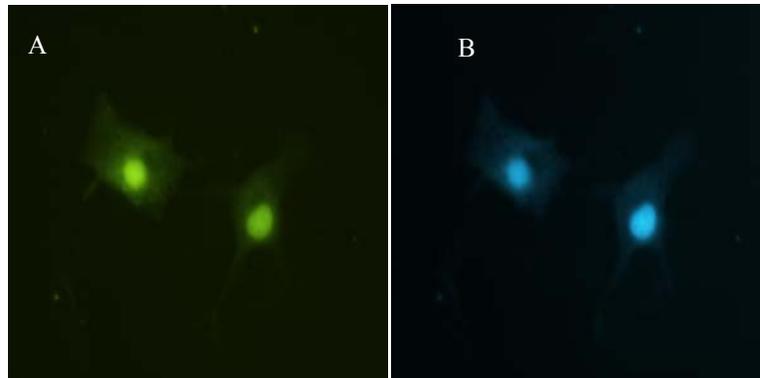


Figura 39. **Distribuição subcelular de GFP-DCRA na linhagem CHO-K1 transientemente transfectada.** A. Célula CHO-K1 expressando pEGFP-DCRA mostrando padrão de distribuição N (forte intensidade no núcleo) B. Célula CHO-K1 expressando pEGFP-DCRA corada com DAPI para indicar a localização do núcleo.

A observação de que a localização de GFP-DCRA em todas as linhagens celulares testadas foi significativamente nuclear com presença citoplasmática (Figura 38B) levantou uma interessante questão sobre a localização celular dessa proteína. Como predições feitas no programa PSORT II sugeriam que se tratava de uma proteína citoplasmática (60,9% de probabilidade) a presença da DCRA no núcleo indicava que poderia haver um sinal nuclear atípico, enquanto que a significativa presença de GFP-DCRA no citoplasma poderia ter sido causada por um motivo hidrofóbico similar a um putativo NES (*nuclear export signal*) ou sinal de retenção citoplasmático.

Sabe-se que o transporte de macromoléculas entre o citoplasma e o núcleo é um processo muito importante para a manutenção da vida e ocorre através de interações diretas entre proteínas específicas que reconhecem o sinal de localização nuclear (NLS).

A localização nuclear das proteínas pode ocorrer através de dois mecanismos, difusão passiva ou transporte ativo (GUIOCHON-MANTEL et al., 1991). No entanto, a maioria das proteínas entra no núcleo através de um processo que envolve gasto de energia e reconhecimento de um sinal de localização nuclear específico.

O NLS é formado por um conjunto de resíduos de aminoácidos básicos (monopartido) ou por dois conjuntos separados por 10 a 12 resíduos de aminoácidos

(bipartido). É reconhecido por um receptor complexo formado pelas importinas α e β . A importina α é uma proteína adaptadora constituída por um domínio de ligação a importina β e um domínio de ligação ao C terminal do NLS. As importinas β são importantes por auxiliarem o importe das importinas α através do poro nuclear (NPC).

Em adição a esses sinais, tem sido mostrado que algumas proteínas possuem sinais de retenção nuclear (NRS) ou sinal de retenção citoplasmático (CRS) que contribuem para a determinação da localização celular (BULLIDO et al., 2000). Há várias maneiras específicas pelo qual o transporte nuclear pode ser regulado, sendo que um dos mais bem conhecidos é a retenção citoplasmática no qual a proteína endereçada ao núcleo é retida no citoplasma pela ligação a uma “estrutura âncora” ou a um parceiro que seqüestra o NLS (BIRD et al., 2001).

JANS et al. (2000) acreditam que o provável reconhecimento de seqüências alvo representa o ponto chave do transporte nuclear. A eficiência desse reconhecimento pode ser modificada pela presença de múltiplos NLS, pela modificação de seqüências flanqueadoras e pela acessibilidade ao NLS pela maquinaria de importe celular (SHANG et al., 1998; GORLICH & MATTAJ, 1996).

Assim com objetivo de testar a hipótese da presença de um putativo sinal NLS atípico formas truncadas com 42, 60, 118, 212, 110 aminoácidos a partir da porção N-terminal e com 45 aminoácidos da porção C-terminal da DCRA foram geradas com base em predição de motivo (domínio estrutural pequeno que pode ser reconhecido em várias proteínas), domínio (porção contínua discreta da seqüência de aminoácidos que pode ser relacionado a uma função particular) e de serinas, treoninas e tirosinas que poderiam ser sítios de fosforilação e assim contribuir para o importe nuclear dessa proteína (ver Tabela 3).

Cada um desses fragmentos tinha como último aminoácido, respectivamente treonina (T), ácido glutâmico (E), glutamina (Q) e isoleucina (I) (Figura 4). No entanto, a transfecção celular das construções formas truncadas da GFP-DCRA apresentaram padrão de distribuição exclusivamente citoplasmático em todas as linhagens celulares testadas descartando, portanto, a possibilidade de NLS atípico, NES ou sinal de retenção citoplasmático (Figura 40).

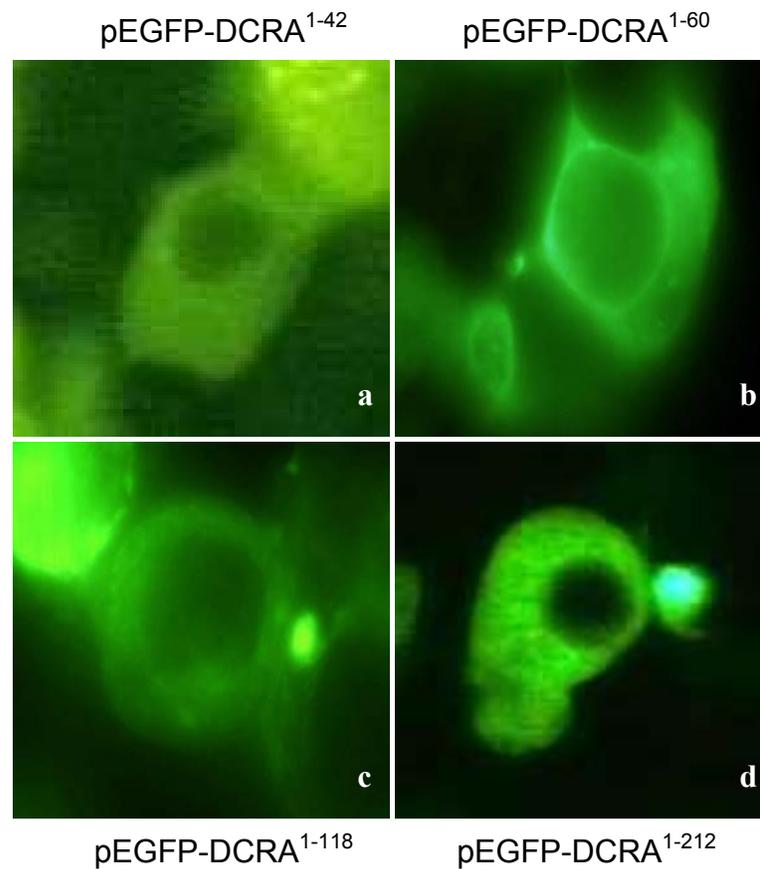


Figura 40. Distribuição subcelular das formas truncadas da GFP-DCRA em células CHO-K1 transientemente transfectadas. Em a, b, c e d são mostrados o padrão de distribuição C (forte intensidade no citoplasma).

Além do mais, as construções que geraram formas truncadas da GFP-DCRA foram realizadas de modo que os “cortes” na proteína não mutassem nenhuma possível serina ou tirosina fosforilada e que os motivos preditos permanecessem preservados, o que indica que tais considerações não foram responsáveis pela localização nuclear da GFP-DCRA.

Para testar qualquer possibilidade de interferência da GFP-tag na localização celular da DCRA a proteína fluorescente foi fusionada no N-terminal da DCRA gerando DCRA-GFP e a DCRA foi fundida a dois epítomos gerando FLAG-DCRA-myc.

Segundo MARGOLIN (2000) o domínio estrutural ou a topologia da proteína em questão determinam se a GFP será fundida no C ou no N-terminal da proteína alvo, talvez por isso, para confirmar se a localização celular era independente da posição da GFP-tag HOJA, WAHLESTED & HÖÖG (2000) clonaram a GFP no N e no C-terminal sempre que na seqüência de aminoácidos das proteínas analisadas havia resíduos hidrofóbicos, estrutura α -hélice ou domínio de membrana.

A análise *in silico* da sequência de aminoácidos da DCRA não detectou resíduos hidrofóbicos, estrutura α -hélice ou domínio de membrana no amino ou no carboxi-terminal que pudessem causar uma localização espúria da DCRA se acaso a sequência codificadora da GFP fosse clonada em fusão somente com uma das extremidade do DCRA. Contudo, a localização das proteínas DCRA-GFP e FLAG-DCRA-*myc* mostrou-se citoplasmática com ausência total no núcleo (Figura 41) confirmando que a proteína DCRA de fato é citoplasmática e quando fusionada ao C-terminal da GFP localiza-se predominantemente no núcleo.

