

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

**"EXPRESSÃO HETERÓLOGA E LOCALIZAÇÃO SUBCELULAR DE SEIS  
PROTEÍNAS POSSIVELMENTE ENVOLVIDAS COM A SÍNDROME DE DOWN"**

**Daniela Morilha Néo-Justino**

**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 E LOCALIZAÇÃO SUBCELULAR DE SEIS  
PROTEÍNAS POSSIVELMENTE ENVOLVIDAS COM A SÍNDROME DE DOWN"**

**Daniela Morilha Néo-Justino**

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**

N438eh

Néo-Justino, Daniela Morilha.

Expressão heteróloga e localização subcelular de seis proteínas possivelmente envolvidas com a Síndrome de Down / Daniela Morilha Néo-Justino. -- São Carlos : UFSCar, 2005.

106 p.

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

1. Genética molecular. 2. Down, Síndrome de. 3. Trissomia do 21. 4. Expressão heteróloga. 5. Localização subcelular. I. Título.

CDD: 574.873228 (20<sup>a</sup>)

Orientador

---

Prof. Dr. Flávio Henrique da Silva

*“A paciência, em verdade, é a força que nos assegura a calma e o conhecimento nas horas amargas.”*

## AGRADECIMENTOS

Ao Prof. Dr. Flávio Henrique da Silva, que me deu a oportunidade de realizar este trabalho e que me proporcionou um grande aprendizado.

À Universidade Federal de São Carlos, Departamento de Genética e Evolução e Programa de Pós-Graduação em Genética e Evolução.

À Fundação de Amparo à Pesquisa do Estado de São Paulo pelo auxílio concedido.

À Banca Examinadora pela análise deste trabalho.

À Dra Ana Paula Ulian de Araújo (IFSC-USP-São Carlos) e Dra Heloísa Selistre de Araújo (UFSCar) pela atenção e uso do laboratório.

À Dra Gláucia Maria Machado Santelli (ICB-USP-São Paulo) pela colaboração e uso do laboratório.

Ao Dr Otávio Henrique Thiemman pela gentileza e uso do laboratório.

À Dra Leila Maria Beltramini (IFSC-USP-São Carlos) pelos ensaios de Dicroísmo Circular.

Às secretárias Leila, Rose, Regiane e Tatiane pela paciência e toda a ajuda prestada.

Aos colegas Raphael Bessa e Patrícia Campana pelo auxílio prestado.

À Bel, Andressa e Suzana que sempre me recebem bem e me ajudaram em tudo.

Aos colegas do laboratório de ontem e de hoje Sandra, Camillo, Patrícia, Wilson, Henrique, Alexandre, Rodrigo, Márcia, Antônia, Andrea, Elisete, César, Reis, Wendy, Andréia, Viviane e Juliana pela convivência agradável e principalmente pela colaboração e solidariedade em todos os momentos.

Às amigas Andrea, Elisete, Flávia, Sandra e Elisandra por toda a ajuda, estímulo, amizade e apoio em todas as etapas deste trabalho.

Ao Santiago pelo apoio, paciência, colaboração, companheirismo e amor em todos os momentos.

À minha família, especialmente aos meus pais que sempre me deram muito amor e todo o apoio e incentivo. Obrigado por tudo!!!!

## LISTA DE ILUSTRAÇÕES

<b>Figura 1:</b> Características apresentadas por portadores da Síndrome de Down.....	4
<b>Figura 2:</b> Mapa esquemático do cromossomo 21 mostrando a localização da DSCR e a DCR2.....	8
<b>Figura 3:</b> Esquema representativo da estratégia utilizada para isolar o gene DCRB.....	16
<b>Figura 4:</b> Sequência parcial de nucleotídeos do RNA mensageiro do gene que codifica a DCRB.....	46
<b>Figura 5:</b> Construção do gene DCRB.....	47
<b>Figura 6:</b> Análise da sequência do DCRB humano.....	51
<b>Figura 7:</b> Análise da expressão do DCRB em linhagens de células tumorais.....	52
<b>Figura 8:</b> Análise do alinhamento da sequência do DCRB humano contra a sequência genômica do chimpanzé ( <i>Pan troglodytes</i> ).....	53
<b>Figura 9:</b> Análise da expressão e solubilidade da proteína DCRB em <i>E.coli</i> (BL21-DE3) utilizando o vetor pET28a.....	56
<b>Figura 10:</b> Análise da co-expressão e solubilidade da proteína DCRB com a chaperonina GroESL em <i>E.coli</i> (BL21-DE3).....	57
<b>Figura 11:</b> Análise da expressão da proteína DCRB em <i>E.coli</i> (BL21-DE3) utilizando o vetor pET37b.....	58
<b>Figura 12:</b> Análise do teste de solubilidade da proteína pET37bDCRB expressa em <i>E. coli</i> . BL21 (DE3).....	59
<b>Figura 13:</b> Análise dos testes de indução da expressão da proteína DCRB clonada no vetor pBad.....	60
<b>Figura 14:</b> Análise do teste de expressão da proteína DCRB no plasmídeo pTYB2.....	62
<b>Figura 15:</b> Análise do teste de solubilidade da proteína DCRB em fusão com a inteína.....	62
<b>Figura 16:</b> Caracterização dos corpos de inclusão produzidos através da expressão da DCRB no vetor pET28a.....	64
<b>Figura 17:</b> Sequência de aminoácidos codificados pela DCRB.....	65
<b>Figura 18:</b> Análise da expressão e solubilidade da proteína DCRB mutada em <i>E.coli</i> (BL21-DE3) em SDS-Page 15%.....	66
<b>Figura 19:</b> Análise do sequenciamento de um clone mutante do DCRB.....	67
<b>Figura 20:</b> Análise da expressão da proteína DCRB 2/3 em <i>E.coli</i> (BL21-DE3). SDS-Page 15% mostrando teste de expressão e purificação da proteína.....	69

<b>Figura 21:</b> Análise do teste de expressão e solubilidade da proteína DCRB na bactéria Rosetta (DE3).....	70
<b>Figura 22:</b> Perfil cromatográfico da eluição da DCRB na coluna Superose12.....	71
<b>Figura 23:</b> Análise da purificação da proteína DCRB.....	72
<b>Figura 24:</b> Espectros de CD.....	74
<b>Figura 25:</b> Análise do reconhecimento da DCRB pelo antissoro contra esta proteína.....	75
<b>Figura 26:</b> Visualização da coloração observada no teste de titulação do anticorpo contra a DCRB.....	76
<b>Figura 27:</b> Análise por “western blot” da expressão e purificação da DCRB em células de mamíferos COS-7.....	78
<b>Figura 28:</b> Análise da expressão da ORF45 em <i>E.coli</i> (B121-DE3). SDS-Page 15% mostrando teste de expressão e solubilidade da proteína ORF 45.....	80
<b>Figura 29:</b> Análise da expressão da ORF 83-2 em <i>E.coli</i> (B121-DE3). SDS-Page 15% mostrando teste de expressão e solubilidade da proteína ORF 83-2.....	81
<b>Figura 30:</b> Análise da localização intracelular da proteína DCRB.....	92
<b>Figura 31:</b> Análise da localização intracelular da proteína DCRB e das suas formas truncadas.....	94
<b>Figura 32:</b> Análise da localização intracelular das proteínas estudadas.....	95
<b>Tabela 1:</b> Diluições realizadas para determinação da titulação do anticorpo.....	37
<b>Tabela 2:</b> Características das novas ORFs do cromossomo 21 selecionadas para estudos.....	38
<b>Tabela 3:</b> Alinhamento contra ESTs humanas.....	51
<b>Tabela 4:</b> Valores correspondentes à leitura da placa de titulação.....	76
<b>Tabela 5:</b> Análise <i>in silico</i> e <i>in vivo</i> das proteínas estudadas.....	91

## SUMÁRIO

<b>1.INTRODUÇÃO</b>	<b>1</b>
1.1 Considerações Gerais.....	1
1.2 O cromossomo 21.....	5
1.3 Região Crítica da Síndrome de Down.....	7
<b>2. OBJETIVOS</b>	<b>15</b>
<b>3. MATERIAL E MÉTODOS</b>	<b>16</b>
3.1 Obtenção do gene DCRB.....	16
3.2 Avaliação da expressão da DCRB.....	16
3.2.1 Extração de RNA de linhagens celulares.....	17
3.2.2 Avaliação da qualidade dos RNAs obtidos.....	18
3.2.3 RT-PCR.....	19
3.3 Descrição dos vetores utilizados.....	20
3.4 Construção dos vetores de expressão.....	22
3.4.1 Construção do vetor pET28aDCRB.....	22
3.4.2 Construção do vetor pET37bDCRB.....	23
3.4.3 Construção do vetor pBADTOPODCRB.....	24
3.4.4 Construção do vetor pCX-TOPODCRB.....	25
3.4.5 Construção do vetor pTYB2DCRB.....	26
3.4.6 Construção do vetor pcDNA3DCRB.....	26
3.5 Mutação sítio dirigida do gene DCRB.....	27
3.6 Deleção de parte da proteína DCRB.....	29
3.7 Teste de solubilidade.....	30
3.8 Purificação de proteínas.....	32
3.8.1 Cromatografia de afinidade.....	32
3.8.2 Cromatografia de exclusão molecular.....	33
3.8.3 Obtenção da proteína DCRB em condições desnaturantes.....	33
3.9 Medidas de CD (Circular Dichroism –CD).....	34
3.10 Produção de anticorpos policlonais.....	34
3.10.1 Detecção de anticorpos.....	35
3.10.2 Titulação de anticorpos.....	36

3.11 Amplificação, clonagem e expressão de oito ORFs do cromossomo 21.....	37
3.12 Construção dos plasmídeos pEGFP-C1 com os genes de interesse.....	40
3.13 Construção das formas truncadas da DCRB no plasmídeo pEGFP-C1.....	41
3.14 Cultura e Transfecção celular.....	42
3.15 Coloração com DAPI.....	44
3.16 Análises <i>in silico</i> .....	44
<b>4. RESULTADOS E DISCUSSÃO</b> .....	<b>46</b>
4.1 O gene DCRB.....	46
4.2 Obtenção do gene DCRB e subclonagem no vetor de expressão pET28a.....	46
4.3 Análise da sequência do gene DCRB humano.....	49
4.4 Análise da expressão da DCRB humana em diferentes tipos celulares.....	51
4.5 Alinhamento do DCRB humano contra a sequência genômica de <i>Pan troglodytes</i> ...	53
4.6 Expressão da proteína DCRB em <i>E. coli</i> .....	55
4.6.1 Expressão da DCRB no vetor pET28a.....	55
4.6.2 Expressão da DCRB no vetor pET37b.....	57
4.6.3 Expressão da DCRB no vetor pBADTOPO.....	59
4.6.4 Expressão da DCRB no vetor pCX-TOPO.....	61
4.6.5 Expressão da DCRB no vetor pTYB2.....	61
4.7 Purificação da proteína DCRB em condições desnaturantes.....	63
4.8 Mutação sítio dirigida do gene DCRB.....	65
4.9 Deleção de parte da proteína DCRB.....	67
4.10 Expressão da DCRB utilizando-se a cepa Rosetta (DE3).....	68
4.11 Análise comparativa dos espectros de CD da proteína DCRB.....	73
4.12 Produção e titulação de anticorpos policlonais.....	75
4.13 Expressão da DCRB em células de mamíferos.....	77
4.14 Amplificação, clonagem e expressão de cinco ORFs do cromossomo21.....	79
4.15 Expressão e solubilidade das proteínas heterólogas em <i>E. coli</i> .....	81
4.16 Localização subcelular das proteínas estudadas neste trabalho.....	88
<b>5. CONCLUSÕES</b> .....	<b>98</b>
<b>6. REFERÊNCIAS BIBLIOGRÁFICAS</b> .....	<b>99</b>

## RESUMO

A Síndrome de Down (SD) é o distúrbio genético mais freqüente em humanos, sendo detectado em média um caso a cada 700 nascimentos. Esta anomalia é decorrente principalmente da trissomia total do cromossomo 21 porém, em alguns casos, ocorre a trissomia parcial deste cromossomo o que possibilitou a definição de uma região comum a todos os portadores denominada Região Crítica da Síndrome de Down (DSCR). Nesta região encontram-se vários genes que possivelmente estão relacionados às características fenotípicas apresentadas pelos indivíduos portadores. Entre eles estão os genes estudados neste trabalho: DCRB ou DSCR4 e as ORFS c21orf45, c21orf59, c21orf83-2, c21orf95 e c21orf101. Com o objetivo de contribuir para os estudos realizados com genes localizados na DSCR foram realizadas a expressão heteróloga em *E. coli* e a localização subcelular das proteínas codificadas pelos genes descritos acima. Assim, os experimentos de expressão heteróloga em *E. coli* demonstraram que as proteínas expressas encontram-se na forma insolúvel para a maioria das condições testadas, entretanto foi possível a expressão de uma pequena fração solúvel da c21orf45, c21orf59 e da DCRB quando utilizou-se células da linhagem Rosetta (DE3) o que permitiu que a última proteína fosse purificada em condições nativas e que ensaios iniciais de estrutura secundária fossem realizados. Experimentos de deleção e mutação sítio dirigida também foram realizados com esta proteína assim como a produção de anticorpos policlonais. Estudos da localização subcelular revelaram que quatro das proteínas analisadas (DCRB, c21orf45, c21orf59 e c21orf83-2) são citoplasmáticas enquanto que duas (c21orf95 e c21orf101) são nucleares. Considerando a escassez de informações a respeito dos genes localizados na DSCR, esses resultados vem contribuir para o conhecimento das proteínas possivelmente envolvidas com a Síndrome de Down.

## ABSTRACT

Down syndrome (DS) is the most common congenital disease occurring in approximately 1 out of 700 live births. In addition to the full HC21 trisomy, rare individuals with clinically recognized DS have only a partial trisomy, carrying a region common to all patients dubbed as Down Syndrome Critical Region (DSCR). Several genes contained in this portion of the HC21 are probably related to the DS phenotypes. In an attempt to characterize some of the DSCR genes, heterologous expression and subcellular localization analyses were performed for the genes DCRB (DSCR4), c21orf45, c21orf59, c21orf83-2, c21orf 95 and c21orf101. Analyses of the recombinant proteins produced in *Escherichia coli* showed that most of them remained in the insoluble fraction. DCRB was expressed as a soluble protein in *E. coli* Rosetta (DE3), purified and subjected to assays for secondary structure analyses. Also, deletion and site directed mutagenesis as well as the production of polyclonal antisera against DCRB were performed. Subcellular localization analyses revealed that the proteins DCRB, c21orf45, c21orf59 and c21orf83-2 were distributed in the cytoplasm of mammalian cells and that c21orf95 and c21orf101 resided in the nucleus. Considering the lack of information concerning these proteins encoded in the DSCR our results allowed a insight into some of the proteins provably involved in Down syndrome.

## 1. INTRODUÇÃO

### 1.1 Considerações Gerais

Entre a população humana tem sido descritas várias alterações cromossômicas como as síndromes de Turner, de Patau e de Down, sendo que a última é a mais freqüente e com conseqüências menos drásticas para os portadores.

Há 137 anos, John Langdon Down publicou a primeira descrição clínica da condição denominada atualmente de Síndrome de Down (SD) e atribuía esta anormalidade à tuberculose materna. Em 1932, Waardenburg sugere que a síndrome era conseqüência de anormalidade cromossômica e, finalmente em 1959 Lejeune et al. descobriram que a trissomia do cromossomo 21 era a responsável por estas anormalidades (Capone, 2001).