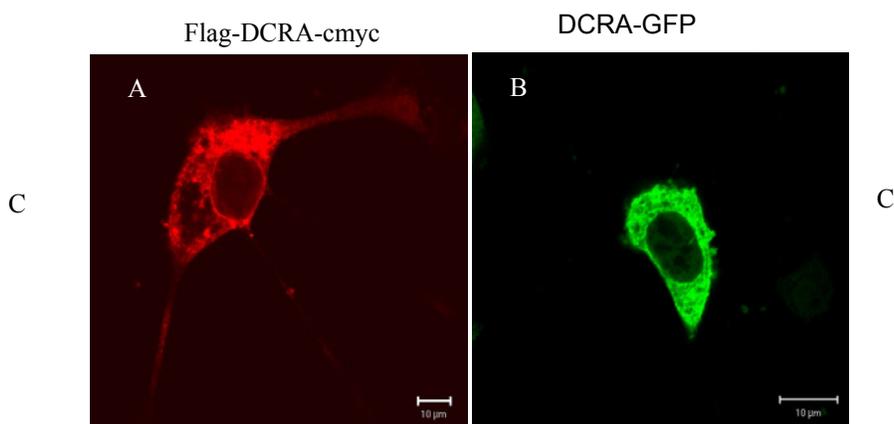


Figura 41. Distribuição subcelular de Flag-DCRA-cmyc e DCRA-GFP em células transientemente transfectadas. Em (A) a célula COS-7 foi tratada com anticorpo anti-*myc* conjugado com Cy3 para detectar DCRA-*myc* e exibe o padrão C (forte intensidade no citoplasma) e (B) DCRA-GFP, na linhagem CHO-K1, também exibe o padrão de distribuição C.

Possíveis explicações para a localização nuclear da GFP-DCRA são: (1) a fusão da GFP no N-terminal da DCRA de alguma maneira possibilitou seu transporte para o núcleo uma vez que cerca de 10% das proteínas com ausência de um core peptídico NLS típico podem entrar no núcleo co-transportadas por um “fator” que possua um NLS (BOULIKAS, 1994) e (2) a fusão GFP-DCRA poderia ser instável na célula. Em suporte a essa explicação está a observação de que um número significativo de fusões com GFP é instável na célula, resultando em um *background* de polipeptídeos de GFP que não estão associados com a proteína de interesse, ou de fusões que não fluorescem (PRESCOTT et al., 1999).

A transaldorase uma proteína presumidamente citoplasmática foi localizada no núcleo quando em fusão com a GFP e embora experimentos posteriores tenham confirmado esse resultado, não está claro porque transaldorase-GFP foi localizada no núcleo (MISAWA, et al., 2000).

Portanto, é aconselhável fusionar a GFP no N e no C-terminal da proteína alvo e confirmar o resultado com imuno-citoquímica para descartar qualquer possibilidade de artefato.

Nossos estudos revelam que DCRA é uma proteína que está no compartimento citoplasmático da célula, indicam a provável existência de um domínio de Vps26 e de sítios de fosforilação de significado ainda desconhecido. A presença desse domínio sugere um possível envolvimento da DCRA no transporte endossomo-Golgi e qualquer desequilíbrio nesta homeostase celular causada pela desregulação da expressão do DCRA poderiam contribuir com a variabilidade fenotípica da SD. Juntos esses dados contribuem para melhor conhecimento da DCRA e estimulam futuras análises para o entendimento da função celular desta proteína, bem como sua implicação na etiologia da Síndrome de Down.

12 Conclusões

- 1- Somente foi possível a expressão da DCRA na forma solúvel quando o gene que codifica esta proteína foi clonado em fusão com seqüências que codificam respectivamente o C-terminal da tiorredoxina e o N-terminal da GFP.
- 2- A retirada dos 42 aminoácidos do N-terminal da DCRA resultou na expressão do produto Trx-DCRA. No entanto, essa fusão era expressa na forma de corpos de inclusão.
- 3- Apesar da obtenção de proteína solúvel a purificação não resultou em proteína em quantidade e grau de pureza adequada para os ensaios de cristalização.
- 4- Análises *in silico* da estrutura primária da DCRA indicaram um putativo domínio de Vps26.
- 5- A seqüência de aminoácidos da DCRA possui 57% e 49% de identidade respectivamente com uma proteína de *Anopheles gambiae* e de *Drosophila melanogaster*, ambas de função desconhecida.
- 6- O antissoro para reconhecimento da DCRA foi satisfatoriamente obtido em camundongo, reconhecendo as proteínas recombinantes tiorredoxina e DCRA. Porém não reconhecendo a DCRA endógena no extrato protéico de célula de mamífero.
- 7- A proteína GFP, por motivos ainda desconhecidos, foi localizada no núcleo quando em fusão com o N-terminal da DCRA.
- 8- A DCRA é uma proteína citoplasmática.
- 9- A fusão da DSCR8 ao C-terminal da Trx aliada às características da linhagem Rosetta permitiram a obtenção de proteína na fração solúvel.

13. Anexos

13.1 Mapas de plasmídeos

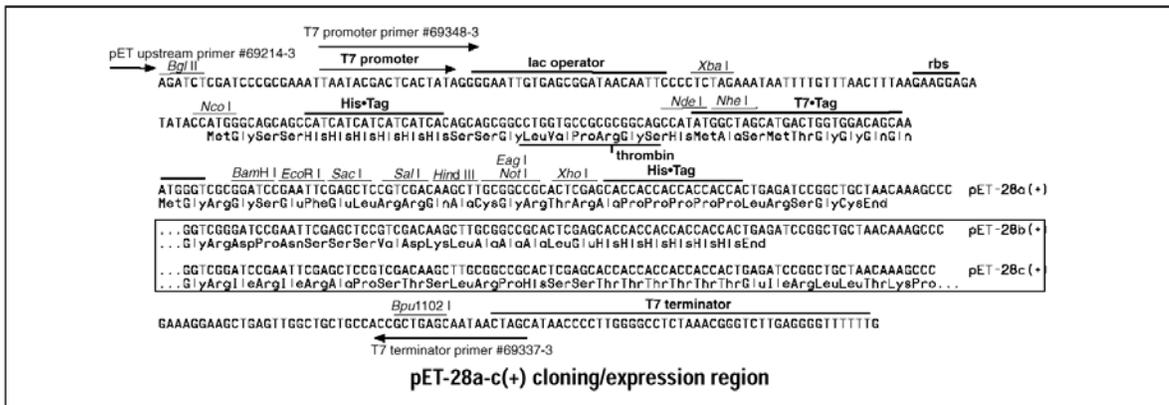
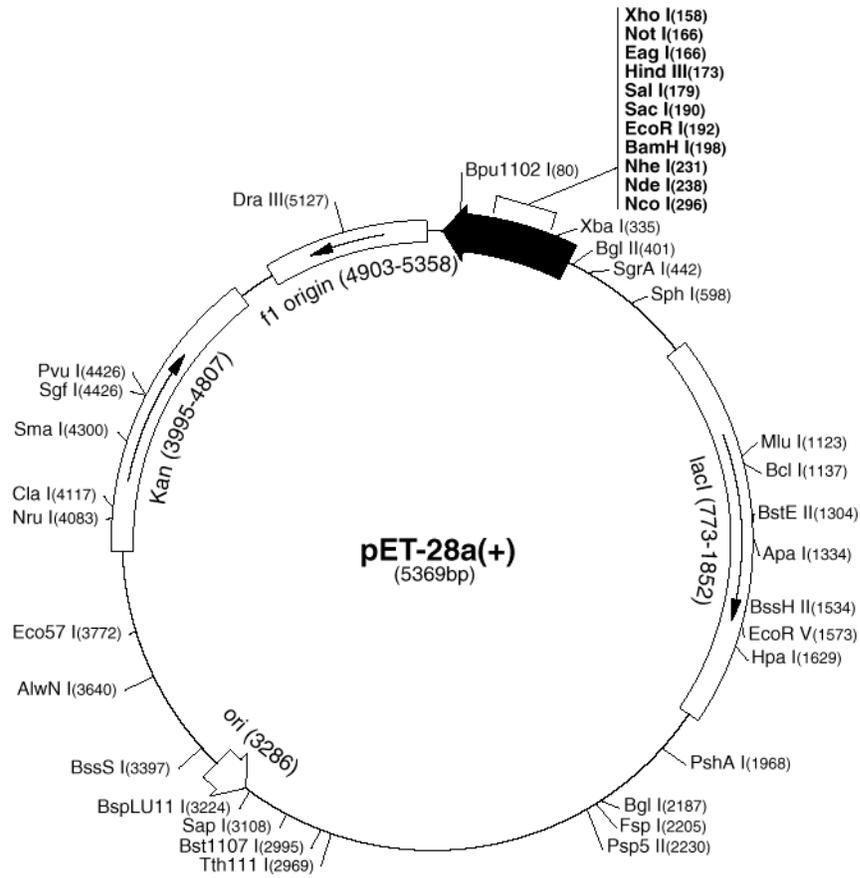


Figura 42. Vektor pET28a (Novagen). Este vetor carrega um His Tag e sítio de clivagem para trombina.

Figura 43. Vetor pET32a (Novagen). Este vetor carrega um His *Tag* e possibilita a clonagem em fase com a seqüência codificadora da Trx.

Figura 44. Vetor pET37b (Novagen). Este vetor carrega um His *Tag* e possibilita a clonagem em fase com a seqüência da CBD.