A Síndrome de Down ocorre em todas as raças e em ambos os sexos e é conseqüência de três eventos principais: trissomia, translocação e mosaicismo. Dentre estas, a mais comum é a trissomia completa do cromossomo 21 (Figura 1B) devido a não disjunção durante a formação dos gametas. Raramente a não disjunção ocorre após a fertilização, resultando em duas “linhagens” celulares, uma com o número normal de cromossomos e a outra com um cromossomo 21 extra. Este evento é denominado mosaicismo e é responsável por de 2 a 4% dos casos. Outro evento é a translocação que normalmente ocorre entre os cromossomos 14 e 21 com uma frequência de 3 a 5%. Embora esse tipo de translocação seja o mais freqüente, casos envolvendo outros cromossomos do complemento cariotípico já foram detectados (Capone, 2001).

A origem do cromossomo extra presente nos portadores da SD é devido a não disjunção na meiose em 95% dos casos, sendo que a maioria deles (88%) ocorre na meiose materna (68% na meiose I e 20% na meiose II) enquanto que 7% é de origem paterna e apenas 5% dos pacientes possuem uma trissomia de origem mitótica (Antonarakis, 1998).

Embora a causa da SD seja conhecida desde 1959, o mecanismo pelo qual três cópias do cromossomo 21 resulta na interrupção do desenvolvimento normal não é bem conhecido. Existem duas hipóteses mais conhecidas denominadas efeito da dosagem gênica e instabilidade no desenvolvimento. De acordo com a primeira hipótese ocorre um desbalanço de dose global ou em um pequeno número de genes do cromossomo 21 localizado na região crítica causando os fenótipos da Síndrome de Down. A segunda hipótese considera que a maioria das manifestações observadas em portadores da SD é causada por um distúrbio não especificamente relacionado à expressão e regulação gênica. (Mao et al., 2003 e Amano et al., 2004).

Para testar estas hipóteses vários estudos tem sido realizados especialmente aqueles relacionados ao efeito da dosagem gênica. Cheon et al. (2003a, d) analisaram a expressão de várias proteínas codificadas no cromossomo 21 usando a técnica de “western blot” e verificaram que nenhuma delas apresentou superexpressão em córtex cerebral de feto de portadores da SD quando comparadas ao controle. Em alguns casos, como para HACS1 a expressão foi significativamente diminuída (Cheon et al., 2003b) em outros, como para a sinaptojanina1 e c21orf2, foi aumentada (Cheon et al., 2003c) e superexpressa como a DSCR5 (Ferrando-Miguel et al., 2004). Assim, esses autores consideram que a hipótese de dosagem gênica não é suficiente para explicar totalmente os fenótipos observados em portadores da síndrome de Down.

Em análises realizadas com modelos de ratos para a síndrome de Down (Ts1Cje) foi observado que os níveis de expressão dos genes localizados na região trissômica aumentou 1,5 vezes, o que não ocorreu para a região euplóide ou para os outros cromossomos (Sago, et al., 1998). Esses resultados indicam que os genes da região trissômica de Ts1Cje são superexpressos de modo dose-dependente e estão relacionados à patologia da SD. Estudos como estes deverão contribuir para um melhor entendimento da

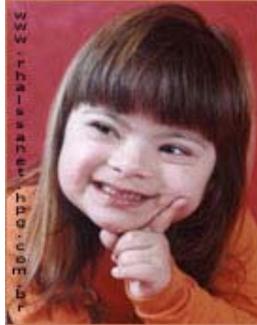
patobioquímica da síndrome de Down e revelar moléculas alvo para intervenções terapêuticas (Amano et al., 2004).

As características clínicas dos portadores da síndrome são facilmente detectadas e incluem principalmente hipotonia neonatal, face achatada, fendas palpebrais oblíquas para cima, orelhas pequenas com implantação baixa, prega palmar horizontal única (Figura 1A). Apresentam também anormalidades cardíacas, gastrointestinais, alterações no sistema endócrino e imunológico e doença de Alzheimer em idade prematura (Epstein, 1986; Hill et al., 2003). Crianças portadoras de até quatro anos de idade apresentam risco aumentado de desenvolver leucemia aguda (Fong e Brodeur, 1987). Por outro lado, a trissomia do 21 parece conferir proteção contra o desenvolvimento de tumores sólidos, sugerindo a existência de genes supressores de tumores nesse cromossomo (Hattori et al., 2000).

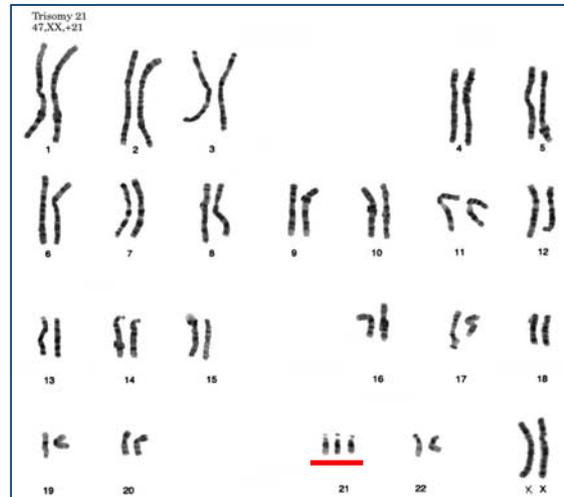
Algumas características são bastante freqüentes como as cardiopatias que ocorrem em 50% dos portadores da síndrome (Korenberg et al., 1992). O retardamento mental e a hipotonia neonatal são as únicas características comuns a 100% dos casos relatados (Korenberg et al., 1994). Indivíduos com SD têm mais problemas comportamentais e psiquiátricos que outras crianças, mas menores que outros indivíduos com retardamento mental (Roizen e Patterson, 2003).

Apesar de ser uma síndrome bastante freqüente e de seus aspectos clínicos serem bastante conhecidos, ainda não se conhece com exatidão os mecanismos moleculares pelos quais a trissomia ocorre, porém sabe-se que tem a sua incidência aumentada com o aumento da idade materna. Em análises realizadas por Hassold e Jacobs (1984) verificou-se uma incidência de 1 caso a cada 700 nascimentos em mulheres com idade materna em torno de 30 anos, o que foi significativamente aumentada (6:700 nascimentos) para mulheres com idade ao redor de 40 anos (Hook et al., 1983 e Vogel et al., 1997).

A)



B)



**Figura 1: Características apresentadas por portadores da Síndrome de Down.**

A) Características fenotípicas e B) Cariótipo de uma portadora da Síndrome de Down. Em destaque a trissomia total do cromossomo 21.

Recentemente, tem sido descrito um segundo possível fator de risco genético materno para a SD. Este fator está relacionado ao polimorfismo de genes envolvidos no metabolismo do ácido fólico, que é essencial no processo de síntese *de novo* dos nucleotídeos precursores da síntese de DNA e reações de metilação do DNA. Assim, a associação entre a deficiência de folato e a hipometilação do DNA sugere que deficiências genéticas e/ou nutricionais que afetam o metabolismo do folato podem estar relacionadas ao aumento de risco da não disjunção e consequentemente da Síndrome de Down (Hobbs et al., 2000).

Devido à importância desta síndrome, visto que ocorre uma alta incidência na população humana mundial em todas as raças e em ambos os sexos, os estudos realizados a respeito dos produtos gênicos do cromossomo 21 são de grande importância.

### **1.2 O cromossomo 21**

O cromossomo 21 está relacionado a algumas doenças genéticas como Alzheimer, desordens autoimunes, epilepsia e principalmente a Síndrome de Down (Guimera et al., 1997) o que o torna alvo de intensas pesquisas científicas que culminaram na publicação da sequência genômica completa deste cromossomo realizada por Hattori et al. (2000).

O cromossomo 21 representa cerca de 1 a 1,5% do genoma humano, sendo o menor cromossomo do complemento autossômico. O braço longo (21q) possui aproximadamente 37 Mb e constitui cerca de 1% do genoma humano e o braço curto (21p) é similar ao braço curto dos outros cromossomos acrocêntricos do complemento cariotípico humano (Antonarakis, 1998) e consiste basicamente de uma região organizadora de nucléolo (RON) que contém múltiplas cópias de genes que codificam RNA ribossômico e uma região mais proximal composta de sequências de DNA altamente repetitivas. Os genes localizados no braço curto parecem não serem essenciais para o desenvolvimento normal, já que a sua duplicação ou deleção usualmente leva a poucas manifestações fenotípicas observáveis (Capone, 2001).

Além de possuir uma baixa densidade gênica quando comparado a outros cromossomos autossômicos (Hattori et al., 2000), o 21 também apresenta uma variação dessa densidade ao longo do braço maior (21q). Na sua parte distal foram identificados 167 genes enquanto na proximal 58 genes, sendo notável uma região de 7 Mb contendo apenas 7 genes (Antonarakis et al., 2001).

Segundo Hattori et al. (2000), a baixa densidade gênica do cromossomo 21 é consistente com a observação de que a trissomia do 21 é uma das únicas trissomias autossômicas viáveis em humano.

Após o sequenciamento total deste cromossomo, verificou-se a existência de 225 genes e 59 pseudogenes. Desses, 127 correspondem a genes conhecidos, 98 representam putativos novos genes preditos *in silico*. Dentre os novos genes, 13 são similares a proteínas conhecidas, 17 ORFs anônimas com características de domínios modulares e os outros 68 genes são unidades de transcrição desconhecidas, sem similaridade com proteínas já descritas (Hattori et al., 2000).

Em trabalhos mais recentes, verificou-se que muitos modelos de genes foram incompletos, alguns estavam incorretos e, para poucos foram consideradas as expressões gênicas (Gardiner et al., 2002). Segundo Roizen e Patterson (2003), o conteúdo gênico do cromossomo 21 é agora estimado em 329 genes, incluindo 165 genes confirmados experimentalmente, 150 modelos de genes baseados em bancos de EST (“Expressed Sequence Tag”) e 14 de predições computacionais.

Fazendo uma análise geral dos genes catalogados no cromossomo 21, observam-se no mínimo dez quinases, cinco genes envolvidos nas vias de ubiquitinação, cinco moléculas de adesão celular, um número de fatores de transcrição e sete canais de íons, entre outros (Hattori et al., 2000). Aproximadamente 35% dos genes do cromossomo 21 são homólogos aos genes de *Drosophila*, 35% a *Caenorhabditis elegans* e 18% a *Saccharomyces cerevisiae* (Antonarakis et al., 2001).

Segundo Roizen e Patterson (2003) alguns genes do cromossomo 21 formam grupos que estão envolvidos no mesmo caminho metabólico ou sistema biológico. A geração de energia nas mitocôndrias e o metabolismo de espécies relativas ao oxigênio estão relacionadas a 16 genes presentes neste cromossomo. Capone (2001) relaciona um

grupo de 10 genes que devem exercer influência na estrutura e função do sistema nervoso central e pode ter um papel na neuropatogênese da Síndrome de Down.

Assim, a completa análise do cromossomo 21 humano terá profundas implicações no entendimento da patogênese e principalmente no desenvolvimento de novas terapias.

### **1.3 Região Crítica da Síndrome de Down**

Os avanços ocorridos no desenvolvimento de técnicas moleculares para análises citogenéticas foram de grande importância, visto que a partir daí foi possível detectar pequenas regiões do cromossomo 21 em portadores de várias características da Síndrome de Down que não apresentavam nenhuma anormalidade cromossômica.

Embora a trissomia total seja a mais freqüente, em alguns casos a trissomia é parcial, o que permitiu a definição de uma região do cromossomo 21 comum a todos os portadores da SD (Korenberg et al., 1988; Epstein et al., 1991 e Fuentes et al., 1995 e 1997).

Vários estudos têm sido realizados com objetivo de delimitar a região mínima responsável pela maioria das características da síndrome, entretanto ainda não há um consenso entre os pesquisadores.

A princípio toda a região 21q22 era considerada região crítica da SD (Epstein, 1986). Posteriormente Korenberg et al. (1988) consideram a porção cromossômica localizada abaixo do locus D21S55 como a região crítica. Entretanto, outros pesquisadores sugerem que a região localizada cerca de 3Mb ao redor da D21S55 seja a região mínima responsável pelas características da Síndrome de Down (Fuentes et al., 1995). Antonarakis (1998) definiu uma região de aproximadamente 4 Mb entre os marcadores D21S17 e ETS2 como a Região Crítica da Síndrome de Down (DSCR) e os genes localizados nessa região

estariam relacionados com o retardo mental, hipotonia muscular, crescimento retardado e hiperlaxia de junções, além de oito características dermatológicas e faciais.

Analisando pacientes com trissomia parcial foi possível mapear uma segunda região que se sobrepõe a primeira e estaria localizada entre os marcadores D21S55 e MX1 e inclui genes responsáveis por seis características dermatológicas e faciais e, em alguns casos, demonstra estar associada à cardiopatia e retardamento mental (Korenberg et al., 1992) e estenose duodenal (Vidal-Taboada et al., 1998) (Figura 2).



**Figura 2: Mapa esquemático do cromossomo 21 mostrando a localização da DSCR e a DCR2.**  
(Fonte: Antonarakis, 1998)

Assim, com o reconhecimento da ligação entre o locus D21S55 e a SD, vários estudos tem sido realizados com o objetivo de se detectar genes localizados nessa e/ou ao redor desta região e que possivelmente estejam relacionados com a síndrome o que deve

ser bastante facilitado com a publicação do mapeamento do cromossomo 21 por Hattori et al. (2000).

Epstein (1986) considerou que o gene SOD-1 estaria na região crítica, já que os portadores da síndrome possuíam um metabolismo oxidativo anormal decorrente da superexpressão deste gene. Entretanto, Mattei et al. (1981) e Yamamoto et al. (1986) analisaram casos em que alguns portadores da síndrome não apresentaram o gene da SOD-1 em triplicata e, portanto com níveis normais de expressão desta enzima o que levou os autores a propor que o gene SOD-1 não era o principal responsável pelas alterações fenotípicas encontradas nos portadores (Rahmani et al., 1989; Epstein et al., 1991 e Fuentes et al., 1995).

Fuentes et al. (1995) identificaram um gene denominado DSCR1 (Down Syndrome Critical Region 1), que está situado a 2Mb acima do locus D21S55 e possui 45 kb com 7 éxons e 6 íntrons. Análises de diferentes cDNAs possibilitaram verificar que os 3 primeiros éxons (1, 3 e 4) são alternativos e, portanto, podem originar proteínas alternativas e os 3 últimos (5, 6 e 7) são comuns a todas as proteínas. Este gene é altamente expresso em cérebro e coração e é um candidato no envolvimento com a patogênese da SD, principalmente em relação ao retardamento mental e a defeitos cardíacos. Recentemente, o gene DSCR1 teve sua localização subcelular analisada por Pfister et al. (2002) que verificaram que esta proteína é de localização nuclear independentemente da isoforma destacando a importância dos 33 aminoácidos da região c-terminal nesta localização. Além disso, estes autores identificaram alguns resíduos de serina e treonina que quando mutados alteram a localização da proteína sugerindo que sua localização subcelular e atividade possam ser reguladas por quinases e/ou fosfatases.