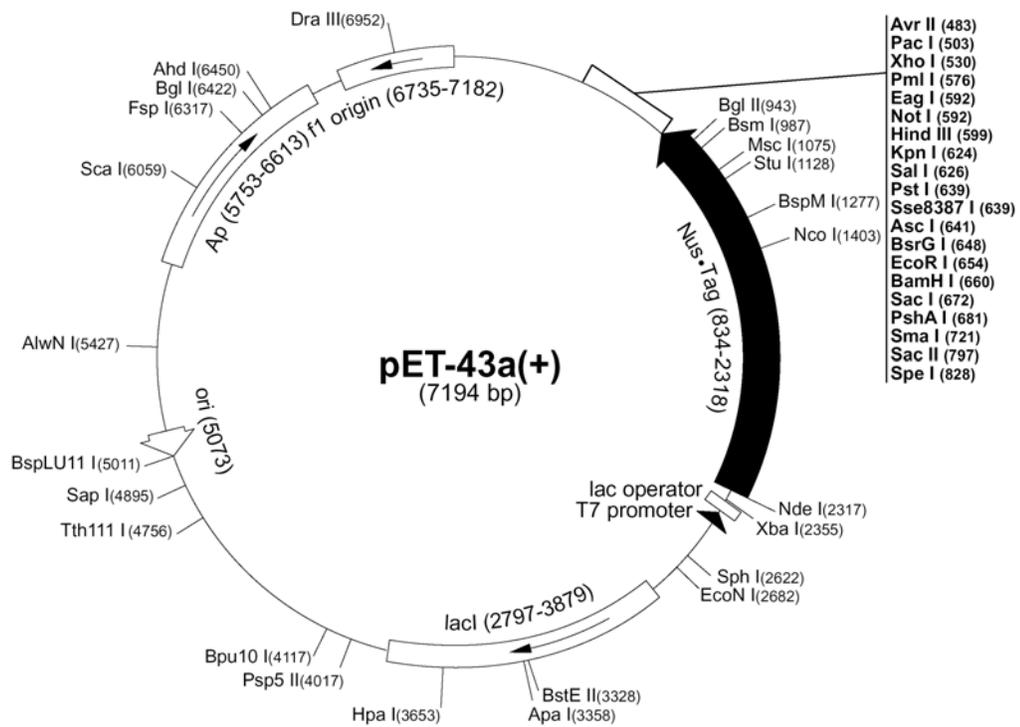
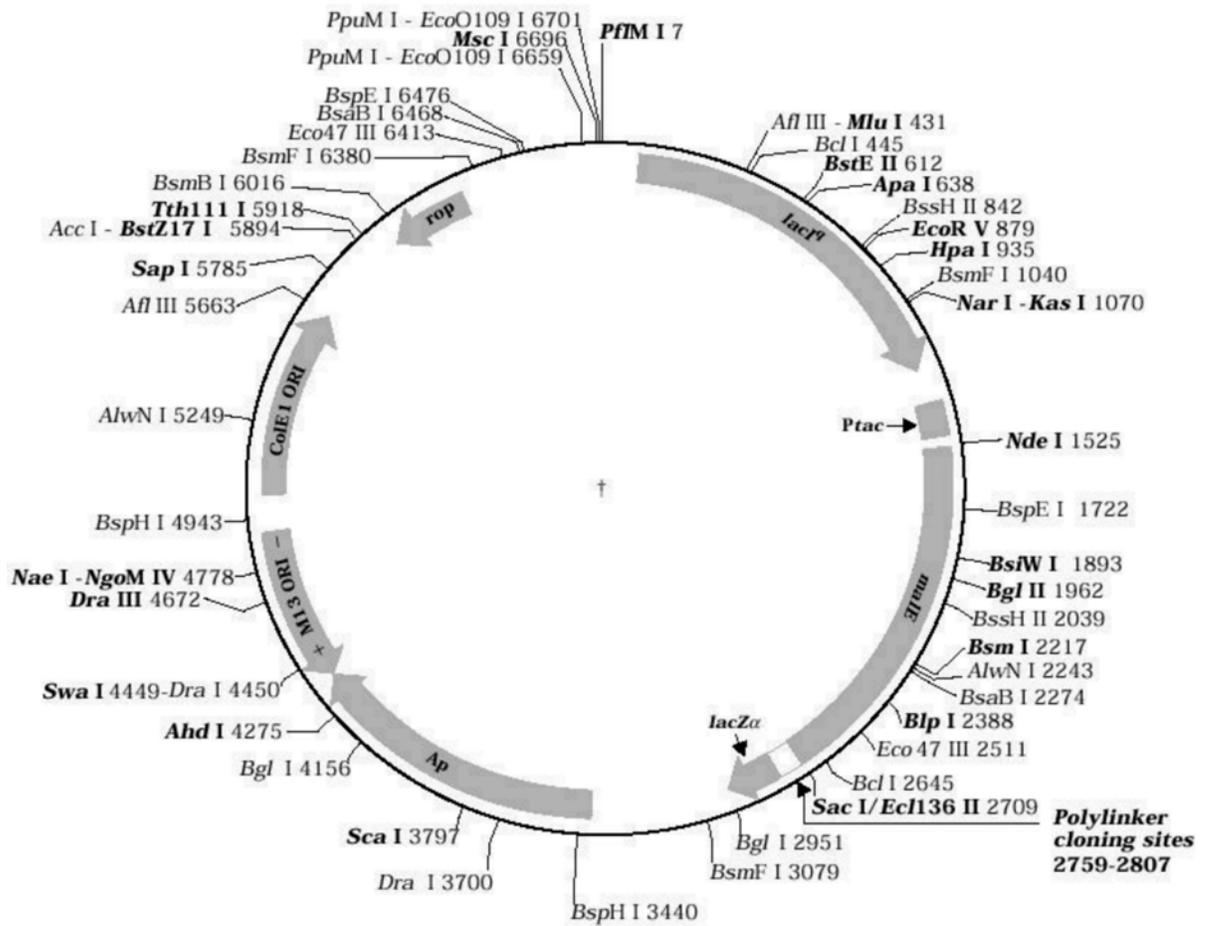


Figura 46. Vetor pET43a (Novagen). Este vetor carrega um His Tag e possibilita a clonagem em fase com a seqüência codificadora da NusA.

.....
 pTYB2 Met Ala Ser Ser Arg Val Asp Gly Gly Arg Glu Phe Leu Glu Pro Gly Cys1 Intein →
 CAT ATG GCT AGC TCG CGA GTC GAC GGC GGC CGC GAA TTC CTC GAG CCC GGG TGC
Nde I *Nhe* I *Nru* I *Sal* I *Not* I *Eco*RI *Xho* I *Sma* I

Figura 47. Vetor pTBY2 (NEB). Este foi utilizado para clonagem do gene DCRA e possibilita a clonagem em fase com a seqüência codificadora da inteína.



pMAL-c2X, pMAL-p2X Polylinker

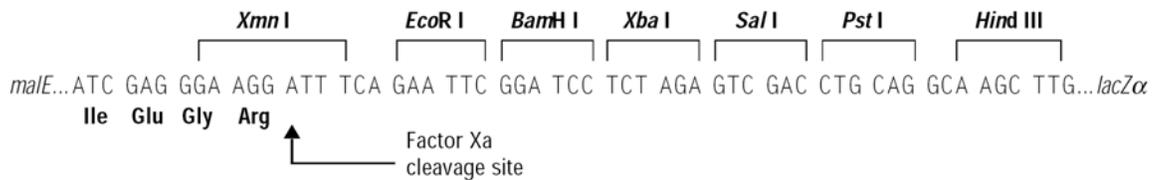


Figura 48. Vetor pMal-c2x (NEB). Este foi utilizado para clonagem do gene DCRA e possibilita a clonagem em fase com a seqüência codificadora da MBP.

Figura 49. Vetor pBad (Invitrogen). Este vetor carrega um His *Tag* e foi utilizado para clonagem do gene DCRA.

Figura 50. Vetor pCX-TOPO (Invitrogen). Este foi utilizado para clonagem do gene DCRA e possibilita a clonagem em fase com a seqüência codificadora da RsaA.

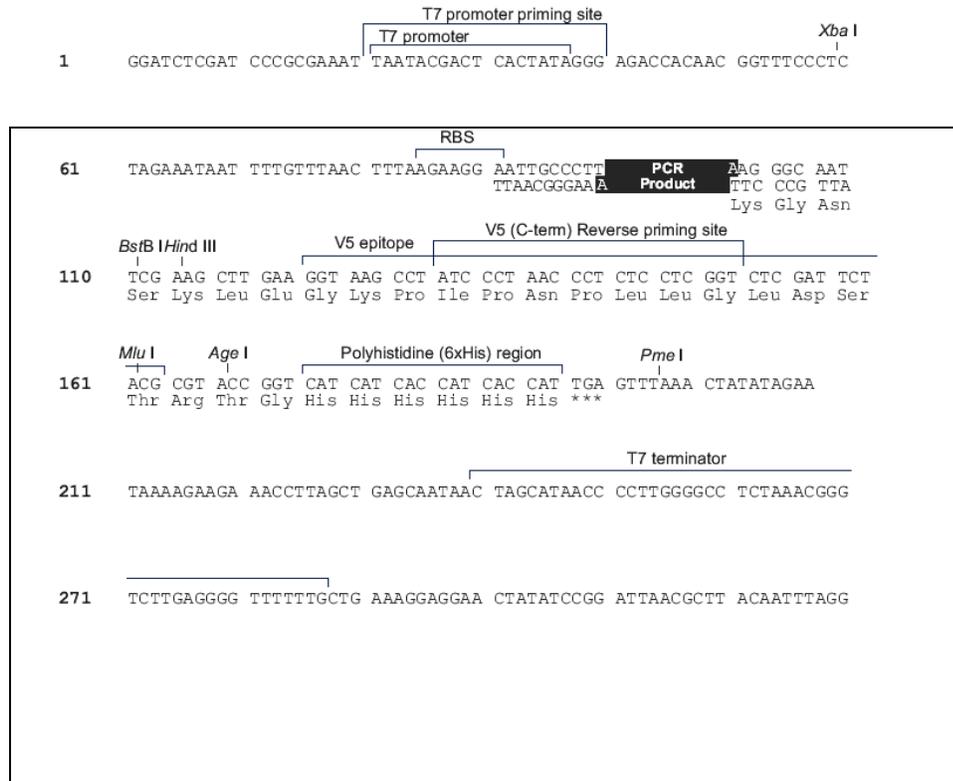


Figura 51. Vetor pCRT7/CT TOPO (Invitrogen). Este vetor carrega um His Tag e foi utilizado para clonagem do gene DCRA.

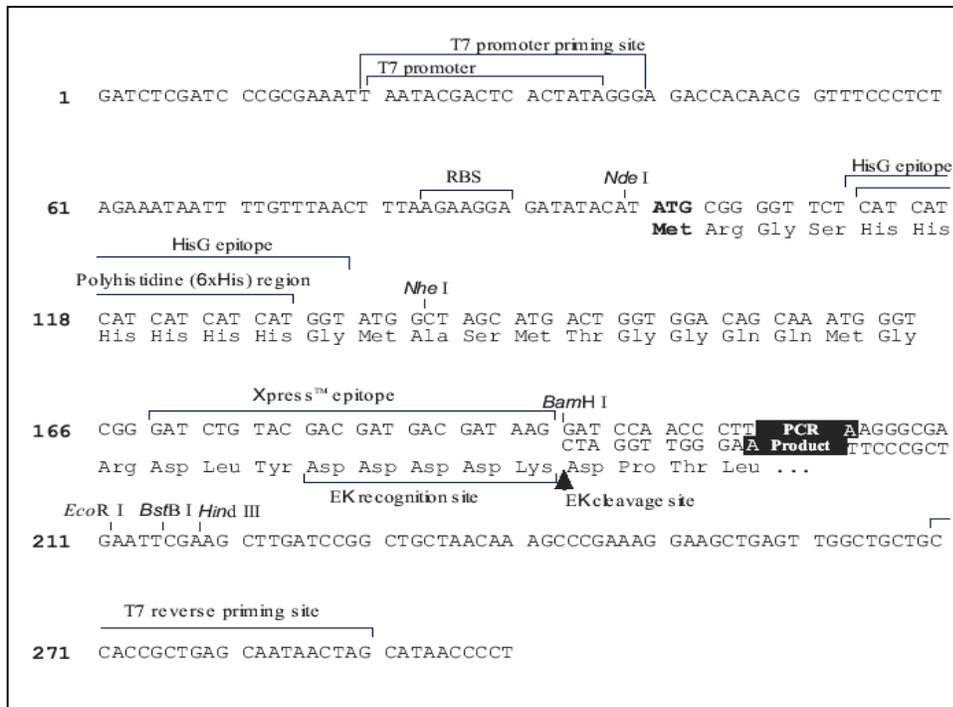


Figura 52. Vetor pCRT7/NT TOPO (Invitrogen). Este vetor carrega um His Tag e foi utilizado para clonagem do gene DCRA.

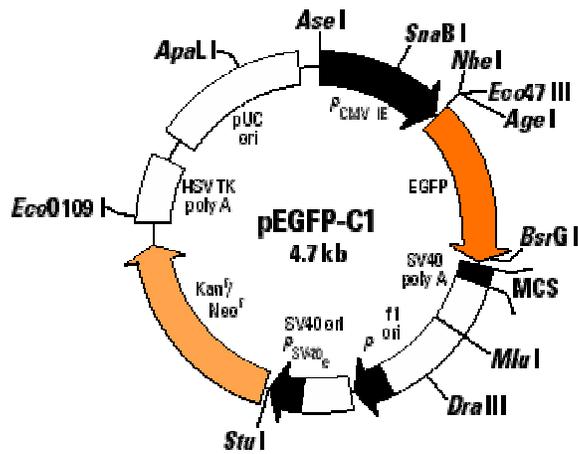


Figura 53. Vetor pEGFP-C1 (Clontech). Este vetor possibilita a clonagem em fase com a seqüência codificadora do C-terminal da GFP.

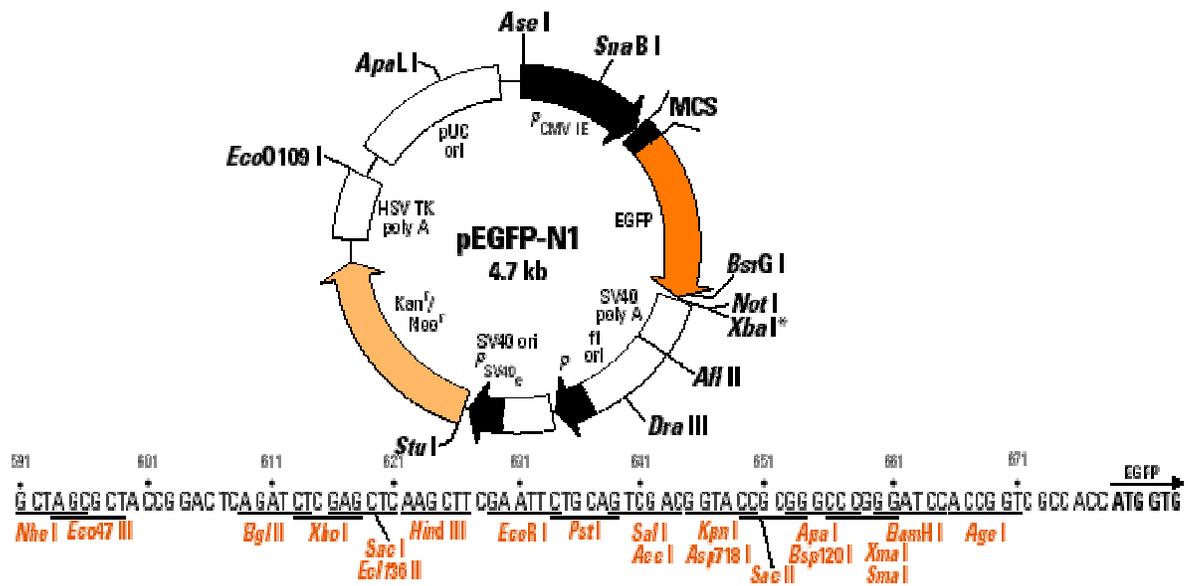


Figura 54. Vetor pEGFP-N1 (Clontech). Este vetor possibilita a clonagem em fase com a seqüência codificadora do N-terminal da GFP.

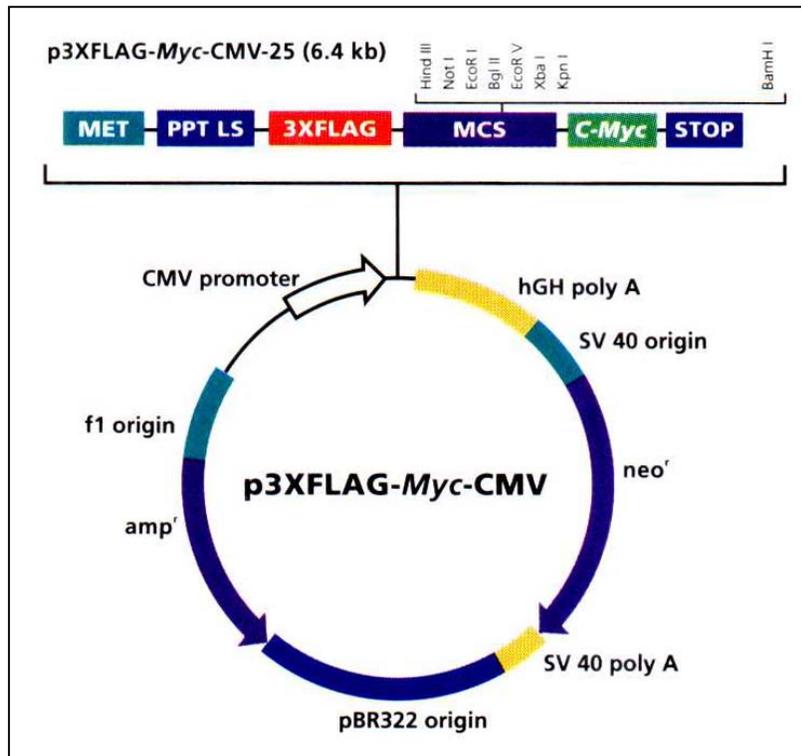
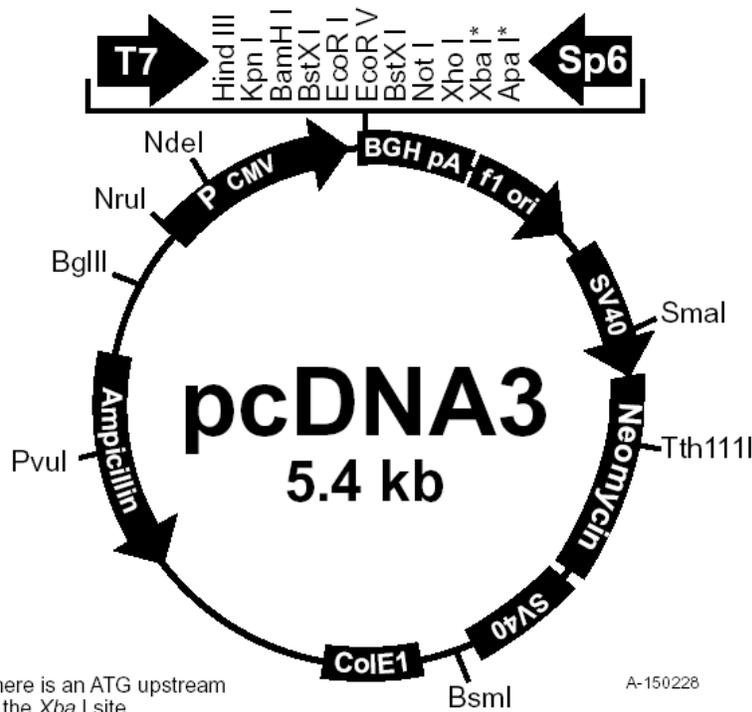