Outros estudos revelaram que a DSCR1 interage com a calcineurina, uma serina treonina-fosfatase que está envolvida em vários processos de sinalização celular

(Rothermel et al., 2000 e Fuentes et al., 2000) e com as proteínas UXT que possui três sítios potenciais de fosforilação e com a APLP1 (Silveira et al., 2004) que estaria envolvida com o mal de Alzheimer. Iizuka et al. (2004) verificaram a interação entre a DSCR1 e a integrina  $\alpha v \beta 3$  e sugeriram que a DSCR1 está envolvida na angiogênese por regulação da adesão e migração de células endoteliais.

Analisando a sequência da proteína DSCR1 verifica-se que esta apresenta similaridade com algumas proteínas, entre elas a proteína codificada pelo gene *adapt 78* de hamster chinês, que é similar a forma alternativa 4 da DSCR1 (Crawford et al., 1997). Considerando que o gene *adapt 78* é induzível por peróxido de hidrogênio, e possivelmente envolvido no metabolismo oxidativo, é de se supor que a forma 4 da DSCR1 possa sofrer as mesmas influências visto que fortes evidências indicam que o estresse oxidativo esteja relacionado com a degeneração neural bastante acentuada em portadores da Síndrome de Down (Crawford et al., 1997).

Outra proteína que apresenta similaridade com a DSCR1 é a proteína ZAKI-4, positivamente regulada pelo hormônio tireoidiano T3 que pode ter um papel vital no desenvolvimento fetal, assim como durante a vida adulta (Miyazaki et al., 1996).

O gene humano MNBH/DYRK1 foi isolado da DSCR e é homólogo ao gene “minibrain” (*mnbr*) de *Drosophila*. Em *D. melanogaster* este gene codifica para uma serina-treonina-quinase que tem um papel essencial na neurogênese pós-embriônica. Guimera et al. (1996 e 1999) verificaram que o MNBH é superexpresso no cérebro de pacientes com Síndrome de Down e em diversas regiões do cérebro de rato, como o bulbo olfatório, cerebelo, córtex cerebral, camada de células piramidais do hipocampo e vários núcleos hipotalâmicos. Este padrão de expressão corresponde às regiões afetadas nos portadores da

SD indicando que a superexpressão deste gene deve resultar no fenótipo observado em indivíduos com trissomia do 21.

Outro gene de *Drosophila* que possui homólogos humanos é o *sim* (*single-minded*), que está envolvido no desenvolvimento do sistema nervoso central. Fan et al. (1996) identificaram dois homólogos deste gene denominados *sim1* e *sim2*. *Sim1* está localizado na região proximal do cromossomo 10 de rato enquanto o *sim2* está na porção distal final do cromossomo 16 que é sintênica a região crítica da Síndrome de Down. Recentemente Vialard et al. (2000) mostraram que o gene *sim2* é expresso durante o início do desenvolvimento fetal do sistema nervoso e em tecidos como palato, facial, ossos do crânio e nas vértebras sendo todas estas regiões afetadas nos portadores da síndrome.

Outro gene localizado entre as regiões D21S343 e D21S268 pertencente a DSCR foi identificado e caracterizado por Vidal-Taboada et al. (1998). O gene denominado oficialmente de DSCR2 foi anteriormente chamado de c21-LRP (chromosome 21 leucine-rich protein) devido a sua relativa abundância de aminoácidos leucina.

Análises da expressão gênica demonstraram que o DSCR2 é altamente expresso em testículo e em células leucêmicas da linhagem Jurkat, o que indica que este gene pode estar relacionado aos fenótipos apresentados pelos portadores de Down, já que estes indivíduos possuem um risco de 10 a 20 vezes maior do que o restante da população de desenvolverem leucemia. Além disso, a superexpressão de c21-LRP em testículos reforça esta ligação, uma vez que os pacientes masculinos com Down possuem a característica de subfertilidade ou esterilidade e hipogonadismo. Assim, esses resultados indicam que a função do DSCR2 pode estar relacionada à proliferação celular (Vidal-Taboada et al., 2000). Outros estudos foram realizados recentemente com a DSCR2 nos quais esta proteína foi expressa e purificada em fusão com a GST (Sommer, 2003) e poderá ser utilizada em estudos estruturais a fim de contribuir para a descoberta da função desta

proteína. Os dados a respeito da localização subcelular desta proteína indicam que ela está presente no retículo endoplasmático de células de mamíferos e não na membrana como sugerido por Vidal-Taboada et al. (1998) (Abrão Possik et al., 2004).

Recentemente os genes DSCR5 e DSCR6 também foram isolados da DSCR. DSCR5 e DSCR6 consistem de 5 e 6 éxons respectivamente e apresentam diferentes inícios de transcrição e eventos de *splicing* alternativo resultando em diferentes RNAs e, conseqüentemente proteínas em ambos os genes. O gene DSCR5 codifica para três diferentes proteínas preditas, apresenta um domínio transmembrana e é expresso em vários tecidos, principalmente fetal como, por exemplo, cérebro, coração e fígado. Por outro lado o DSCR6 codifica duas isoformas e é expresso em um número limitado de tecidos e em baixos níveis (Shibuya et al., 2000).

Outros dois genes localizados na região crítica da SD foram descritos por Nakamura et al. (1997a, b) e denominados DCRA e DCRB, respectivamente.

**Comentário:** A descrição do DCRA foi feita em um trabalho diferente da descrição do B.

O DCRA (Down Syndrome Critical Region gene A) ou DSCR3 (Down Syndrome Critical Region gene 3) contém 8 éxons e codifica uma proteína de 297 aminoácidos. Este gene é expresso em vários tecidos como cérebro fetal e adulto, coração, pulmão, fígado e rim. Além disso, possui similaridade com o DCRA de camundongo (92%) e com o gene H $\beta$ 58 de rato, essencial na embriogênese, o que indica que o gene DCRA possa estar relacionado com o desenvolvimento embrionário em humanos. Embora ainda não se conheça a função dessa proteína, a superexpressão do DCRA pode causar alterações, levando ao desenvolvimento de algumas das características da SD. Estudos recentes indicam que a localização subcelular desta proteína é citoplasmática (Corrêa, 2004) e que possui um domínio conservado de Vps26, uma proteína relacionada ao tráfego celular.

O gene DCRB (Down Syndrome Critical Region gene B) ou também denominado DSCR4 (Down Syndrome Critical Region gene 4) é objeto do presente trabalho. Ele contém 3 éxons e seu cDNA codifica uma proteína de 118 aminoácidos que não apresenta similaridade com nenhuma proteína já descrita. Análises do padrão de expressão do DCRB demonstraram expressão preferencial em placenta, embora também seja expresso em coração e músculo esquelético de adultos (Nakamura et al., 1997b).

Outro transcrito, o c21orf65 (DSCR8) também chamado de MMA-1 (malignant melanoma-associated 1) está localizado na DSCR. A expressão de MMA-1 foi detectada exclusivamente em testículo de amostras de tecidos normais, mas exibe um alto nível de expressão em tecidos tumorais especialmente o melanoma (de Wit et al., 2002). Outros transcritos denominados c21orf (*Homo sapiens* chromosome 21 open reading frame) tem sido mapeados nesta região e codificam proteínas de função desconhecida e alguns deles são analisados no presente trabalho.

Recentemente, Takamatsu et al. (2002) usando predições computacional de genes e subsequente clonagem do cDNA, identificaram dois novos genes designados DSCR9 e DSCR10 que estão na chamada Região Crítica da Síndrome de Down (DSCR). Ambos os genes são expressos preferencialmente em testículo e codificam proteínas desconhecidas de 149 e 87 resíduos de aminoácidos, respectivamente. Outras análises demonstram que estes genes são exclusivos de genomas de primatas. Assim, segundo os autores, esses genes teriam evoluído como uma nova classe de genes em primatas.

Engidawork et al. (2003) descreveram uma proteína hipotética de 28,5 kDa que é superexpressa em cérebro fetal de portadores da SD e sugerem uma possível associação desta proteína com a neuropatologia da síndrome. Esta proteína hipotética tem homologia com a proteína denominada Epoxide hidrolase, cuja expressão é induzida por estresse oxidativo, que é uma condição muito freqüente entre os pacientes com SD. Portanto,

devido a similaridade entre esta proteína e a proteína hipotética de 28,5 kDa os autores sugerem que o estresse oxidativo pode ser responsável pela superexpressão da proteína hipotética de 28,5 kDa observada em cérebro fetal de portadores da Síndrome de Down.

Enfim, os genes localizados na DSCR podem estar envolvidos com a Síndrome de Down e necessitam ser estudados detalhadamente a fim de possibilitar um melhor entendimento sobre o desenvolvimento da SD.

## 2. OBJETIVOS

Levando-se em conta a alta incidência e o impacto da Síndrome de Down na população humana mundial, estudos tanto em relação à origem como quanto ao desenvolvimento da anomalia são de grande importância.

Assim sendo, tornam-se objetivos deste trabalho:

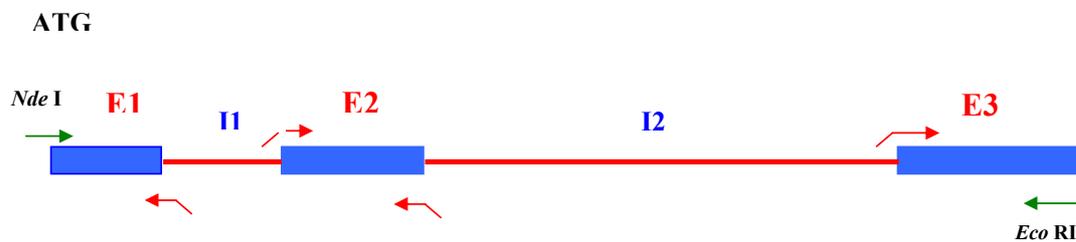
- a. Isolar o gene DCRB e as nove ORFs (Open Reading Frame) localizadas no cromossomo 21;
- b. Subclonagem dos genes de interesse em vetores para a expressão heteróloga em *Escherichia coli*;
- c. Purificação da proteína recombinante DCRB;
- d. Produção de anticorpos policlonais contra a DCRB humana em camundongo;
- e. Estudos iniciais sobre a estrutura secundária da proteína DCRB utilizando Dicroísmo Circular;
- f. Predição *in silico* da localização subcelular das proteínas estudadas;
- g. Localização subcelular das proteínas de interesse em células de mamíferos.

### 3. MATERIAL E MÉTODOS

#### 3.1 Obtenção do gene DCRB

O gene DCRB foi amplificado a partir de DNA genômico através de um método por nós denominado “splicing in vitro” que consiste em uma única reação em cadeia da polimerase (PCR) utilizando-se de conjuntos de "primers" que flanqueiam os éxons internos em adição a um par que flanqueia os dois éxons externos do gene a ser amplificado contendo sítios de restrição para clonagem.

O esquema da construção do gene é mostrado na figura abaixo.



**Figura 3: Esquema representativo da estratégia utilizada para isolar o gene DCRB.** O cDNA completo foi amplificado em um único passo, utilizando DNA genômico humano como molde e oligonucleotídeos específicos flanqueando os éxons DCRB, com sobreposições entre eles, na concentração final de 0,1 uM. Na mesma reação foram incluídos oligonucleotídeos que flanqueiam todo o cDNA, em concentração final de 1 uM. O oligonucleotídeo 5' flanqueador contém o sítio para a enzima *Nde I* e o 3' para a *Eco RI*.

#### 3.2 Avaliação da expressão da DCRB

Linhagens celulares tumorais, disponibilizadas pela ATCC (“American Type Culture Collection”) foram cultivadas e seus RNAs extraídos para avaliar o padrão de expressão da DCRB. Para tanto, foi obtido o maior número de linhagens tumorais possível, dos mais diferentes tipos de tecido. Foram elas: CaSki (carcinoma epidermóide da cérvix

uterina), Hela (adenocarcinoma de útero), A172 (glioblastoma), T98G (glioblastoma multiforme), HL-60 (leucemia), H358, (adenocarcinoma de pulmão), H1155 (adenocarcinoma de pulmão), Du145 (carcinoma de próstata), PC3 (adenocarcinoma de próstata), SCABER (carcinoma de células escamosas da bexiga), FaDu (carcinoma de células escamosas da faringe), MDA-436 (adenocarcinoma de mama), MCF-7 (adenocarcinoma de mama), SW-480 (adenocarcinoma coloretal), SaOS-2 (osteossarcoma), A2058 (melanoma), SKmel-28 (melanoma maligno) e placenta normal, que foi utilizada como controle positivo. Cada linhagem foi cultivada segundo as especificações do fornecedor.

### **3.2.1 Extração de RNA de linhagens celulares**

As linhagens tumorais de interesse foram cultivadas em meio apropriado até obtenção de massa suficiente ( $3 \times 10^6$  células) para a extração de RNA. Os RNAs da placenta normal e linhagens celulares foram extraídos pelo método de sedimentação em cloreto de céσιο, devido à boa quantidade e qualidade obtida. Seguindo-se esse método, a placenta, ainda congelada, foi macerada e homogeneizada em 9 ml de solução de lise (4M isotiocianato de guanidina/25mM citrato de sódio pH 7,0/0,1M  $\beta$ -mercaptoetanol). No caso das linhagens, o meio foi aspirado e a solução de lise adicionada à garrafa de cultura ( $160\text{cm}^2$ ). Em seguida, o lisado foi transferido para um tubo de ultracentrífuga contendo 4ml de solução de CsCl (5,7M CsCl e 25mM AcNa) e então centrifugado a 150.000g por 17 horas a 20°C (rotor SW40Ti, Beckman). Após a centrifugação, formou-se um precipitado (RNA) e o sobrenadante (proteínas e DNA) que foi descartado. Finalmente, a parede interna do tubo foi limpa e o precipitado de RNA solubilizado em 50 a 200 $\mu$ l de

água DEPC (di-etil pirocarbonato). A quantificação foi feita em seguida, em espectrofotômetro apropriado.

### **3.2.2 Avaliação da qualidade dos RNAs obtidos**

A qualidade dos RNAs foi avaliada de duas maneiras: verificou-se se houve degradação do material e se o mesmo estava contaminado com DNA genômico. A integridade dos RNAs foi visualizada aplicando-se 1µg de RNA total, em gel de agarose 1,0%. Antes de ser aplicado, o RNA foi desnaturado a 65° C por 5 minutos, sendo mantido em condição desnaturante em tampão de amostra contendo uréia. (7M uréia e 30% glicerol). Todos os procedimentos foram feitos com materiais próprios para uso exclusivo de RNA. A coloração foi feita com brometo de etídio e o gel visualizado sob luz UV. Desta maneira, foram considerados íntegros, os RNAs que apresentaram as bandas correspondentes aos RNAs ribossômicos 28S e 18S bem evidentes, sendo a primeira com o dobro da intensidade da segunda.

A contaminação com DNA genômico foi verificada através do teste de hMLH1 (human mut-L homologue 1). Este teste compreende uma PCR, na qual utiliza-se 200ng de RNA total e oligonucleotídeos (sense: 5' TGG TGT CTC TAG TTC TGG 3' e reverse: 5' CAT TGT TGT AGT AGC TCT GC 3') que se anelam aos íntrons 12 e 13 do gene hMLH1, de maneira que, qualquer amplificação presente será devido à contaminação com DNA (tamanho esperado do fragmento: 250pb). As seguintes condições de amplificação foram utilizadas: 1,5 mM MgCl<sub>2</sub>, 0,1 mM dNTPs, 1,6 µM de cada oligonucleotídeo, 1U de Taq DNA polimerase (Gibco), em tampão apropriado. A reação foi iniciada com desnaturação a 94°C por 5 minutos, seguida de 35 ciclos de: 45 segundos a 94°C, 45 segundos a 55°C e 1 minuto a 72°C, e extensão final de 6 minutos a 72°C. Os produtos de

amplificação foram visualizados em gel de poliacrilamida, corado em prata. Além disso, foi feito o tratamento com DNase (Gibco), segundo instruções do fabricante. Para confirmar se a contaminação havia sido eliminada, o teste de hMLH1 foi repetido.