Figura 55. Vetor p3XFLAG-Myc-CMV (Sigma). Este vetor possibilita a clonagem em fase com o a seqüências codificadoras dos epítotos *Flag* e *c-Myc*.



```

224   ATATACGCGT TGACATTGAT TATTGACTAG TTATTAATAG TAATCAATTA CGGGGTCATT
284   AGTTCATAGC CCATATATGG AGTTCGCGCT TACATAACTT ACGGTAAATG GCCCGCCTGG
344   CTGACCGCCC AACGACCCCC GCCCATTGAC GTCAATAATG ACGTATGTTC CCATAGTAAC
404   GCCAATAGGG ACTTTCCATT GACGTCAATG GGTGGACTAT TTACGGTAAA CTGCCCACTT
464   GGCAGTACAT CAAGTGTATC ATATGCCAAG TACGCCCCCT ATTGACGTCA ATGACGGTAA
524   ATGGCCCCGC TGGCATTATG CCCAGTACAT GACCTTATGG GACTTTCCTA CTTGGCAGTA
584   CATCTACGTA TTAGTCATCG CTATTACCAT GGTGATGCGG TTTTGGCAGT ACATCAATGG
644   GCGTGGATAG CGGTTTGACT CACGGGGATT TCCAAGTCTC CACCCCATTTG ACGTCAATGG
704   GAGTTTGTTT TGGCACAAA ATCAACGGGA CTTTCCAAA TGTCGTAACA ACTCCGCCCC
764   ATTGACGCAA ATGGGCGGTA GGCCTGTACG GTGGGAGGTC TATATAAGCA GAGCTCTCTG
824   GCTAACTAGA GAACCCACTG CTTACTGGCT TATCGAAATT AATACGACTC ACTATAGGGA
884   GACCCAAGCT TGGTACCGAG CTCGGATCCA CTAGTAACGG CCGCCAGTGT GCTGGAATTC
944   TGCAGATATC CATCACACTG GCGGCCGCTC GAGCATGCAT CTAGAGGGCC CTATTCTATA
1004  GTGTCACCTA AATGCTAGAG CTCGCTGATC AGCCTCGACT GTGCCTCTA GTTGCCAGCC
1064  ATCTGTGTGT TGCCCTCCC CCGTGCCTTC CTTGACCCTG GAAGGTGCCA CTCCCACTGT
1124  CCTTCCTAA TAAATGAGG AAATTGCAT

```

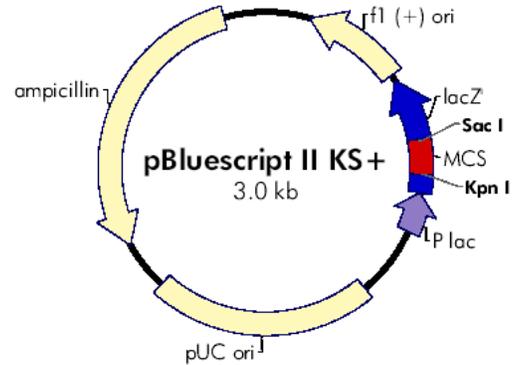
Annotations in the sequence:

- 5' end of hCMV promoter/enhancer (arrow pointing to TGACATTGAT)
- enhancer region (5' end) (bracket above AGTTCGCGCT)
- enhancer region (3' end) (bracket above TGGCACAAA)
- CAAT (bracket above ATGGGCGGTA)
- putative transcriptional start (arrow above ATGGGCGGTA)
- TATA (bracket above TATATAAGCA)
- 3' end of hCMV (arrow pointing to GAGCTCTCTG)
- 17 promoter (bracket above AATACGACTC)
- Hind III, Kpn I, BamH I, Bst XI, Eco R I (restriction sites above 884-900)
- Eco R V, Bst X I, Not I, Xho I, Xba I, Apa I (restriction sites above 900-920)
- Sp6 promoter (bracket above TGCAGATATC)
- BGH polyA (bracket above TGGTACCGAG)

NOTE: There is an ATG upstream of the *Xba* I and *Apa* I sites

Figura 56. Vetor pcDNA3 (Invitrogen). Este vetor é utilizado para expressão em células de mamífero.

f1 (+) origin 135–441
 β -galactosidase α -fragment 460–816
multiple cloning site 653–760
lac promoter 817–938
pUC origin 1158–1825
ampicillin resistance (*bla*) ORF 1976–2833



pBluescript II KS (+/-) Multiple Cloning Site Region
 (sequence shown 598–826)

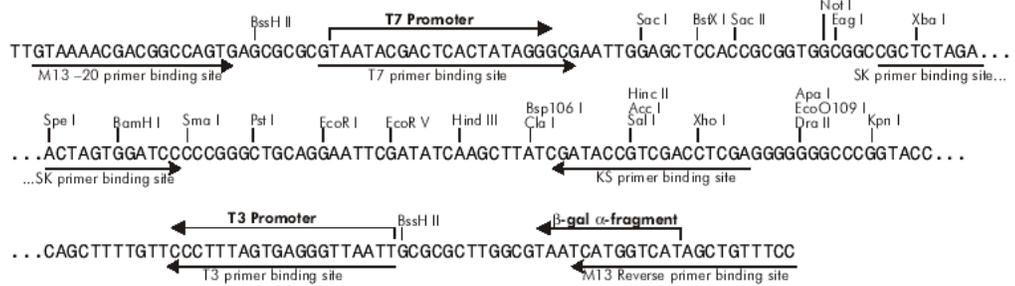


Figura 57. Vetor pBluescript II KS⁺ (Stratagene). Este vetor foi utilizado para a clonagem do gene DCRA.

13.2 Meios de cultura

Meio PYE

0,2% de peptona

0,1% de extrato de levedura

0,02% $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$

0,01% $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ (necessário para o crescimento ótimo da *Caulobacter*)

- 1- Para 1 L, combinar 2 g de peptona, 1 g de extrato de levedura, 1 mL de solução de $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ a 20%, e 1 mL de $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ a 10% em 1 L de água Milli Q. Para meio sólido adicionar 1g de agar por litro.
- 2- Autoclavar por 20 minutos.
- 3- Esfriar o meio a 55° C e adicionar 2 $\mu\text{g}/\text{mL}$ de cloranfenicol (se necessário).
- 4- Armazenar a temperatura ambiente.

Meio M11

- 1- Para preparar 1 L de M11, adicionar a 925 mL de água Milli Q: 5 mL de imidazol 1M, pH 7,0 e 1 mL tampão fosfato de potássio, pH 6,8.
- 2- Autoclavar por 20 minutos.
- 3- Quando a solução atingir 55° C adicionar: 15 mL de glicose a 10%, 15 mL de glutamato a 10%; 1,25 mL de NaCl 100 mM e 2,5 mL da Base Mineral de Hutners'.
- 4- Armazenar a 4° C.

Base Mineral de Hutner's

Ácido nitrilo acético (NTA) 20g

$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 54,5 g

$\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ a 10% 66,7 mL

$(\text{NH}_4)_6 \text{Mo}_7\text{O}_{24} \cdot 4\text{H}_2\text{O}$ 18,5 mg

$\text{FeSO}_4 \cdot \text{H}_2\text{O}$ 198 mg (323,7 mg se for $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$)

Metais 44 100 mL

- 1- Para 1 L, adicionar o NTA em 750 mL de água Milli Q e vagarosamente adicionar 14,6 g de pastilhas de KOH.
- 2- Adicionar os demais componentes na ordem em que são listados.
- 3- Ajustar o pH para 6,8 com HOH 5N
- 4- Completar o volume para 1L.
- 5- Autoclavar por 20 minutos.

Metais 44

EDTA	2,5 g
ZnSO ₄ .7H ₂ O	10,95 g
FeSO ₄ .7H ₂ O	5 g
MnSO ₄ .H ₂ O	1,54 g
CuSO ₄ .5H ₂ O	392 g
Co(NO ₃) ₂ .6H ₂ O	250 mg
Na ₂ B ₄ O ₇ .10H ₂ O	177 mg (200 mg se for Na ₂ B ₄ O ₇ .8H ₂ O)

- 1- Para 1 L adicionar EDTA em 800 mL de água Milli Q.
- 2- Dissolver os demais reagentes na ordem em que são listados. **Não adicionar o próximo reagente até que os componentes anteriores tenham sido dissolvidos completamente.** Adicionar gotas de H₂SO₄ para retardar a precipitação.
- 3- Completar o volume para 1L. **Não autoclavar.** Armazenar a 4° C.

14 Referências Bibliográficas

ALLISON, D. B. et al. Decreased resting metabolic rate among persons with Down Syndrome. **Internat. J. Obesity**. **19**: 858-61, 1995.

ANDRADE, M. A.; O'DONOGHUES. I.; ROST, B. Adaptation of protein surfaces to subcellular location. **J. Mol. Biol.** **276**: 517-525, 1998.

ANTONARAKIS S. E. et al. Differential gene expression studies to explore the molecular pathophysiology of Down Syndrome. **Brain Research Reviews**. **36**: 265-274. 2001.

ANTONARAKIS, S. E. 10 Years of genomics, chromosome 21, and Down Syndrome. **Genomics** **51**: 1-16, 1998.

ANTONARAKIS, S. E. Human chromosome 21: genome mapping and exploration circa 1993. **Trends Genet.** **9**: 142-148, 1993.

ANTONARAKIS, S. E. Parental origin of the extra chromosome in trisomy 21 as indicated by analysis of DNA polymorphism. **N. Engl. J. Med.** **324**:872-876, 1991.

AUSUBEL, F. et al. *Current protocols in Molecular Biology*. 1 Ed. USA: John Wiley & Sons, Inc., 1994.

BHANDARI, P.; & GOWRISHANKAR, J. An Escherichia coli host strain useful efficient overproduction of cloned gene products with NaCl as the inducer. **J Bacteriol.** **179**:4403-6, 1997

BIRCHLER, J. A. et al. Dosage-dependent gene regulation in multicellular eukaryotes: implications for dosage compensation, aneuploid syndromes, and quantitative traits. **Dev. Biol** **234**: 275-288, 2001.

BIRD C. H. et al. Nucleocytoplasmic Distribution Of The Ovalbumin Serpin PI-9 Requires a Nonconventional Nuclear Import Pathway And Export Factor Crm 1. **Molecular and Cellular Biology**. **21**(6):5396-5407 2001.

BLOM, N.; GAMMELTOFT, S.; BRUNAK, S. Sequence and Structure Based Prediction Of Eukaryotic Protein Phosphorylation Sites. **J. Mol. Bio.** **294** (5):1351-1362, 1999.

BOULIKAS, T. Putative nuclear localization signals (NLS) in protein transcription factors. **J. Cell Biochem.** **55**(1):32-58, 1994.

BOVICELLI, L. et al. Reproduction in Down Syndrome. **Obstet Gynecol.** **59**(6 Suppl): 13S-7S, 1982.

BRADEN, C.; & TOOZE, J. *Introduction to protein structure. Determination of protein structures*. New York. Garland Publishing, Inc. 1991.

BRADFORD, M. M. A rapid and sensitive Method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. **Anal. Biochem.** **72**: 248-254, 1976.

BRAUN, P. et al. Proteome-scale purification of human proteins from bacteria. **PNAS** **99**: 2654-2659, 2002.

- BULLIDO R. et al. Several protein regions contribute to determine the nuclear and cytoplasmic localization of the influenza A virus nucleoprotein. *J of General Virology* **81**: 135-142, 2000.
- BURNS, G. W.; & BOTTINO, P. J. **Genética**. Rio de Janeiro. 6ª Ed. Guanabara Koogan, 1991. 381 p.
- CAPONE, G.T. Down Syndrome: Advances in Molecular Biology and the Neurosciences. **Developmental and Behavioral Pediatrics**. **22**(1):40-59, 2001.
- CEDANO, J. et al. Relation Between amino acid composition and cellular localization of proteins. **J. Mol. Biol.** **266**:594-600, 1997.
- CHA, H. J. et al. Observations of Green Fluorescent Protein as a Fusion Partner in Genetically Engineered *Escherichia coli*: Monitoring Protein Expression and Solubility. **Biotechnology and Bioengineering**. **67**: 565-574. 2000.
- CHARLOTTE, A. et al. Polymorphisms in Genes Involved in Folate Metabolism as Maternal Risk Factors for Down Syndrome. **AM. J. Hum. Genet.** **67**: 623-630, 2000.
- CHRAST, R. et al. The mouse brain transcriptome by SAGE: differences in gene expression between P30 brains of the partial trisomy 16 mouse model of Down syndrome (Ts65Dn) and normal. **Genome Research** **10**: 2006-2021, 2000.
- CORPET, F. Multiple sequence alignment with hierarchical clustering. **Nucl. Acids Res** **16**(22):10881-10890, 1988.
- DABROWSKI, S.; BRILLOWSKA, A.; KUR, J. Use of the Green Fluorescent Protein Variant (YFP) to Monitor MetArg Human Proinsulin Production in *Escherichia coli*. **Protein Expression Protein** **16**: 315, 323, 1999.
- DAVIS, G. D. et al. New fusion protein systems designed to give soluble expression in *Escherichia coli*. **Biotechnology and bioengineering** **65**(4): 382-388, 1999.
- DELABAR, J. M. et al. Molecular mapping of 24 features of Down's Syndrome on chromosome 21. **Eur. J. Hum. Genet** **1**: 114-24, 1993.
- De WIT, N. J., et al. Expression profiling of MMA-1a e splice variant MMA-1b: new cancer/testis antigens identified in human melanoma. **Int. J. Cancer** **98**: 547-553, 2002.
- DIECKMAN, L. et al. High throughput methods for gene cloning and expression. **Protein Expression and Purification** **25**: 1-7, 2002
- DOWN, J. L. H. Observation on an ethnic classification of idiots. **London Hosp. Clin. Lect. Rep** **3**: 259, 1886.
- EDGAR, A. J.; & POLAK, J. M. Human homologues of yeast vacuolar protein sorting 29 e 35. **Biochemical and Biophysical Research Communication** **277**: 622-630, 2000.
- EGEO, A. et al. Identification and characterization of a new human cDNA from chromosome 21q22.3 encoding a basic nuclear protein. **Hum. Genet.** **8**: 289-93, 1998.
- EKI, T. et al. Cloning and characterization of a novel gene, DCRR1, expressed from Down's syndrome critical region of human chromosome 21q22.2. **DNA Res** **7** (3-4): 153-64, 1977.

- ENGIDAWORK, E.; & LUBEC, G. Protein expression in Down Syndrome. **Amino Acids** **21**: 331-61, 2001.
- EPSTEIN, C. J. *The consequence of chromosomes imbalance: principles, mechanism and models*. Cambridge University Press, Cambridge. 1986.
- FEDOROV, A. N.; & BALDWIN T. O. Cotranslational protein folding. **J. Biol. Chem.** **272**: 32715-32718, 1997.
- FERRE-D'AMARÉ, A. R.; & BURLEY, S. K. Use of dynamic light scattering to assess crystallizability of macromolecules and macromolecular assemblies. **Structure** **2**: 357-359 1994
- FERRIERES, G. et al. Human cardiac troponin I: precise identification of antigenic epitopes and prediction of secondary structure. **Clinical Chemistry** **44**: 487-493, 1998.
- FISCH, H. et al. The influence of paternal age on Down Syndrome. **The J. Urol.** **169**: 2275-2278, 2003.
- FOX, J. D.; KAPUSTR. B.; WAUGH, D. S. Single amino acid substitutions on the surface of *Escherichia coli* maltose-binding protein can have a profound impact on the solubility of fusions proteins. **Protein Science** **10**: 622-630, 2001.
- FUENTES, J. J. et al. A new human gene from Down syndrome critical region encodes a proline-rich protein highly expressed in fetal brain and heart. **Hum. Molec. Genet** **4**: 1935-1944, 1995.
- FUENTES, J. J. et al. DSCR1, overexpressed in Down syndrome, is an inhibitor of calcineurin-mediated signaling pathways. **Hum. Molec. Genet.** **9**: 1681-1690, 2000.
- GARDINER, K.; & DAVISSON, M. The sequence of human chromosome 21 and implications for research into Down syndrome. **Genome Biology** **1**: 2.1-2.9, 2000.
- GARDINER K. et al. Annotation of Human Chromosome 21 for Relevance to Down Syndrome: Gene Structure and Expression Analysis. **Genomics.****79**: 833-843, 2002.
- GARDINER, et. al. Mouse of Down syndrome: how useful can they be? Comparison of the gene content of human chromosome 21 with orthologous mouse genomic regions. **Gene** **318**: 137-147, 2003.
- GOLSTEIN, H. Menarche, menstruation, sexual relations and contraception of adolescent females with Down syndrome. **Eur. J. Obstet. Gynecol. Reprod. Bio.** **27**: 343-9, 1998.
- GORLICH, D.; & MATTAJ, I. W. Nucleocytoplasmic transport. **Science.** **271**: 1513-1518, 1996.
- GREBER-PLATZER, S. et al. Evidence against the current hypothesis of “gene dosage effects” of trisomy 21: ets-2, encoded on chromosome 21” is not overexpressed in heart of patients with Down syndrome. **BBRC** **254**: 395-399, 1999.
- GUIMERA, J. et al. Cosmid contig and transcriptional map of three regions of human chromosome 21q22: Identification of 37 novel transcripts by direct selection. **Genomics** **45**: 59-67, 1997.
- GUIOCHON-MANTEL et al. Nucleocytoplasmic shuttling of the progesterone receptor. **EMBO J.** **10**:3851-3859, 1991.