### 3.2.3 RT-PCR

Para a síntese de cDNA foram utilizados 2 µg de RNA total e ainda: 0,025 µg/µl de oligo dT, 0,5 mM dNTPs em 12 µL de reação. Este foi incubado a 65°C durante 5 minutos e armazenado em gelo. Em seguida, foram adicionados 1X *First strand buffer*, 0,01M DTT, 200U de RNase *out-inhibitor* e 40U de SuperScript II (Gibco), em volume final de 20 µl. A reação seguiu-se com incubação a 42°C por 1 hora e depois a 70°C por 15 minutos para inativação da enzima.

Antes do cDNA ser utilizado na PCR de interesse, foi feito um teste para avaliar a qualidade do mesmo. Para tal, fez-se uma PCR utilizando-se oligonucleotídeos específicos para os éxons 6 e 7 do gene GAPDH (gliceraldeído desidrogenase) (sense: 5' CTG CAC CAC CAA CTG CTT A 3' e reverse: 5' CAT GAC GGC AGG TCA GGT C 3') que foi utilizado como controle positivo. A reação foi feita nas seguintes condições: 1µl de cDNA, 1,0mM MgCl<sub>2</sub>, 0,1mM dNTPs, 0,5µM de cada oligonucleotídeo e 1U de Taq *DNA* polimerase (Gibco), em tampão apropriado e volume final de 20µl, iniciando-se com desnaturação a 94°C por 5 minutos, seguida de 22 ciclos de: 1 minuto a 94°C, 45 segundos a 60°C e 1 minuto a 72°C, e extensão final de 10 minutos a 72°C.

Para o DCRB foram usadas as seguintes condições: 94°C por 3 minutos, seguida de 30 ciclos de: 30 segundos a 94°C, 45 segundos a 56°C e 1 minuto a 72°C.

Todos os procedimentos realizados para a avaliação da expressão da DCRB foram realizados com a colaboração do doutorando Raphael Bessa Parmigiani e sua orientadora Profa Dra Ana Maria Camargo do Instituto Ludwig de São Paulo.

### **3.3 Descrição dos vetores utilizados**

#### **pET28a (Novagem)**

Este vetor possui uma seqüência na região N-terminal do sítio de múltipla clonagem que codifica seis histidinas denominadas “His-Tag”. Assim o gene de interesse é clonado em fusão com esta seqüência que facilita a purificação da proteína expressa em cromatografia de afinidade. A cauda de histidinas pode ser removida após a clivagem com a trombina. A seleção das bactérias é feita com o antibiótico canamicina.

#### **pET37b (Novagen)**

Este vetor possibilita a clonagem em fusão com a proteína CBD (“cellulose binding domain”) e S-Tag. Além disso, apresenta uma seqüência sinalizadora que faz com que a proteína seja enviada para o espaço periplasmático. Esta característica facilita a purificação da proteína de interesse através de choque osmótico. Este vetor utiliza o promotor do bacteriófago T7 sendo que a transcrição da proteína recombinante é feita pela T7 RNA polimerase. A proteína de fusão e a proteína de interesse podem ser separadas pela clivagem com o Fator Xa e Trombina. A seleção é feita pelo antibiótico canamicina.

#### **pBAD TOPO (Invitrogen)**

Este vetor foi desenvolvido para a rápida clonagem de produtos de PCR amplificados com a Taq polimerase. A ligação inserto-vetor é catalisada pela

topoisomerase I que tem a alta capacidade de clivar e religar dupla fita de DNA. A expressão em *E. coli* é dirigida pelo promotor araBAD que é aumentada na presença de L-arabionose. A purificação da proteína de interesse pode ser feita através de cromatografia de afinidade. A seleção das bactérias é feita pelo antibiótico ampicilina.

#### **pCX TOPO (Invitrogen)**

Este vetor permite a rápida clonagem dos produtos de PCR amplificados com Taq polimerase que são ligados através da ação da topoisomerase. A proteína RsaA em fusão com o C-terminal da proteína alvo é direcionada e secretada como agregados protéicos para o meio de cultura de *Caulobacter crescentus*. A proteína recombinante pode ser purificada com 90% de pureza por um simples passo de filtração. A seleção é feita pelo antibiótico cloranfenicol.

#### **pTYB2 (New England Biolabs)**

Este plasmídeo utiliza o promotor T7 e permite a clonagem do gene de interesse em fusão com a seqüência que codifica a inteína. Esta proteína possui a capacidade de autoclivagem na região N-terminal na presença de tióis como DTT,  $\beta$  mercaptoetanol e cisteína. Na porção C-terminal da inteína está localizada a proteína CBD (“chitin binding domain”) que tem afinidade por quitina, possibilitando a fácil purificação da proteína de interesse sem a necessidade do uso de proteases. A seleção é feita pelo antibiótico ampicilina.

### **pcDNA3 (Invitrogen)**

Por possuir o promotor CMV (Human Cytomegalovirus), este plasmídeo permite altos níveis de expressão de proteínas recombinantes em células de mamíferos. Além disso, possui um sinal SV40 que permite a terminação eficiente da transcrição e poliadenilação do RNA mensageiro e o BGH (Bovine Growth Hormone) que também codifica um sinal de poliadenilação e uma seqüência de terminação de transcrição que aumenta a estabilidade do RNA mensageiro. A seleção de transfectantes estáveis em células de mamíferos é feita pelo antibiótico geneticina e a ampicilina é utilizada para a seleção do vetor em *E. coli*.

As cepas bacterianas utilizadas tanto na propagação quanto na expressão foram as compatíveis com cada sistema selecionado.

## **3.4 Construção dos vetores de expressão**

### **3.4.1 Construção do vetor pET28aDCRB**

O produto amplificado do gene DCRB, assim como o vetor pET28a (Novagen), foram clivados com as enzimas de restrição *Nde* I e *Eco* RI e posteriormente purificados de um gel de agarose 1-1,5% utilizando-se o kit Concert<sup>tm</sup> Rapid Gel Extraction System (Gibco BRL).

Em seguida plasmídeo e inserto foram ligados e posteriormente células competentes de *E. coli* DH5- $\alpha$  (CaCl<sub>2</sub>) (Sambrook e Russel, 2001) foram transformadas com o produto de ligação e plaqueadas em LB ágar contendo o antibiótico canamicina (25 mg/ml).

A confirmação dos recombinantes foi realizada através da análise de restrição e o clone selecionado foi sequenciado pelo método dideoxi (Sanger, 1977) em um

sequenciador automático modelo ABI Prism 377 DNA Sequencer. Após o sequenciamento, para verificar a integridade do gene e a junção vetor-inserto, a seqüência obtida foi submetida ao alinhamento através do software Multalin versão 5.4.1 (<http://www.toulouse.inra.fr/multalin.html>) (Corpet et al., 1988).

O gene DCRB clonado no pET28a foi expresso em BL21(DE3) e co-expresso em BL21(DE3) pT-GroE. As células transformadas foram cultivadas em meio seletivo contendo o antibiótico adequado e a expressão da proteína de interesse foi induzida por 4 horas na fase logarítmica do crescimento por isopropil- $\beta$ -D-tiogalactopiranosídeo (IPTG) na concentração final de 0,4 e 1mM separadamente.

#### **3.4.2 Construção do vetor pET37bDCRB**

O plasmídeo pET37b foi clivado com a enzima de restrição *Nco* I e tratado com a nuclease “Mung-bean” de acordo com as especificações do fabricante de maneira que esta extremidade ficasse não coesiva. Em seguida o plasmídeo foi clivado com a enzima *Eco* RI, gerando um plasmídeo com extremidades não coesiva/*Eco* RI.

O inserto foi retirado do plasmídeo pET28aDCRB através da clivagem com a enzima *Nde* I, preenchido com dNTPs pela atividade da Klenow, gerando uma extremidade não coesiva e, em seguida, clivado com *Eco* RI, gerando um inserto com extremidades compatíveis as do plasmídeo pET37b clivado.

O plasmídeo e inserto foram purificados de gel de agarose 1-1,5% utilizando-se o Kit Concert<sup>tm</sup> Rapid Gel Extraction System (Gibco BRL) e ligados. Posteriormente, células  $\text{CaCl}_2$  competentes de bactérias *E. coli* DH5- $\alpha$  foram transformadas e plaqueadas em meio seletivo contendo o antibiótico canamicina (25 mg/ml)

A seleção dos recombinantes foi realizada através de análise de restrição e confirmados através de sequenciamento como descrito anteriormente.

Após extração plasmidial, os recombinantes denominados pET37bDCRB foram utilizados para transformar cepas bacterianas de expressão do tipo *E. coli* BL21 (DE3) competentes. Uma das colônias transformantes foi repicada em 3 ml de meio líquido LB contendo o antibiótico adequado (pré-inóculo) o qual permaneceu sob agitação a 200 rpm a 37°C durante 12 horas.

O pré-inóculo foi então diluído em 10 ml de meio LB líquido contendo o antibiótico adequado permanecendo sob agitação a 37°C por aproximadamente 3 horas até atingir densidade ótica de 0,8 em absorbância de 600nm. Neste momento, a expressão foi induzida com a adição de IPTG numa concentração final de 0,4 mM.

### **3.4.3 Construção do vetor pBADTOPODCRB**

Para esta clonagem foram sintetizados oligonucleotídeos (Life Technologies) sendo que o sense foi construído a partir do códon de iniciação ATG original do gene e o reverse apresenta deleção para o códon de terminação e a inserção de um novo sítio para a enteroquinase (em azul sublinhado). A seguir a sequência dos oligonucleotídeos:

Oligo sense: 5' ATG TCG TTA ATC ATC TTG AC 3'

Oligo reverse 5' CTT GTC GTC GTC GTC GGT TGA TGG GCT TGC 3'

A amplificação do inserto foi feita a partir do molde pET28aDCRB e em seguida o produto de amplificação foi purificado de gel de agarose 1,5% utilizando-se o kit NucleoSpin Extract (Bioeasy).

A ligação foi realizada conforme orientação do fabricante e então, células Top10 quimicamente competentes foram transformadas e plaqueadas em meio LB ágar contendo

o antibiótico adequado. A seleção dos recombinantes foi realizada através de PCR de colônias e confirmados através de sequenciamento como descrito anteriormente.

Após a seleção das colônias recombinantes, realizou-se a extração plasmidial e em seguida, estes foram utilizados na transformação de *E. coli* LMG 194 para os teste de expressão.

As colônias transformantes foram inoculadas em meio RM e ampicilina (100 mg/ml) e cultivadas até atingir densidade ótica de 0,5 a 600 nm de absorbância, quando foram induzidas através da adição de diferentes concentrações de L-arabinose e mantido sob agitação a 200 rpm a 37°C por 4 horas.

#### **3.4.4 Construção do vetor pCX-TOPODCRB**

A amplificação do DCRB foi realizada utilizando-se dos mesmos oligonucleotídeos e nas mesmas condições descritas no item 3.4.3. Em seguida o produto de amplificação foi purificado utilizando-se o kit Concert Rapid PCR Purification System (Gibco BRL) e então ligado ao plasmídeo conforme orientação do fabricante. Em seguida, células Top 10F' quimicamente competentes foram transformadas com o produto de ligação e plaqueadas em meio LB ágar contendo o antibiótico adequado. A seleção dos recombinantes foi realizada através de PCR de colônias e confirmados através de sequenciamento como descrito anteriormente.

Colônias recombinantes foram selecionadas e utilizadas na transformação de células competentes de B5 BAC (*Caulobacter crescentus*) para a análise da expressão protéica. Estas colônias foram inoculadas em meio PYE líquido e cultivadas sob agitação a 30°C por 18 horas.

Após este período, esta cultura foi transferida para o meio M11 e incubada a 30°C sob agitação suave (100 rpm) durante no máximo cinco dias.

#### **3.4.5 Construção do vetor pTYB2DCRB**

O plasmídeo pTYB2 foi linearizado com as enzimas *Nde* I e *Sma* I e purificado resultando num plasmídeo com extremidades *Nde* I/*Sma* I.

O inserto DCRB foi amplificado utilizando-se os oligonucleotídeos sense com o sítio de *Nde* I e o reverse utilizado para a clonagem nos plasmídeos pBAD e PCX-TOPO. O produto amplificado foi purificado usando o kit Concert Rapid PCR Purification System (GIBCO BRL), clivado com *Nde* I e novamente purificado. O produto de PCR clivado e purificado foi ligado ao plasmídeo pTYB2 e em seguida, bactérias *E. coli* DH5- $\alpha$  CaCl<sub>2</sub> competentes foram transformadas e plaqueadas em meio LB ágar contendo o antibiótico ampicilina. A seleção dos recombinantes foi realizada através de PCR de colônias e confirmados através de sequenciamento como descrito anteriormente.

Posteriormente a seleção da colônia recombinante foi realizada a extração de DNA plasmidial que, em seguida, foi utilizado na transformação de bactérias competentes de *E. coli* ER2566.

Uma das colônias transformadas foi cultivada sob agitação a 200 rpm a 37°C até atingir densidade ótica de 0,5 a 600nm de absorbância, quando foi induzida com a adição de IPTG numa concentração final de 0,3 mM.

#### **3.4.6 Construção do vetor pcDNA3DCRB**

O plasmídeo pcDNA3 foi clivado com as enzimas *Eco* RV e *Not* I e purificado resultando num plasmídeo com extremidades não coesiva/*Not* I

O inserto foi retirado a partir do plasmídeo pET28aDCRB através da clivagem com a enzima *Xba* I, preenchido com dNTPs pela atividade da Klenow, gerando uma extremidade não coesiva e, em seguida, clivado com *Not* I, gerando um inserto com extremidades compatíveis a do plasmídeo pcDNA3. Como este vetor não possui nenhum tipo de fusão que possibilita a purificação em cromatografia de afinidade, esta estratégia permitiu que o gene DCRB pudesse ser clonado fusionado a seqüência que codifica a cauda de histidina do pET28a para a posterior purificação da proteína estudada.

O plasmídeo e inserto foram purificados de gel de agarose 1-1,5% utilizando-se o Kit Concert<sup>™</sup> Rapid Gel Extraction System (Gibco BRL) e ligados. Posteriormente, células CaCl<sub>2</sub> competentes de bactérias *E. coli* DH5- $\alpha$  foram transformadas e plaqueadas em meio seletivo.

A seleção dos recombinantes foi realizada através de PCR de colônias e confirmados através de sequenciamento como descrito anteriormente.

Um dos plasmídeos recombinantes foi utilizado para a transfecção de células de mamíferos pelo método de transferência mediada por lipídeos catiônicos (como descrito posteriormente) e as células transfectantes foram selecionadas com o antibiótico geneticina (500 ug/ml) durante um mês. Após esse período as células foram cultivadas em meio seletivo com geneticina (100 ug/ml) até atingir a confluência desejada e então foram soltas com uma solução de tripsina e utilizadas para a análise de expressão e solubilidade da proteína de interesse.