- GURUPRASAD, K.; REDDY, B. V.; PANDIT, M. W. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. **Protein Eng.** **4**: 155-161, 1990.
- GUSMÃO, F. A. F.; TAVARES, E. J. M.; MOREIRA, L. M. A. Idade materna e síndrome de Down no nordeste do Brasil. **Cad. Saúde Pública, Rio de Janeiro.****19**: 973-978, 2003.
- GUZMAN, L. M, et al. Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. **J Bacteriol.** **177**:4121-30, 1995.
- HAFT, C. R. et al. Human orthologs of yeast vacuolar protein sorting proteins Vps26, 29, and 35: assembly into multimeric complexes. **Molecular Biology of the Cell.** **11**: 4105-4116, 2000.
- HAMMARBERG, B. et al. Dual affinity fusion approach and its use to express recombinant human insulin-like growth factor II. **Proc. Natl. Acad. Sci. USA** **86**: 4367-4371, 1989.
- HAMMARSTRÖM M, et al. Rapid screening for improved solubility of small human proteins produced as fusion proteins in *Escherichia coli*. **Protein Science** **11**: 313-321, 2002.
- HANNING, G.; & MAKRIDES, S. C. Strategies for optimizing heterologous protein expression in *Escherichia coli*. **Trends in Biotechnol.** **16**: 54-60, 1998.
- HATTORI, M., et al. The DNA sequence of human chromosome 21. **Nature** **405**: 311-318, 2000.
- HOJA M. R.; WAHLESTED, C.; HÖÖG, C. A visual intracellular classification strategy for uncharacterized human proteins. **Exp. Cell Research.** **259**: 239-246, 2000.
- HOOK, E. G. Epidemiology of Down syndrome. In: Pueschel, S. M.; RINDERS, J. E.: *Down Syndrome. Advances in Biomedicine and Behavioral Science*. Cambridge. Ware Press. pp 11. 1982.
- JONASSON, P. et al. Genetic design for facilitated production and recovery of recombinant proteins in *Escherichia coli*. **Biotechnol. Appl. Biochem.** **35**: 91-105, 2002.
- JANS D. A. et al. Nuclear Targeting Signal Recognition: A Key Control Point In Nuclear Transport? **BioEssays** **22**:532-544, 2000.
- JYOTHY, A. et al. Parental age and origin of extra chromosome 21 in Down Syndrome. **J. Hum. Genet.** **46(6)**: 347-50, 2001
- KATTAN. H. A. et al. A pilot study of the relationship between Dow's Syndrome and hearing loss. **Saudi. Med. J.** **21(10)**: 931-3, 2000.
- KIEFHABER, T. et al. Protein aggregation in vitro and in vivo: A quantitative model of the kinetic competition between folding and aggregation. **Biotechnology** **9**: 825-829. 1991.
- KIM B. Y. et al. Molecular characterization of mammalian homologues of class C Vps proteins that interact with syntaxin-7. **J. Biol. Chem.** **276**: 29393-29402, 2001.
- KLAPPER, M. H. The independent distribution of amino acid near neighbor pairs into polypeptides. **BBCA.** **78**: 1018-1024, 1977.
- KOLASKAR, A. S.; & TONGAONKAR, P. C. A semi-empirical method for prediction of antigenic determinants on protein antigens. **FEBS lett.** **276**: 172-174, 1990.

- KORENBERG, J. R. et al. Down syndrome phenotypes: The consequences of chromosomal imbalance. **Proc. Natl. Acad. Sci. USA.** **91**: 4997-5001, 1994.
- KORENBERG, J. R., et al. Molecular definition of a region of chromosome 21 that cause features of the Down Syndrome phenotype. **Am. J. Hum. Genet.** **47**: 236-246, 1990.
- KURTEN, R. C.; CADENA, D. L.; GILL, G. N. Enhanced degradation of EGF receptor by a sorting nexin, SNX1. **Science** **272**: 1008-1010, 1996.
- LAEMMLI, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **Nature** **227**: 680-685, 1970.
- LAVALLIE, E. R. et al. A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. **Biotechnology** **11**: 187-1893, 1993.
- LEE, J. J. et al. Identification and characterization of a novel, evolutionary conserved gene disrupted by the murine H beta 58 embryonic lethal transgene insertion. **Development** **115**(1): 277-88, 1992.
- LEJEUNE, J.; GAUTIER, M.; TURPIN, R. A study of somatic chromosomes in nine infants with mongolism. **CR Acad Sci (Paris)** **240**: 1026-1027, 1959.
- LESHIN L. Trissomy 21: The Story of Down Syndrome. 2000 In. www.ds-health.com/trisomy.htm. Acessado em 02/07/2003.
- LICENT, D. Localization of 102 exon to a 2,5 Mb region involved Down syndrome. **Hum Mol Genet.** **8**: 1305-11, 1995
- LINDWALL, G. et al. A sparse matrix approach to the solubilization of overexpressed proteins. **Protein Engineering** **13**:67-71, 2000.
- MAO, R. et al. Global up-regulation of chromosome 21 gene expression in the developing Down syndrome brain. **Genomics** **81**: 457-467, 2003.
- MARGOLIN, W. Green fluorescent protein as a reporter for macromolecular localization in bacterial cells. **Methods** **20**: 62-72, 2000.
- MAZZONI, D. S.; ACKKLEY, R. S.; NASDH, D. J. Abnormal pinna type and hearing loss correlations in Down's syndrome. **J. Intellect. Disabil. Res.** **38**: 549-560, 1994.
- McGUFFIN, L. J.; BRYSON, K.; JONES, D. T. The PSIPRED protein structure prediction server. **Bioinformatics.** **16**: 404-405, 2000.
- MIKKELSEN, M. Down's syndrome cytogenetics epidemiology. **Hereditas** **86**: 45-59, 1977.
- MISAWA, K et al. A method to identify cDNAs based on localization of green fluorescent protein fusion products. **PNAS** **28**:3062-3066, 2000.
- MORRIS, J.K.; MUTTON, D. E.; ALBERMAN, E. Revised estimatives of the maternal age specific live birth prevalence of Down's Syndrome. **Physiol. Res.** **49**: 227-231, 2000.
- MURBY, M.; UHLÉN, M.; STAHL, S. Upstream strategies to minimize proteolytic degradation upon recombinant production in *Escherichia coli*. **Protein Expression and Purification.** **7**: 129-136, 1996.

- NAKAI, K.; & HORTON, P. PSORT: a program for detecting the sorting signals of proteins and prediction their localization. **Trends Biochem. Sci.**, **24**(1):34-35, 1999.
- NAKAMURA, A., HATTORI, M., SAKAKI, Y. A novel gene isolated from human placenta located in Down syndrome critical region on chromossome 21. **DNA Res.** **4** (5): 321-4, 1997 a.
- NAKAMURA, A., HATTORI, M., SAKAKI, Y. Isolation of a novel human gene from Down syndrome critical region of chromosome 21q22.2. **J. Biochem.** (Tokyo). **122** (4): 872-7, 1997b.
- NETZER, W. J.; & HARTL, F. U. Recombinant of protein domains facilitated by co-translational folding in eukaryotes. **Nature** **388**: 343- 349, 1997.
- NILSSON, J. et al. Affinity fusion strategies for detection, purification, and immobilization of recombinant proteins. **Protein Expression and Purification** **11**: 1-16, 1997.
- NIZETIC, D. Functional genomics of the Down Syndrome. **CMJ**: **42**. 421-427, 2001.
- OHIRA, M. et al. A 1,6 Mb P1-basead physical map pf the Down Syndrome Region on Chomosome 21. **Genomics** **33**: 65-74, 1996.
- OLLIS, D.; & WHITE, S. *Protein crystallization. Methods in Enzymology*. Academic Press, Inc. V 182, 1990.
- OLOF, E. et al. prediction subcellular localization of proteins based on their N-terminal acid sequence. **J. Mol. Biol.** **3000**: 1005-1016, 2000.
- PENROSE, L. S. The relative effects of paternal and maternal age in mongolism. **J. Genet.** **27**: 219, 1933.
- PETERSEN, M. B.; & MIKKELSEN, M. Nondisjunction in trisomy 21: origin and mechanisms. **Cytogenet Cell Genet** **91**(1-4): 199-203, 2000.
- PFEFFER, S. R. Membrane transport: Tetramer to the rescue. **Current Biology.** **11**: R109-R111, 2001
- PFISTER, S. C. et al. Mutation analyses of the signal involved in the subcelular localization of DSCR1. **BMC Cell Biology** **3**:24, 2002.
- PINES, J. GFP in Mammalian Cells. **Technical Focus.** **11**(8):320-328, 1995.
- PRASHER, D. C. et al. Primary structure of the Aequorea victoria green-fluorescent protein. **Gene** **111**: 229-233, 1992.
- PRESCOTT, M. et al. The length of polypeptide linker affects the stability of green fluorescent protein fusion proteins. **Analytical Biochemistry**, **273**: 305-307, 1999.
- PRYOR, K, D.; & LEITING, B. High-Level Expression of Soluble Protein in *Escherichia coli* Using a His₆-Tag and Maltose-Binding-Protein Double-Affinity Fusion System. **Protein Expression and Purification** **10**: 309-319, 1997.
- RAHMANI, J. R. et al. Critical role of the D21S55 region on chromosome 21 in the pathogenesis of Down syndrome. **Proc. Natl. Acad. Sci. USA.** **86**: 5958-5962, 1989.
- RECHSTEINER, M.; & ROGERS, S. W. PEST sequences and regulation by proteolysis. **Trends Biochem Sci** **21**: 267-271, 1996.