### **3.5 Mutação sítio dirigida do gene DCRB**

A mutação sítio dirigida é uma técnica comumente utilizada para estudos de estrutura e função de proteínas.

Para a realização da mutação sítio dirigida foi utilizado o kit U.S.E. Mutagenesis (Amersham Biosciences) que usa dois sistemas de oligonucleotídeos para gerar sítios de mutação específicos.

Assim foram utilizados os seguintes oligonucleotídeos:

Oligo de seleção (kit): 5' CTC TGA CTG GTG ACG CGT CAA CCA AGT C 3'.

Este oligonucleotídeo altera o sítio de restrição da enzima *Sca* I para *Mlu* I contido no vetor utilizado (em vermelho sublinhado).

Oligo da mutação (sintetizado): 5' TTT GGG CAA AAG TTC TGA TTT GGA A 3'. Este oligonucleotídeo insere a mutação desejada (em azul sublinhado).

Para posterior seleção do plasmídeo mutante, o oligonucleotídeo de mutação e o de seleção devem ser desenhados de maneira que seja complementar a mesma fita do plasmídeo molde.

Assim, o primeiro passo foi a realização da síntese da fita mutante. Para isso foi feita uma reação de amplificação contendo o plasmídeo molde (pET28aDCRB), os oligonucleotídeos de seleção e mutação, tampão e água. Esta solução foi incubada por 5 minutos a 100°C e 0°C respectivamente e logo em seguida, mantida a temperatura ambiente por 30 minutos. Após esse período, foram acrescentados uma mistura de nucleotídeos (kit) e uma mistura de reação (kit), que contém a polimerase. Esta solução foi incubada a 37°C por uma hora e em seguida a 85°C por 15 minutos.

Posteriormente foi realizada a primeira seleção com a enzima de restrição, na qual a reação descrita acima foi incubada por 2 horas a 37°C com 10U de *Sca* I. Nesse passo o plasmídeo selvagem, no caso o pET28aDCRB, foi linearizado e o mutante mantido íntegro já que neste, o sítio para a enzima *Sca* I foi alterado para *Mlu* I pelo oligo de seleção.

De posse do plasmídeo mutante, foi realizada a transformação de bactérias *E. coli* NM522*mutS* eletrocompetentes e, em seguida, a extração do DNA plasmidial. Este DNA foi então submetido a uma nova seleção após a incubação com 15U da enzima *Sca I* a 37°C por 2 horas, que tem a finalidade de garantir uma maior proporção de plasmídeos mutantes.

Novamente foi realizada a transformação de bactérias *E. coli* DH5- $\alpha$  quimicamente competentes utilizando-se a mistura da reação da segunda seleção e então plaqueadas em meio LB ágar contendo o antibiótico canamicina (25 mg/ml).

Em seguida, 20 clones foram selecionados e submetidos à extração plasmidial e posteriormente analisados através de sequenciamento pelo método dideoxi (Sanger et al., 1977) em um sequenciador automático modelo ABI Prism 377 DNA Sequencer, para verificar a inserção da mutação desejada.

### 3.6 Deleção de parte da proteína DCRB

Resolvemos também deletar parte da região codificadora do C-terminal da proteína DCRB e para isso foi realizada uma reação de PCR utilizando-se como molde o plasmídeo pET28aDCRB e os seguintes oligonucleotídeos:

Oligo sense: 5' TCG AAG GTC GTC ATA TGT CGT TAA TCA TCT TGA C3'

Oligo reverse: 5' AAG AAT TCA AGA AGC TGG CAG G 3'

As letras em verde e azul sublinhadas são correspondentes aos sítios das enzimas *Nde I* e *Eco RI* respectivamente.

O produto amplificado, denominado por nós de DCRB 2/3, foi ligado no vetor pGEM-T (Promega) e os transformantes foram analisados por restrição com as enzimas *Nde I* e *Eco RI*. O inserto correspondente ao DCRB 2/3 foi posteriormente purificado de

um gel de agarose 1,5% utilizando-se o kit Concert<sup>tm</sup> Rapid Gel Extraction System (Gibco BRL).

Em seguida o inserto foi subclonado no vetor pET28a clivado com as mesmas enzimas descritas acima e posteriormente células competentes de *E. coli* DH5- $\alpha$  (CaCl<sub>2</sub>) foram transformadas e plaqueadas em LB ágar contendo o antibiótico adequado.

A confirmação dos recombinantes foi realizada através da análise de restrição e posterior sequenciamento.

Tanto o gene DCRB mutado quanto o DCRB 2/3 foram expressos em *E. coli* BL21(DE3). As células transformadas foram cultivadas em meio seletivo contendo o antibiótico adequado e a expressão da proteína de interesse foi induzida por 4 horas na fase logarítmica do crescimento por isopropil- $\beta$ -D-tiogalactopiranosídeo (IPTG) na concentração final de 0,4 mM.

### **3.7 Teste de solubilidade**

Para a determinação da forma de expressão de proteínas recombinantes é necessário que se faça um teste de solubilidade no qual será possível verificar se a proteína de interesse foi expressa na forma solúvel ou insolúvel. Para isso, as células são cultivadas e induzidas a expressar a proteína de interesse. Após esse período as células são precipitadas por centrifugação e posteriormente ressuspensas em um tampão adequado, lisadas e novamente centrifugadas. Assim, as proteínas são separadas em duas frações, sobrenadante (forma solúvel) e precipitado (forma insolúvel). A seguir a solubilidade da proteína de interesse é analisada em SDS-PAGE.

O plasmídeo de expressão pET37B requer uma forma particular de realização deste teste que é através do choque osmótico descrito a seguir.

Após a verificação da expressão da proteína de interesse, o próximo passo é a realização do teste de solubilidade. Para isso um transformante de *E. coli* BL21 (DE3) foi inoculado como descrito anteriormente e utilizado como pré-inóculo.

O pré-inóculo (5 ml) foi então diluído em 495 ml de meio LB líquido contendo o antibiótico adequado e permaneceu sob agitação a 200 rpm a 37°C por aproximadamente 3 horas até atingir densidade ótica de 0,8 em absorvância de 600nm, quando a expressão foi induzida através da adição de IPTG numa concentração final de 0,4 mM.

Após a indução durante 4 horas a cultura foi centrifugada a 6.500g por 15 minutos a 4°C. O precipitado resultante foi submetido ao choque osmótico.

Para este choque, o precipitado foi ressuspense em 400 ml de tampão 1 (Tris-HCl 30mM; sacarose 20%, pH 8,0) e então adicionou-se 1ml de EDTA 500 mM e a solução resultante permaneceu no agitador por 10 minutos. Em seguida, centrifugou-se a 10.000g a 4°C por 10 minutos. As bactérias foram ressuspensas em 400 ml de MgSO<sub>4</sub> 5mM gelado e permaneceu no gelo sob agitação suave por 10 minutos e em seguida foram novamente centrifugadas (10.000g a 4°C por 10 minutos) e coletado o sobrenadante (fração periplasmática) para a análise em gel de poliacrilamida.

O precipitado foi novamente ressuspense em Tris-HCl 20mM pH 7,5 gelado e incubado a 30°C por 15 minutos. Após sonicação (10 pulsos de 1 minuto por 1 minuto de descanso), a amostra lisada foi centrifugada a 14.000g, por 10 minutos a 4°C, e amostras do sobrenadante e do precipitado foram analisadas em SDS-PAGE 10% corado com azul de Coomassie.

### **3.8 Purificação de proteínas**

Para que se possa determinar a estrutura e função de uma determinada proteína é desejável que esta esteja na sua forma pura e sem a formação de agregados protéicos. A purificação de proteínas é realizada através da combinação de vários procedimentos tais como cromatografias de afinidade, de exclusão molecular, de troca iônica entre outros.

#### **3.8.1 Cromatografia de afinidade**

Foram realizados testes de purificação com as proteínas DCRB e DCRB 2/3 utilizando-se a técnica de cromatografia de afinidade que tem por princípio a imobilização de um ligante a uma matriz de agarose. Assim, foi utilizada uma coluna contendo níquel imobilizado (ligante) [Ni-NTA (Qiagen)] para o qual as histidinas, codificadas pela seqüência presente no vetor, tem afinidade. Através dessa interação a proteína de interesse fica retida na coluna. A eluição da proteína de interesse é realizada adicionando-se um competidor, no caso o imidazol, ao ligante de maneira que essa seja liberada da matriz.

Após o empacotamento de 5 ml de resina de níquel em uma coluna de 15 ml da Bio-Rad, esta foi lavada com 5 volumes de água milliQ autoclavada e equilibrada com 5 volumes de tampão de sonicação (Tris-HCl 10mM pH 9,0; NaH<sub>2</sub>PO<sub>4</sub> 50mM e NaCl 100 mM). Em seguida, foi realizada a aplicação de 25 ml da amostra na coluna que posteriormente, a fim de se retirar as proteínas contaminantes, foi lavada com 3 volumes de tampão de sonicação. Para a eluição foi utilizado um gradiente de 10 a 250 mM de imidazol em tampão de sonicação. As amostras provenientes da eluição foram analisadas em SDS-PAGE 15%.

Posteriormente as amostras das proteínas DCRB e DCRB 2/3 foram concentradas utilizando-se uma membrana centricon 10 (centricon de poro de 10kDa) e finalmente

submetidas à cromatografia de exclusão molecular e a medidas de Dicroísmo Circular (CD) respectivamente.

### **3.8.2 Cromatografia de exclusão molecular**

Esta técnica utiliza colunas empacotadas com um gel contendo esferas de tamanho específico que formam uma rede tridimensional proporcionando assim, a separação das moléculas em solução de acordo com o tamanho. Portanto, quando uma solução protéica é passada por este tipo de coluna, as moléculas maiores que os poros não vão se difundir por eles e eluem precocemente. Porém, as moléculas menores passam através dos poros e por isso são eluídas tardiamente.

A coluna Superose 12 (Amersham Biosciences) é uma coluna de exclusão molecular e foi utilizada neste trabalho. Assim a coluna foi previamente lavada com água milliQ autoclavada e posteriormente equilibrada com 2 volumes do tampão de sonicação. Amostras obtidas na cromatografia de afinidade foram dialisadas e concentradas a 250 µl e aplicadas na Superose 12 operada em sistema HPLC. O fluxo foi de 1,0 ml/minuto e a eluição foi realizada utilizando-se o mesmo tampão.

Posteriormente a amostra da proteína DCRB foi concentrada utilizando-se centricon 10 e finalmente submetidas a medidas de Dicroísmo Circular (CD).

### **3.8.3 Obtenção da proteína DCRB em condições desnaturantes**

Após o isolamento dos corpos de inclusão com PBS (1X), contendo Triton X-100 a 1%, estes foram solubilizados em tampão contendo Tris-HCl 100 mM, pH 8,5; NaCl 100mM e 6M de uréia e incubados a temperatura ambiente por trinta minutos. Em seguida, a amostra foi diluída em mesmo volume de tampão redutor (Tris-HCl 0,05M, pH 8,5; DTT

10mM) e novamente incubada a temperatura ambiente por trinta minutos. A seguir, a solução foi diluída em um volume de tampão de oxidação (Tris-HCl 0,05M, pH 8,5; uréia 3M, cisteína 5mM e cistina 1mM) e esta solução foi dialisada contra um tampão contendo Tris-HCl 0,05M, pH 8,5; uréia 1M, cisteína 5mM e cistina 1mM. A cada 12 horas o tampão de diálise foi trocado por um novo de mesma constituição, porém sem uréia. Este procedimento foi realizado por 3 a 4 vezes ou até que a solução não contivesse mais uréia.

Após o renovelamento, a amostra foi novamente dialisada contra um tampão contendo apenas Tris-HCl 10mM pH8,5, para a retirada do par redox cisteína-cistina.

Posteriormente a amostra foi concentrada utilizando-se centricon 10 e finalmente submetida a medidas de Dicroísmo Circular (CD) para monitorar o renovelamento.

Esses experimentos foram realizados em colaboração com a Profa Dra Ana Paula Ullian de Araújo do Grupo de Biofísica do Instituto de Física da USP-São Carlos.

### **3.9 Medidas de CD (Circular Dichroism –CD)**

As medidas de CD foram realizadas em um espectropolarímetro Jasco J715 (Jasco Corporation, Japão). Os espectros foram obtidos usando soluções de proteína em uma cubeta cilíndrica de quartzo, de 1mm de caminho óptico. Os espectros foram registrados num intervalo de 195 a 250 nm, a temperatura ambiente, com uma média de 16 varreduras.

Todas as análises de CD foram realizadas com o auxílio da Dra Patrícia Campana em colaboração com a Profa Dra Leila Maria Beltramini do Grupo de Biofísica do Instituto de Física da USP-São Carlos.

### **3.10 Produção de anticorpos policlonais**

A produção de anticorpos policlonais foi realizada através da aplicação da proteína DCRB em camundongos. Para isto, a proteína DCRB foi extraída de um gel de

poliacrilamida da seguinte maneira: amostras das proteínas totais foram aplicadas num gel de poliacrilamida 15% e o fragmento do gel contendo a banda correspondente à proteína DCRB com aproximadamente 13kDa foi recortado e macerado em 500 µl de água destilada incubada a 10°C durante a noite. Em seguida este sobrenadante foi passado por uma resina Sephadex G25 (Amersham Pharmacia) para a retirada de restos de acrilamida.

A proteína resultante foi quantificada pelo método de Bradford (1976) e através de injeção subcutânea, procedeu-se a imunização de quatro exemplares machos de camundongos com 50µl da proteína “purificada” (2µg/µl) acrescida de adjuvante completo (Freund’s Adjuvant complete - Sigma) na razão de 1:1. Após 45 dias foi realizada nova imunização na qual foram aplicados 50µl da proteína “purificada” (2µg/µl) acrescida de adjuvante incompleto (Freund’s Adjuvant incomplete - Sigma) na razão de 1:1.

Amostras de sangue dos animais foram retiradas antes e após a imunização, centrifugadas a 13.000 rpm por 15 minutos para a separação do soro dos animais e estocados na geladeira.

### **3.10.1 Detecção de anticorpos**

Após dez dias da segunda dose da proteína, amostras de soro dos camundongos foram submetidas a uma análise de “western blot”, na qual amostras de proteína foram aplicadas em um gel de poliacrilamida 15% e submetidas à eletroforese durante 1 hora e 40 minutos a 140V e 250mA. Em seguida o gel, a membrana de nitrocelulose e dois papéis de filtro foram imersos num tampão de transferência (Tris 200mM; glicina 50mM) e então montados como um “sanduíche” e novamente submetidos à eletroforese (2h, 140V, 160mM) para que ocorresse a transferência das proteínas do gel para a membrana de nitrocelulose. Terminada a transferência, a membrana foi corada durante 5 minutos com a

solução de Ponceau (Ponceau 0,5%, ácido acético 0,1%) para a visualização e marcação das bandas de interesse.

Após a coloração, a membrana foi incubada em um tampão de bloqueio (leite em pó desnatado; TBS 1X) durante a noite a 10°C. No dia seguinte, esta membrana foi lavada por quatro vezes, durante cinco minutos cada, com TBS (1X) (Tris 70mM; NaCl 100mM). Em seguida procedeu-se a incubação com soro proveniente do camundongo imunizado na razão de 1:10.000 em TBS (1X) por duas horas. Novamente a membrana foi lavada com TBS (1X) como descrito anteriormente e então incubada com o anticorpo “Anti-IGg mouse” (Sigma) na razão de 1:10.000 em TBS (1X) por duas horas e lavada novamente com TBS (1X). Após este procedimento a membrana foi revelada com uma solução NBT-BCIP AP (“Western Blue”-Sigma).

### **3.10.2 Titulação de anticorpos**

A titulação dos anticorpos foi realizada em placas de 96 amostras e através da técnica denominada ELISA (“Enzyme Linked Immunosorbent Assay”). Esta técnica se baseia na reação entre antígeno-anticorpo contra a proteína de interesse, sendo que o produto da reação é medido colorimetricamente através de um espectrofotômetro a 490nm.

Foi realizada a fixação de 50ng de proteína de interesse em 100 µl de Na<sub>2</sub>CO<sub>3</sub> 50mM pH 9,6 em 16 poços da placa por 24 horas a 10°C. Após este período, cada poço foi lavado por 6 vezes com PBS (1X) acrescido de Tween 20 0,05% e então bloqueado por 2 horas a 37°C com BSA 1% diluído em PBS (1X) acrescido de Tween 20 0,05%. Novamente os poços foram lavados por 2 vezes com PBS (1X) 0,05% Tween 20 e incubados por 2 horas a 37°C com o soro do camundongo diluído em PBS (1X) conforme a tabela 1. Após nova lavagem como descrito anteriormente, foi feita a incubação com

“Anti-IGg mouse” (Sigma) na razão de 1:4.000 em PBS (1X) durante 2 horas e então revelado com uma solução NBT-BCIP AP (“Western Blue”-Sigma). A leitura das placas foi realizada no leitor de placas DYNEX a 490nm.

**Tabela 1: Diluições realizadas para determinação da titulação do anticorpo.**

<b>Amostras</b>	<b>Concentração</b>	<b>Diluição</b>
Controle	1:50	20µl soro (camundongo não imunizado) + 1000µl PBS (1X)
C1	1:50	20µl soro (camundongo imunizado) + 1000µl PBS (1X)/Tween20 (C1)
C2	1:100	0,5 ml (C1) em 0,5 ml PBS (1X)/Tween20 (C2)
C3	1:500	2µl soro em 1000µl de PBS (1X)/Tween20 (C3)
C4	1:1.000	0,5 ml (C3) em 0,5 ml PBS (1X)/Tween20 (C4)
C5	1:2.000	0,5 ml (C4) em 0,5 ml PBS (1X)/Tween20 (C5)
C6	1:4.000	0,5 ml (C5) em 0,5 ml PBS (1X)/Tween20 (C6)
C7	1:8.000	0,5 ml (C6) em 0,5 ml PBS (1X)/Tween20 (C7)
C8	1:16.000	0,5 ml (C7) em 0,5 ml PBS (1X)/Tween20 (C8)
C9	1:32.000	0,5 ml (C8) em 0,5 ml PBS (1X)/Tween20 (C9)
C10	1:64.000	0,5 ml (C9) em 0,5 ml PBS (1X)/Tween20 (C10)
C11	1:96.000	0,75 ml (C10) em 0,5 ml PBS (1X)/Tween20 (C11)
C12	1:128.000	0,5 ml (C11) em 0,5 ml PBS (1X)/Tween20 (C12)
C13	1:192.000	0,75 ml (C12) em 0,5 ml PBS (1X)/Tween20 (C13)
C14	1:256.000	0,5 ml (C13) em 0,5 ml PBS (1X)/Tween20 (C14)
C15	1:512.000	0,5 ml (C14) em 0,5 ml PBS (1X)/Tween20 (C15)

### 3.11 Amplificação, clonagem e expressão de oito ORFs do cromossomo 21

As ORFs foram selecionadas segundo os seguintes critérios: tamanho do gene, presença de homologia em camundongo, presença de domínios protéicos conhecidos e nível de expressão. Os dados referentes a estas ORFs estão sumarizados na tabela a seguir:

**Tabela 2: Características das novas ORFs do cromossomo 21 selecionadas para estudos.**

ORF	Isoforma	Aminoác.	Identidade	Domínios	Nível de Expressão
c21orf 18	Forma B	439	-----	-----	-----
c21orf 45	-----	233	Presente	Ausente	Alto
c21orf 59	-----	-----	-----	-----	Alto
c21orf 63/PRED34	-----	440 splice alternativo (70aa)	Homólogo F32A73A <i>C. elegans</i>	Binding D-galactoside L-rhamnose <b>Pfam PF02140</b>	Alto
c21orf 83	A, B e C	353/298	88%	Zinc finger do tipo C <sub>2</sub> H <sub>2</sub> <b>Pfam domain PF00096</b>	Alto
c21orf 95 (CYYR1)	-----	154	84%	Domínio rico em prolina no carboxi terminal <b>Prosite signature PS50099</b>	Alto
c21orf 101 (MRPS6)	-----	125	86%	Membro da família da proteína ribossomal S6 <b>InterProdomain IPR000529</b>	Alto
c21orf 102	-----	257	80%	Domínio rico em leucina	-----

Fonte: Reymond et al. (2001); Gardiner et al. (2002) e Reymond et al. (2002).

Para a amplificação das ORFs foram realizadas reações de PCR utilizando-se como molde uma biblioteca de cérebro fetal humano (Gibco BRL), um painel de cDNA de múltiplos tecidos (Clontech) além dos seguintes oligonucleotídeos:

**c21orf 18**

Oligo sense: 5' TGC TAG CCAT GCA GAA AGG AAA AGG GAG 3'

Oligo reverse: 5' GGA TCC TCA GGT AAA AGC TGT TTG CA 3'

**c21orf 45**

Oligo sense: 5' TCA TAT GGC AGG CGT TCG GTC ACT 3'

Oligo reverse: 5' GGA TCC TCA GCT TTT ACA AGT GGC AA 3'

**c21orf 59**

Oligo sense: 5' TCA TAT GGT TCT GCT GCA CGT GAA 3'

Oligo reverse: 5' GGA TCC TCA TCT TGG TCT CCA CTT TA 3'

**c21orf 63**

Oligo sense: 5' TCA TAT GCT TCT GCC GGG ACG CGC 3'

Oligo reverse: 5' GGA TCC TCA GTA GAA CTG GCC CAT GT 3'

**c21orf 83-1**

Oligo sense: 5' TCA TAT GCT GCA GAA GCA CGT TCA 3'

Oligo reverse: 5' GGA TCC TCA GTA GCT GTA CAT CTG CT 3'

**c21orf 83-2**

Oligo sense: 5' TCA TAT GCA GAG GGT CGC GCG GCT 3'

Oligo reverse: 5' GGA TCC TTA ATG CCT TTC ACA ATC CAA TGG 3'

**c21orf 95**

Oligo sense: 5' TGC TAG CCAT GGA CGC TCC GAG GCT ACC 3'

Oligo reverse: 5' GGA TCC TTA TTT CCT TGC GTT TCC AG 3'

**c21orf 101**

Oligo sense: 5' TCA TAT GCC CCG CTA CGA GCT GGC 3'

Oligo reverse: 5' GGA TCC TCA CTT CTT CCT CTT CTT TG 3'

**c21orf 102**

Oligo sense: 5' TCA TAT GGG CAC CGT GCG CCC ACC 3'

Oligo reverse: 5' GGA TCC TAG GGC CCC GGG CCG ATG G 3'

As letras em vermelho, verde e azul sublinhadas são correspondentes aos sítios reconhecidos pelas enzimas de restrição *Nhe* I, *Nde* I e *Bam* HI respectivamente. Os produtos amplificados foram purificados do gel de agarose 1,5% utilizando-se o kit

Concert<sup>tm</sup> Rapid Gel Extraction System (Gibco BRL) e ligados ao vetor pBSIIKS (Stratagene). Os recombinantes foram selecionados através da análise de restrição e, em seguida, subclonados no vetor pET28a clivado com as enzimas correspondentes. Os clones positivos foram transformados em BL21(DE3) e a análise da expressão gênica foi realizada como descrito anteriormente.

### **3.12 Construção dos plasmídeos pEGFP-C1 com os genes de interesse**

Para a localização intracelular das proteínas estudadas neste trabalho, foi utilizado o plasmídeo pEGFP-C1 (plasmídeo para expressão em fusão com a “Green Fluorescent Protein”-Clontech), tratado com as enzimas de restrição *Bgl* II, preenchido com dNTPs por atividade da Klenow polimerase e clivado com *Eco* RI ou *Bam* HI. Após esses tratamentos os plasmídeos foram separados em gel de agarose 1% e posteriormente purificados utilizando-se o kit Concert<sup>tm</sup> Rapid Gel Extraction System (Gibco BRL).

Para o isolamento dos genes de interesse estes foram subclonados a partir da construção do vetor pET28a que foi clivado com a enzima *Nde* I (com exceção para a *c21orf95* que foi clivado com *Nhe* I), preenchidos com dNTPs por atividade da Klenow polimerase e posteriormente clivados com *Eco* RI ou *Bam* HI. Após esse tratamento os fragmentos com extremidades não coesiva/*Eco* RI ou não coesiva/*Bam* HI foram isolados em gel de agarose 1,5% e posteriormente purificados utilizando-se o kit descrito acima. Após a purificação os fragmentos de interesse (não coesiva/*Eco* RI ou não coesiva/*Bam* HI) e os plasmídeos pEGFP-C1 (não coesiva/*Eco* RI ou não coesiva/*Bam* HI) foram ligados. Em seguida, os produtos das ligações foram introduzidos em células competentes de *E. coli* DH5- $\alpha$  (CaCl<sub>2</sub>) e plaqueados em LB ágar contendo o antibiótico canamicina (25 mg/ml).

Para a obtenção dos plasmídeos recombinantes os transformantes foram analisados através de PCR de colônias com oligonucleotídeos que flanqueiam os respectivos genes e os clones positivos foram sequenciados e alinhados como descrito em anteriormente.

O DNA plasmidial de cada construção foi utilizado para transfectar as células de mamífero como descrito na seção 3.14.

### 3.13 Construção das formas truncadas da DCRB no plasmídeo pEGFP-C1

Para a construção das formas truncadas da DCRB foram sintetizados dois oligonucleotídeos de forma que as regiões amplificadas entrem em fase de leitura com a GFP. Assim, a partir do plasmídeo pEGFPDCRB, foram realizadas ampliações utilizando-se os seguintes oligonucleotídeos:

Oligo sense: 5' TCG AAG GTC GTC ATA TGT CGT TAA TCA TCT TGA 3'

Oligo reverse 5' AAG AAT TCA AGA AGC TGG CAG G 3'

As letras em azul e verde sublinhadas são correspondentes aos sítios das enzimas *Nde* I e *Eco* RI respectivamente.

Esta construção permitiu a amplificação da primeira parte do gene, isto é, da parte que codifica o N-terminal da DCRB. O produto amplificado foi clivado com *Nde* I, preenchido com dNTPs pela atividade da Klenow polimerase e, em seguida clivado com *Eco* RI gerando um fragmento não coesiva/*Eco* RI.

Para a segunda parte, que codifica o C-terminal da DCRB, os seguintes oligonucleotídeos foram utilizados:

Oligo sense: 5' TCT GCC AGC TTC TTC TAC GG 3' (oligonucleotídeo que flanqueia a região próxima aos 197pb)

Oligo reverse: 5' GGA ATT CAG GTT GAT GGG CTT GCG CT 3' (oligonucleotídeo que flanqueia o gene com sítio de *Eco* RI em verde sublinhado).

O fragmento amplificado foi clivado com *Eco* RI e, portanto possui extremidades não coesiva/*Eco* RI.

Assim, essas regiões foram ligadas a um plasmídeo com extremidade não coesiva/*Eco* RI gerado através da clivagem do pEGFP-C1 com *Bgl* II, preenchido com klenow e posteriormente clivado com *Eco* RI.

Os produtos das ligações foram utilizados na transformação de bactérias *E. coli* DH5- $\alpha$  CaCl<sub>2</sub> competentes e os recombinantes foram selecionados através de PCR de colônias.

Após a confirmação, o DNA plasmidial foi utilizado para transfectar as células de mamíferos como descrito em 3.14.

### **3.14 Cultura e Transfecção celular**

Células das linhagens CHO-K1 ("Chinese Hamster Ovary cells"), COS-7 ("SV40 African Green Monkey Kidney") e HEK 293 ("Human Embryonic Kidney"), foram inicialmente descongeladas e semeadas em garrafas de 75 cm<sup>2</sup>, numa densidade de aproximadamente 10<sup>4</sup> células/cm<sup>2</sup> e mantidos em meio de cultura DMEM, contendo soro fetal bovino 10%, glutamina 2mM e penicilina e estreptomicina 100U/ $\mu$ l, até atingir semiconfluência (24 horas).

Para transfecção celular foi utilizado o método de transfecção com lipossomos, utilizando lipofectamina (Gibco BRL), que permite eficiência de transferência do DNA cerca de 10 vezes maior que o fosfato de cálcio. O procedimento apresenta as seguintes etapas:

As células foram soltas utilizando-se uma solução de tripsina e semeadas em placas de 35mm (Corning-Costar), numa concentração de  $10^5$  células/placa e incubadas a  $37^{\circ}\text{C}$ , com 5% de  $\text{CO}_2$  por 24 horas, junto a uma lamínula de vidro previamente esterilizada;

Aproximadamente 5  $\mu\text{g}$  de DNA de interesse foi misturado com 100 $\mu\text{l}$  de meio sem soro (MSS). Paralelamente, cerca de 5 $\mu\text{l}$  de lipofectamina foram misturados em 100 $\mu\text{l}$  de MSS. As duas soluções descritas acima foram misturadas devagar e cuidadosamente incubadas à temperatura ambiente por 15 minutos.

O meio das placas foi removido e as células lavadas com PBS (1X). A mistura das soluções descritas acima, acrescida de 800  $\mu\text{l}$  de MSS, foi colocada nas placas, agitando-as suavemente com movimentos circulares e a seguir foram incubadas por 5 horas à  $37^{\circ}\text{C}$  em câmara úmida e 5%  $\text{CO}_2$ . Após este período, o meio das placas foi suplementado com soro (concentração final 10%) e as células foram incubadas por 24 horas na estufa de  $\text{CO}_2$ .

Decorrido este período, o meio de cultura foi removido e as células aderidas à lamínula foram lavadas em PBS (1X) e fixadas com paraformaldeído 3,7% (pH 7,3) a temperatura ambiente por 30 minutos. Após a fixação, a solução foi descartada e a lamínula foi novamente lavada com PBS (1X). A lamínula foi montada de forma invertida sobre a lâmina de vidro, colocando sobre esta 10 $\mu\text{l}$  de glicerol e vedando-se suas extremidades com esmalte comercial.

Análises preliminares foram realizadas em microscópio de epifluorescência convencional (Olympus BX50), com excitação em filtro azul a 480nm que permitiu a detecção das imagens em verde, e registros em câmara acoplada (Nikon), usando um filme de 400 ASA (Kodak Gold). 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, em colaboração com a Profa Dra Gláucia Maria Machado Santelli do Instituto de Ciências Biomédicas da USP-São Paulo.