- REDDY, J. V.; & SEAMAN M. N. J. Vps26p, a component of retromer, directs the interactions of Vps35p in endosome-to-Golgi Retrieval. **M. Biol. Cell.** **12**: 3242-3256, 2001.
- REYMOND, A. et al. From PREDs and Open Reading Frames to cDNA Isolation: Revesting the Humana Chromosome 21 Transcription Map. **Genomics.** **78**: 46-54, 2001.
- REYMOND, A. et al. Nineteen additional unpredicted transcripts from human chromosome 21. **Genomics** **79**: 824-832, 2002.
- RINAS, U.; & BAYNES, J. E. Overexpression of bacterial hemoglobin causes incorporation of pre- β -lactamase into cytoplasmic inclusion bodies. **Appl. Environ. Microbiol.** **59**: 561-566, 1993.
- ROIZEN, N. J.; & PATTERSON D. Down's syndrome. **The Lancet.** **361**: 12811289, 2003.
- ROST, B.; & SANDER C. Prediction protein structure at better than 70% accuracy. **J. Mol Biol.** **232**: 584-99, 1993.
- ROTHERMEL, B. et al. A protein encoded within the Down Syndrome Critical Region is enriched in striated muscles and inhibits calcineurin signaling. **JBC**: **275**: 9719-8725, 2000.
- RÜCKER, E. et al. Rapid evaluation and optimization of recombinant protein production using GFP tagging. **Protein Expression and Purification.** **21**: 220-223, 2001.
- SACHDEV, D.; & CHIRGWIN J. M. Order of fusions between bacterial and mammalian proteins can determine solubility in *Escherichia coli*. **Biochemical and Biophysical Research Communications** **244**: 933-937, 1998.
- SAMBROOK. J.; & RUSSEL, D. W. *Molecular Cloning: A Laboratory Manual.* 3rd Ed. Cold Spring Harbor Press, New York: Cold Spring Harbor Laboratory, 2001.
- SANGER, F.; NICLEN, S.; COULSON, A. R. DNA sequencing with chain terminating inhibitors. **Proc. Natl. Acad. Sci. USA.** **74**: 5463-5467, 1977.
- SATI, S. P. et al. Extra terminal residues have a profound effect on folding and solubility of a *Plasmodium falciparum* sexual stage-specific protein over-expressed in *Escherichia coli*. **Eur. J. Biochem** **269**: 5259-5263, 2002.
- SCHEIN, C. H.; MATHIEN, H. M.; NOTEBORN, H. M. Formation of soluble recombinant proteins in *E. coli* is favored by lower growth temperature. **Biotechnology** **6**: 291-294, 1988.
- SCHWALBACH, G. et al. Production of fluorescent single-chain antibody fragments in *Escherichia coli*. **Protein Expression and Purification** **18**: 121-132, 2000.
- SEPULVEDA, D. et al. Low spinal and pelvic bone mineral density among individuals with Down Syndrome. **Am. J. Ment. Retar.** **100**: 109-14, 1995.
- SHANG, LI, et al. Identification of a novel cytoplasmic protein that specifically binds to nuclear localization signal motifs. **J. Biol. Chem.** **273**(11): 6183-6189, 1998.
- SHAPIRO, B. L. The Down Syndrome critical region. **J. Neural. Trans (Suppl)** **57**: 41-60, 1999
- SHAPIRO, B. L. Whither Down syndrome critical regions? **Hum. Genet** **99**: 442-423, 1997

- SHIBUYA K. et al. Isolation of two novel genes, DSCR5 e DSCR6, from Down Syndrome Critical Region on human chromosome 21q22.2. **Biochem. Biophys. Res. Commum** **271**: 693-698, 2000.
- SHOTT, S. R. et al. Hearing loss in children with Down Syndrome. **Int. J. Pediatr. Otorhinolaryngol.** **61**(3):199-205, 2001.
- SHIH, Y, P. et al. High-throughput screening of soluble recombinant proteins. **Protein Science** **11**: 1714-1719, 2002.
- SILBER, K, R. ; KEILER, K. C.; SAUER, R. T. Tsp: a tail-specific protease that selectively degrades proteins with nonpolar C termini. **Proc. Nat. Acad. Sci USA** **89**: 295-299, 1992.
- SINET, P. M. et al. Mapping of the Down Syndrome phenotype on chromosome 21 at the molecular level. **Biomed. & Pharmacother.** **48**: 247-252, 1994.
- SKOSYREV, V. S. et al. EGFP as a fusion partner for the expression and organic extraction of small polypeptides. **Protein Expression & Purification** **27**: 55-62, 2003.
- STEVENS R. C. Design of high-throughput methods of protein production for structural biology. **Structure Fold Des.** **8** (9): R177-85, 2000.
- STOLL, V. S. et al. Thioredoxin fusion protein of VanH, a D-lactate dehydrogenase from *Enterococcus faecium*: cloning, expression, purification kinetic analysis and crystallization. **Protein Science** **7**: 1147-1155, 1998.
- TAKAMATSU, K. et al. Identification of two novel primate-specific genes in DSCR. **DNA Res.** **9**: 89-97, 2002.
- THULINE, H. C.; & PUESCHEL, S. M. Cytogenetics in Down syndrome. *In*: PUESCHELL, S. M.; RYNDERS, J. E. *Down Syndrome Advances in Biomedicine and the Behavioral Science*. Cambridge: Ware Press (pub).1982. p 133.
- TOYODA, A. et al. Comparative genomic sequence analysis of the human chromosome 21 Down syndrome critical region. **Genome Res.** **12**: 1323-1332, 2002.
- TSUGAWA, A. et al. *E. coli* NusA protein binds in vitro to an RNA sequence immediately upstream of the boxA signal of bacteriophage lambda. **EMBOJ.** **4**:2337-2342, 1985.
- TSUKAHARA, F. et al. Molecular characterization of the mouse mtprd gene, a homologue of Human TPRD: Unique gene expression suggesting its critical role in the pathophysiology of Down Syndrome. **J. Biochem.** **123**: 1055-1063, 1998.
- VAN BEIJNUM, J. R. et al. Target validation for genomics using peptide-specific phage antibodies: a study of five gene products overexpressed in colorectal cancer. **Int. J Cancer** **101**(2): 118-27, 2002.
- VARSHAVSHY, A. The N-end rule: Functions, mysteries, uses. **Proc. Natl. Acad. Sci USA** **93**: 12142-12149, 1996.
- VIDAL-TABOADA, J. M. et al. Down Syndrome critical region gene 2: expression during mouse development and in human cell lines indicates a function related to cell proliferation. **Biochem Biophys Res Commum** **272**(1): 156-63, 2000.
- VEGA, R. B. et al. Multiple domains of MCIP1 contribute to inhibition of calcineurin activity. **J. Biol. Chem.** **277**: 30401-30407, 2002.

WALDO, G. S. et al. Rapid protein-folding assay using green fluorescent protein. **Nature Biotechnol** 17: 691-695, 1999.

WALL, J. G.; & PLUCKTHUN, A. Effects of overexpressing folding modulations on the in vivo folding of heterologous proteins in *Escherichia coli*. **Curr Opin Biotechnol** 6: 507-516, 1995.

WANG, A. et al. A novel method for increasing the expression level of recombinant proteins. **Protein Expression and Purification** 30: 124,133, 2003.

WU, C. F. et al. GFP-visualized immobilized enzymes: Degradation of paraoxon via organophosphorus hydrolase in a packed column. **Biotechnology and Bioengineering** 77: 212-217, 2002.

YASUWAVA, T. et al. Increase of solubility of foreign proteins in *Escherichia coli* by coproduction of the bacteria thioredoxin. **J. Biol. Chem.** 270: 25328-25331, 1995.

YOKOYAMA, S. Protein expression systems for structural genomics and proteomics. **Current Opinion in Chemical Biology** 7: 39-43, 2003.