### **3.15 Coloração com DAPI**

As células cultivadas em lamínulas e transfectadas com os plasmídeos de interesse foram permeabilizadas em uma solução de Triton X-100 a 0,2% por 5 minutos. Após este período, a solução foi aspirada e as células lavadas com PBS (1X) por 3 vezes durante 5 minutos cada. Em seguida as células foram incubadas a temperatura ambiente numa solução de DAPI (10mg/ml) diluída 1:5.000 por 5 minutos e novamente lavadas com PBS (1X) por 3 vezes durante 5 minutos cada. Posteriormente as células foram fixadas e então montadas com glicerol.

### **3.16 Análises *in silico***

Para as análises *in silico* da localização subcelular foram utilizados os seguintes programas:

**PSORT II** (<http://psort.nibb.ac.jp/form2.html>) (Nakai and Horton, 1999), que faz a predição de sinalização e sítios de localização baseado na seqüência de aminoácidos do N-terminal da proteína.

**NCBI BLAST** (<http://www.ncbi.nlm.nih.gov/BLAST/>) que foi utilizado para o alinhamento das proteínas de interesse com as proteínas depositadas no banco de dados do Centro Nacional de Informação para a Biotecnologia.

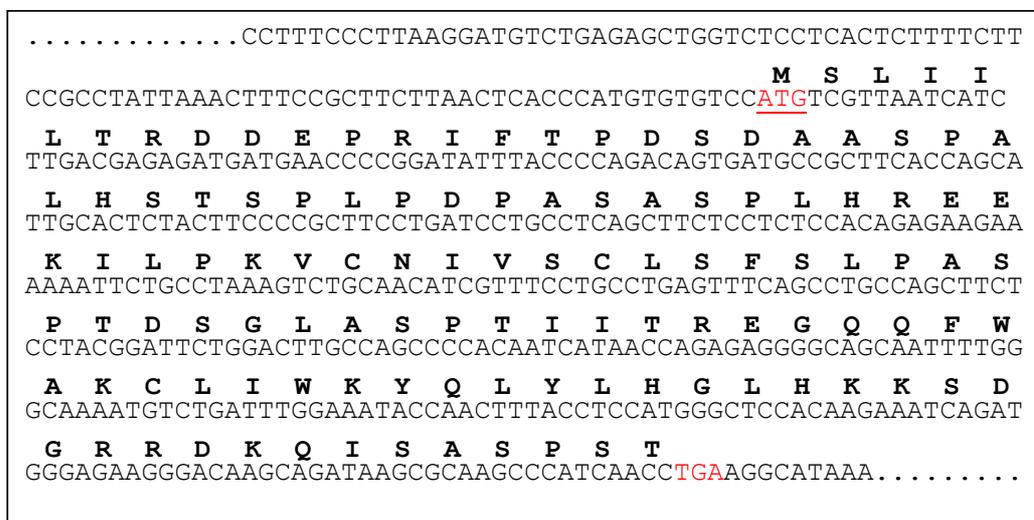
**Predict NLS** (<http://maple.bioc.columbia.edu/predictNLS/>) (Cokol et al., 2000), utilizado para a predição de sinal de localização nuclear.

**TargetP** (<http://www.cbs.dtu.dk/services/TargetP>) (Olof et al., 2000), faz a predição de localização subcelular baseado em rede neural utilizando apenas informações da seqüência N-terminal da proteína.

## 4. RESULTADOS E DISCUSSÃO

### 4.1 O gene DCRB

Nakamura et al. (1997b) publicaram a descrição de um novo gene localizado na região crítica da Síndrome de Down o qual foi denominado DCRB (Down Syndrome Critical Region gene B). Este gene consiste de três éxons com um total de 1095 pb e codifica uma ORF (“Open Reading Frame”) de 118 resíduos de aminoácidos, preferencialmente expressa em placenta. A seqüência de aminoácidos desta proteína (Figura 4) não demonstrou homologia com qualquer proteína conhecida.



**Figura 4:** Seqüência parcial de nucleotídeos do DNA do gene que codifica a DCRB. A região codificante da proteína é mostrada a partir do códon de iniciação (em vermelho sublinhado).

### 4.2 Obtenção do gene DCRB e clonagem no vetor de expressão pET28a

O DCRB completo foi amplificado em um único passo, utilizando DNA genômico humano como molde e oligonucleotídeos específicos flanqueando os éxons do gene DCRB e oligonucleotídeos que flanqueiam todo o cDNA.



Através do RT-PCR, seqüências de RNA transcritos podem ser reversamente transcritos em seqüências de DNA complementar apenas utilizando-se uma polimerase denominada transcriptase reversa. Uma variante do RT-PCR é o DDRT-PCR que permite a identificação e clonagem de genes diferencialmente expressos.

O Flip-PCR permite a inversão da seqüência de pequenas porções de DNA dentro de regiões regulatórias a fim de examinar a importância dos elementos “cis-acting” na atividade transcricional.

Outra técnica que utiliza os princípios da PCR é o chamado PCR recursivo (Prodromou e Pearl, 1992) que permite a síntese de um gene através de amplificação de oligonucleotídeos com regiões de sobreposição sem a utilização de um DNA molde.

A metodologia empregada para a amplificação do gene DCRB é uma variação da técnica de PCR recursivo (a qual não utiliza moldes) (Prodromou e Pearl, 1992) e foi denominada por nós “splicing in vitro”. Através desta técnica, o cDNA completo foi amplificado em um único passo, utilizando-se DNA genômico humano como molde e oligonucleotídeos específicos flanqueando cada éxon do DCRB com sobreposições entre eles em concentração final de 0,1 uM. Na mesma reação também foram incluídos oligonucleotídeos que flanqueiam todo o cDNA, em concentração final de 1 uM.

Em relação ao método “splicing in vitro” o PCR recursivo apresenta a desvantagem de utilizar altas concentrações de oligonucleotídeos, além do número de bases utilizado na síntese desses oligonucleotídeos assim como das regiões de sobreposição, serem maiores aumentando os custos da técnica. Além disso, o “splicing in vitro” apresenta a vantagem de não haver a necessidade de se utilizar polimerases e ciclos térmicos específicos visto que a amplificação foi precisa e resultou em banda de amplificação única.

Outro aspecto interessante deste método, é que ele gera cDNA a partir de DNA genômico dispensando o uso de bibliotecas de cDNA, RNAm e também do uso de transcriptase reversa e tecidos específicos para o isolamento de genes de interesse.

Embora tenham sido detectados alguns clones com mutações, a técnica de “splicing *in vitro*” torna-se uma boa alternativa para o isolamento de genes visto que é facilmente executada e relativamente barata.

### 4.3 Análise da seqüência do gene DCRB humano

As análises *in silico* do DCRB humano demonstraram que essa ORF está dividida exatamente em três éxons quando comparado ao DNA genômico humano com sítios aceptores e doadores de *splicing* (Figura 6).

Além disso, quando submetida a uma análise contra ESTs humanas foi observado a correspondência com várias ESTs o que é uma evidência indireta de que esse gene é expresso (Tabela 3).

#### (A)

>[gi|7768731|dbj|AP001730.1](http://gi|7768731|dbj|AP001730.1)  Homo sapiens genomic DNA, chromosome 21q, section 74/105  
Length = 340000

**EXON1:** Identities = 233/233 (100%)

```

Query: 1      cagtcacgggccagcctttcccttaaggatgtctgagagctggctcctcactcttttct 60
             |||
Sbjct: 322979 cagtcacgggccagcctttcccttaaggatgtctgagagctggctcctcactcttttct 322920

Query: 61      tccgcctattaaactttccgcttcttaactcaccatgtgtgtccatgtcgtaatacatc 120
             |||
Sbjct: 322919 tccgcctattaaactttccgcttcttaactcaccatgtgtgtccatgtcgtaatacatc 322860

Query: 121     ttgacgagagatgatgaacccggatatttaccacagacagtgatgccgcttcaccagca 180
             |||
Sbjct: 322859 ttgacgagagatgatgaacccggatatttaccacagacagtgatgccgcttcaccagca 322800

Query: 181     ttgcactctacttccccgcttctctgatcctgcctcagcttctcctctccacag 233
             |||
Sbjct: 322799 ttgcactctacttccccgcttctctgatcctgcctcagcttctcctctccacag 322747

```

**EXON2:** Identities = 104/104 (100%)

Query: 232 agagaagaaaaaattctgcctaaagtctgcaacatcgtttctgcctgagtttcagcctg 291  
 |||  
 Sbjct: 322029 agagaagaaaaaattctgcctaaagtctgcaacatcgtttctgcctgagtttcagcctg 321970

Query: 292 ccagcttctcctacggattctggacttgccagccccacaatcat 335  
 |||  
 Sbjct: 321969 ccagcttctcctacggattctggacttgccagccccacaatcat 321926

**EXON3:** Identities = 759/760 (99%)

Query: 336 aaccagagaggggcagcaattttgggcaaaatgtctgatttggaaataccaactttacct 395  
 |||  
 Sbjct: 256600 aaccagagaggggcagcaattttgggcaaaatgtctgatttggaaataccaactttacct 256541

Query: 396 ccattgggtccacaagaaatcagatgggagaaggacaagcagataagcgcaagcccatc 455  
 |||  
 Sbjct: 256540 ccattgggtccacaagaaatcagatgggagaaggacaagcagataagcgcaagcccatc 256481

Query: 456 aacctgaaggcataaacacatccagccacctccttctgatcagcagcaaagctgacgtt 515  
 |||  
 Sbjct: 256480 aacctgaaggcataaacacatccagccacctccttctgatcagcagcaaagctgacgtt 256421

Query: 516 ttgatctccatctgtctgattcttgtgtctacttctcagtttacaactccagtgggaaag 575  
 |||  
 Sbjct: 256420 ttgatctccatctgtctgattcttgtgtctacttctcagtttacaactccagtgggaaag 256361

Query: 576 aaagagctttatttacagaccataaaaaatcccatcagtgctgctcccctgctgagaggcc 635  
 |||  
 Sbjct: 256360 aaagagctttatttacagaccataaaaaatcccatcagtgctgctcccctgctgagaggcc 256301

Query: 636 atgtgagaccataatggaaaaacaacagccataatggcagcatggcagtggaagggtttgt 695  
 |||  
 Sbjct: 256300 atgtgagaccataatggaaaaacaacagccataatggcagcatggcagtggaagggtttgt 256241

Query: 696 cttgtgccaggccttgcggtcatgcaagtttcttgggatcctgttgggaccagccact 755  
 |||  
 Sbjct: 256240 cttgtgccaggccttgcggtcatgcaagtttcttgggatcctgttgggaccagccact 256181

Query: 756 caccaaggctgagtaggtccacaataatggggactttctaccagactcacagagaactg 815  
 |||  
 Sbjct: 256180 caccaaggctgagtaggtccacaataatggggactttctaccagactcacagagaactg 256121

Query: 816 ctgggttttgggaaggggtgtgcgtgtcttggggcatggaagttggggttatagtgag 875  
 |||  
 Sbjct: 256120 ctgggttttgggaaggggtgtgcgtgtcttggggcatggaagttggggttatagtgag 256061

Query: 876 acccagaggatgagaaaacttctctgcctcagcagaagagtggcagctgagagagaggca 935  
 |||  
 Sbjct: 256060 acccagaggatgagaaaacttctctgcctcagcagaagagtggcagctgagagagaggca 256001

Query: 936 agaaactcgccccactgtggactggggcagagagattttgaggagaatgaaatccagaa 995  
 |||  
 Sbjct: 256000 agaaactcgccccactgtggactggggcagagagattttgaggagaatgaaatccagaa 255941

Query: 996 actctgtgtggtattagtttgtatccagaggtgaccctcttctcaaggaaatgggtgtc 1055  
 |||  
 Sbjct: 255940 actctgtgtggtattagtttgtatccagaggtgaccctcttctcaaggaaatgggtgtc 255881

Query: 1056 atcaatcttctacactattaaagatataaagtctctggca 1095  
 |||  
 Sbjct: 255880 atcaatcttctacactattaaagatataaagtctctggca 255841

(B)

Éxon	Sítio aceptor	Sítio doador	Tamanho éxon
1	---	CTCTCCACAGgtattgatcc	233 pb
2	atcttctagAGAAGAAAA	CCACAATCATgtaagcaat	102 pb
3	ttcactcagAACCCAGAGAG	---	760 pb

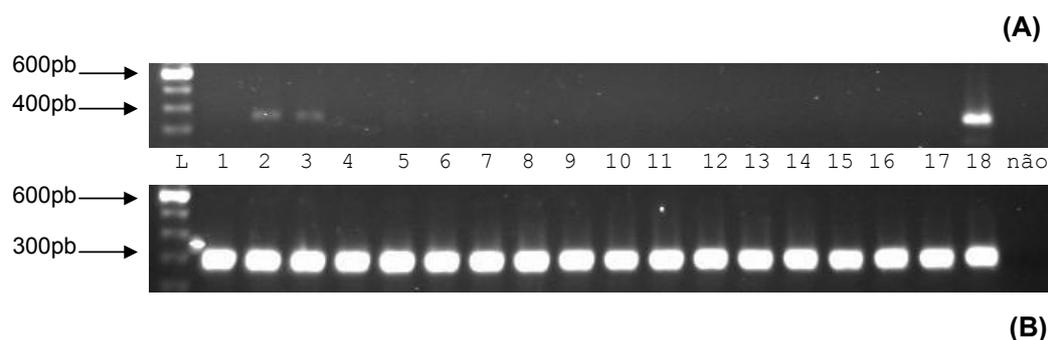
**Figura 6: Análise da seqüência do DCRB humano.** (A) Análise do alinhamento do DCRB contra a seqüência de DNA genômico humano (*Homo sapiens*). Em vermelho, os sítios iniciador e terminador respectivamente. (B) Análise dos sítios doadores e aceptores de *splicing*

**Tabela 3: Alinhamento da DCRB contra ESTs humanas.** Evidência indireta de que o gene é expresso.

Clone	Tamanho (pb)	Tecido	Identidade	Posição de alinhamento contra a DCRB
BX112350	760	cDNA de placenta <i>Homo sapiens</i>	99%	33-791
AI817192	586	“pool”	99%	517-1088
BU753721	489	Placenta	99%	626-1094
AA863381	434	“pool” tumor de célula germinativa	99%	665-10095
CB993317	818	Placenta	100%	13-335

#### 4.4 Análise da expressão da DCRB humana em diferentes tipos celulares

Foram realizadas análises da expressão da DCRB através da técnica de RT-PCR utilizando-se linhagens de células tumorais. Das 17 linhagens testadas em apenas 2 (H358 – Pulmão e A2058 – Melanoma) foi observada uma banda de amplificação correspondente ao DCRB sendo que o nível de expressão é bem menor do que o encontrado em células de placenta que foram usadas como controle positivo (Figura 7).



**Figura 7: Análise da expressão do DCRB em linhagens de células tumorais através de RT-PCR observado em gel de agarose. (A) DCRB e (B) GAPDH. L: marcador de peso molecular (100 pb), 1: H1155 – Pulmão, 2: H358 – Pulmão, 3: A2058 – Melanoma, 4: SkMel25 – Melanoma, 5: SAOS-2 – Osteosarcoma, 6: Du145 – Próstata, 7: PC3 – Próstata, 8: Hela – Útero, 9: CasKi – Útero, 10: HL60 – Leucemia, 11: SCABER – Bexiga, 12: A172 – Glioblastoma, 13: T98G – Glioblastoma, 14: MCF-7 – Mama, 15: MDA436 – Mama, 16: SW480 – Cólon, 17: FADu – Cabeça e Pescoço e 18: Placenta normal (controle positivo).**

Alguns anos depois, Toyoda et al. (2002) publicaram que o gene DCRB havia sido incorretamente predito. Os autores fizeram esta afirmativa levando-se em consideração a bem estabelecida conservação entre as regiões cromossômicas MMU16 e HSA21 de rato e humano respectivamente. Eles verificaram que o DCRB não está presente na seqüência genômica de rato, não existe identidade entre ESTs de rato e humano e através de análises computacionais não se observou a existência de genes candidatos na região da seqüência genômica de rato. Assim os autores consideram que o gene DCRB foi incorretamente predito.

Entretanto, nossas análises demonstram que o gene DCRB está organizado em três éxons bem definidos, inclusive com sítios aceptores e doadores de *splicing*, que possui similaridade com ESTs humanas depositadas em banco de dados, que é expressa em linhagens de células tumorais de pulmão e melanoma e em células COS-7 (resultados

descritos posteriormente). Esses resultados nos permite dizer que há fortes indícios a favor da existência da DCRB.

#### 4.5 Alinhamento do DCRB humano contra a seqüência genômica de *Pan troglodytes*

Recentemente a comunidade científica tomou conhecimento da seqüência genômica do chimpanzé (*Pan troglodytes*) que sabidamente apresenta grande similaridade com o genoma humano. Assim, pudemos fazer uma análise comparando a seqüência genômica da DCRB humana com o genoma do chimpanzé, no qual pode-se observar que o DCRB humano é encontrado e tem uma organização genômica idêntica em chimpanzé (Figura 8). Além disso, as ORFs da DCRB humana e chimpanzé apresentam alta similaridade com apenas duas diferenças nas posições 99 e 339 que não refletem mudanças de aminoácidos nas seqüências preditas das proteínas analisadas.

>[gi|37537393|dbj|BS000126.1](https://www.ncbi.nlm.nih.gov/nuccore/gi|37537393|dbj|BS000126.1)  Pan troglodytes chromosome 22 clone:PTB-128A24, map 22, complete sequences Length = 207747

**EXON1:** Identities = 58/58 (100%)

```
Query: 1      cagtcacgggccagcctttcccttaaggatgtctgagagctgggtctcctcactctttt 58
             |||
Sbjct: 159440 cagtcacgggccagcctttcccttaaggatgtctgagagctgggtctcctcactctttt 159383
```

Identities = 163/165 (98%)

```
Query: 69      ttaaaactttccgcttcttaactcaccatgtgtgtccatgtcggttaatcatcttgacgag 128
             |||
Sbjct: 159384 ttaaaactttccgcttcttaactcaccatgtgtgtccatgtcggttaatcatcttgacgag 159325
```

```
Query: 129     agatgatgaaccccgatatttaccocagacagtgatgccgcttcaccagcattgcactc 188
             |||
Sbjct: 159324 agatgatgaaccccgatatttaccocagacagtgatgccgcttcaccagcattgcactc 159265
```

```
Query: 189     tacttccccgcttccctgatcctgcctcagcttctcctctccacag 233
             |||
Sbjct: 159264 tacttccccgcttcccgatcctgcctcagcttctcctctccacag 159220
```

**EXON2:** Identities = 104/104 (100%)

Query: 232 agagaagaaaaaattctgcctaaagtctgcaacatcgtttctgcctgagtttcagcctg 291  
 |||  
 Sbjct: 158515 agagaagaaaaaattctgcctaaagtctgcaacatcgtttctgcctgagtttcagcctg 158456

Query: 292 ccagcttctcctacggattctggacttgccagccccacaatcat 335  
 |||  
 Sbjct: 158455 ccagcttctcctacggattctggacttgccagccccacaatcat 158412

**EXON3:** Identities = 752/760 (98%)

Query: 336 aaccagagaggggcagcaatTTTgggcaaaatgtctgatttggaaataccaactttacct 395  
 |||  
 Sbjct: 93856 aaccagagaggggcagcaatTTTgggcaaaatgtctgatttggaaataccaactttacct 93797

Query: 396 ccatgggctccacaagaaatcagatgggagaagggacaagcagataaagcgcaagccatc 455  
 |||  
 Sbjct: 93796 ccatgggctccacaagaaatcagatgggagaagggacaagcagataaagcgcaagccatc 93737

Query: 456 aaactgaaggcataaaaccacatccagccacctccttctgatcagcagcaagctgacgtt 515  
 |||  
 Sbjct: 93736 aaactgaaggcataaaaccacatccagccacctccttctgatcagcagtaagctgacgtt 93677

Query: 516 ttgatctccatctgtctgattcttgtgtctacttctcagtttacaactccagtgaggaaag 575  
 |||  
 Sbjct: 93676 ttgatctccatctgtctgattcttgtgtctacttctcagtttacaactccagtgaggaaag 93617

Query: 576 aaagagctttatttacagaccataaaaaatcccatcagtgctgctcccctgctgagaggcc 635  
 |||  
 Sbjct: 93616 aaagagctttatttacagaccataaaaaatcccatcagtgctgctcccctgctgagaggcc 93557

Query: 636 atgtgagaccatattgaaaaacaacagccataatggcagcatggcagtggaagggtttgt 695  
 |||  
 Sbjct: 93556 atgtgagaccatattgaaaaacaacagccataatggcagcatggcagtggaagggtttgt 93497

Query: 696 cttgtgccaggccttgcggctcatgcaagtttcttggatcctgttgggaccagccact 755  
 |||  
 Sbjct: 93496 cttgtgccaggccttgcggctcatgcaagtttcttggatcctgttgggaccagccact 93437

Query: 756 caccaaggctgagtaggtccacaaataatggggactttctaccagactcacagagaactg 815  
 |||  
 Sbjct: 93436 caccaaggctgagtaggtccacaaataacggggactttctaccagactcacagagaactg 93377

Query: 816 ctgggtttttgggaaggggtgctgcttcttggggcatggaagttggggttatagtgag 875  
 |||  
 Sbjct: 93376 ctgggtttttgggaaggggtgctgcttcttggggcatggaagttggggttatagtgag 93317

Query: 876 acccagaggatgagaaaaacttctctgcctcagcagaagagtggcagctgagagagaggca 935  
 |||  
 Sbjct: 93316 acccagaggatgagaaaaacttctctgcctcagcagaagagtggcagctgagagagaggca 93257

Query: 936 agaaactcgcaccactgtggactggggcagagagatTTTgaggagaatgaaatccagaa 995  
 |||  
 Sbjct: 93256 agaaactcgcaccactgtggaccgggcagagagatTTTgaggagaatgaaatccagaa 93197

Query: 996 actctgtgtggtattagtttggatccagaggggtgaccctcttctcaaggaaatgggtgtc 1055  
 |||  
 Sbjct: 93196 actctgtgtggtattagtttggatccagaggggtgaccctcttctcaaggaaatgggtgtc 93137

Query: 1056 atcaatTTTctacactattaaagatataaagttccttgga 1095  
 |||  
 Sbjct: 93136 atcaatTTTctacactattaaagatataaagttccttgga 93097

**Figura 8: Análise do alinhamento da seqüência do DCRB humano contra a seqüência genômica do chimpanzé (*Pan troglodytes*). O DCRB humano e de chimpanzé tem mesma organização genômica. Em vermelho, os sítios iniciador e terminador respectivamente.**

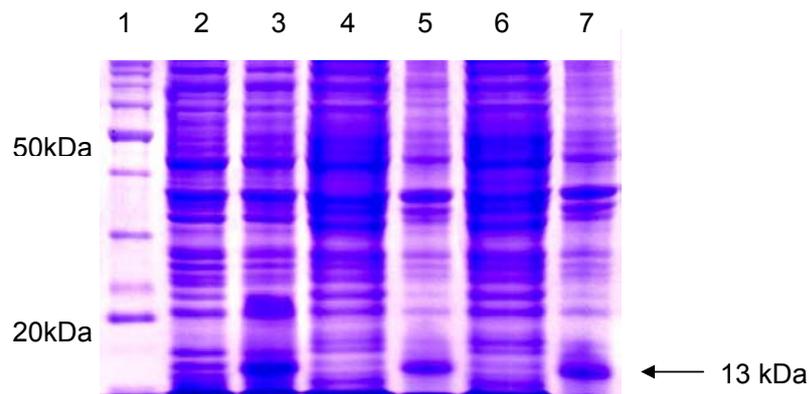
Recentemente, Takamatsu et al. (2002) usando predições computacional de genes e subsequente clonagem do cDNA, identificaram dois novos genes designados DSCR9 e DSCR10 que estão na chamada Região Crítica da Síndrome de Down (DSCR). Ambos os genes são expressos preferencialmente em testículo e codificam proteínas desconhecidas de 149 e 87 resíduos de aminoácidos, respectivamente. Outras análises demonstram que estes genes são exclusivos de genomas de primatas. Segundo os autores esses genes teriam evoluído como uma nova classe de genes em primatas evento este que também poderia ter ocorrido com o gene DCRB, reforçando as evidências de que as conclusões de Toyoda et al. (2002) estão equivocadas.

#### **4.6 Expressão da proteína DCRB em *E. coli***

##### **4.6.1 Expressão da DCRB no vetor pET28a**

O plasmídeo recombinante denominado pET28aDCRB foi usado para transformar as bactérias de expressão *E. coli* BL21(DE3) as quais foram induzidas com IPTG a uma concentração final de 0,4 mM por 4 horas. As análises de expressão protéica realizadas através de SDS-PAGE 15% (Laemmli, 1970) corado com azul de Coomassie revelam a superexpressão da proteína DCRB com um tamanho aproximado de 13kDa, conforme esperado. Além disso, os testes de solubilidade revelaram que a proteína resultante se encontrava na forma insolúvel (Figura 9).

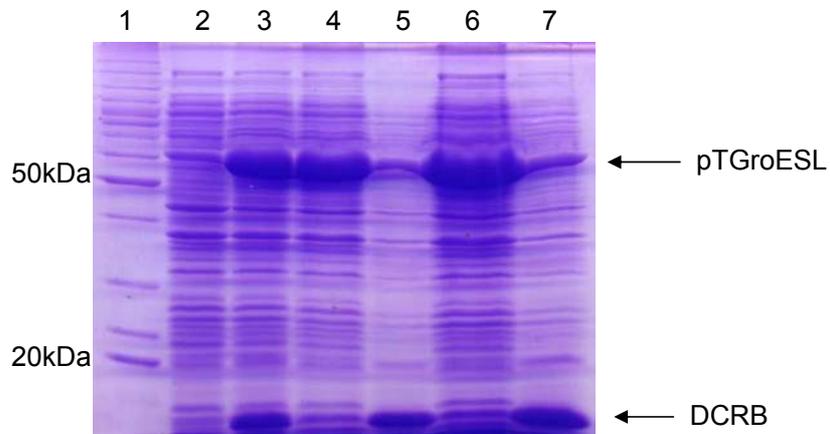
Algumas alterações como variação da concentração do indutor e a indução em temperaturas e agitação mais baixas foram realizadas, mas não houve alteração na solubilidade da proteína estudada.



**Figura 9: Análise da expressão e solubilidade da proteína DCRB em *E.coli* BL21 (DE3) utilizando o vetor pET28a.** SDS-PAGE 15% mostrando em 1-marcador de massa molecular; 2-proteínas totais bacterianas, sem indução; 3-indução com IPTG por 4 horas; 4 e 6-fração correspondente ao sobrenadante (37°C e 20°C respectivamente) e 5 e 7-fração correspondente ao precipitado (37°C e 20°C respectivamente). A seta indica a proteína expressa de aproximadamente 13kDa.

Outra estratégia muito utilizada para aumentar a solubilidade de proteínas expressas em *E. coli* é a co-expressão com chaperoninas (discutido na seção 4.15). Neste trabalho foi utilizada a co-expressão com a chaperonina GroESL e os resultados podem ser observados na figura 10.

A co-expressão da DCRB com a GroESL aumentou a solubilidade da proteína de interesse, porém em quantidades que não foram suficientes para a realização da purificação e posterior utilização em análises de estrutura secundária. Para isso é necessário que as proteínas expressas, especialmente as de função desconhecida, estejam na forma solúvel. Assim, com esse objetivo, o gene DCRB foi clonado em outros plasmídeos e sua expressão e solubilidade foram analisados.

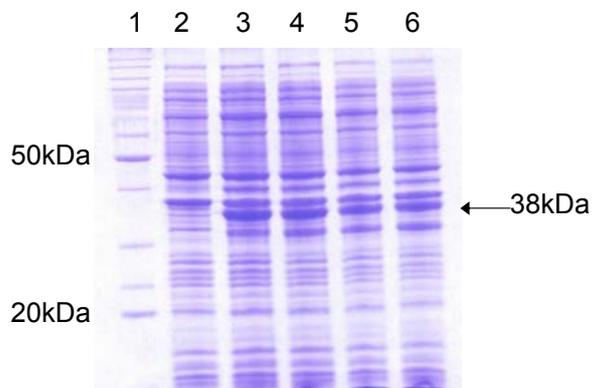


**Figura 10: Análise da co-expressão e solubilidade da proteína DCRB com a chaperonina GroESL em *E.coli* BL21 (DE3).** SDS-PAGE 15% mostrando em 1-marcador de massa molecular; 2- proteínas totais bacterianas, sem indução; 3-indução com IPTG por 4 horas; 4 e 6-fração correspondente ao sobrenadante (37°C e 20°C respectivamente) e 5 e 7- fração correspondente ao precipitado (37°C e 20°C respectivamente). As setas indicam a proteína expressa GroESL e a proteína DCRB de aproximadamente 13kDa.

#### 4.6.2 Expressão da DCRB no vetor pET37b

O vetor pET37b codifica uma seqüência sinalizadora que faz com que a proteína de fusão seja enviada para o espaço periplasmático. Esta característica permite a formação das pontes dissulfeto em um ambiente menos redutor que o citoplasma favorecendo o correto enovelamento das proteínas que contém pontes dissulfeto (Murby et al., 1996 e Jonasson et al., 2002) o que poderia ser interessante uma vez que a proteína DCRB possui três aminoácidos cisteína. Além disso, a purificação da proteína de interesse é facilitada já que a quantidade de contaminantes é bastante reduzida.

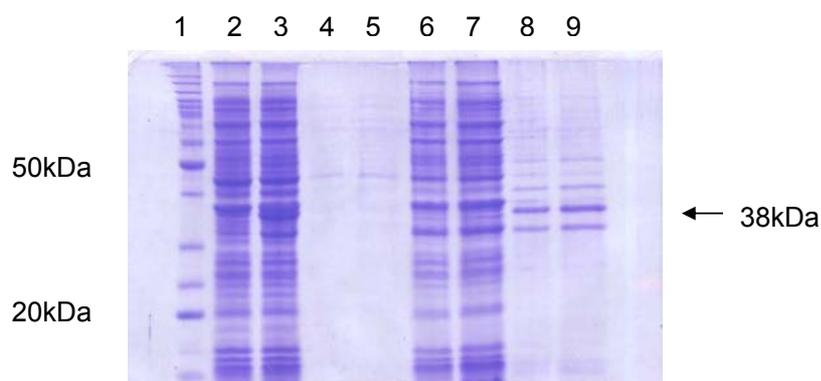
Assim, após a indução da expressão, esta foi analisada em gel de SDS-PAGE 10% corado com azul de Coomassie onde observa-se a expressão da proteína de fusão CBD-DCRB de aproximadamente 38kDa (Figura 11).



**Figura 11: Análise da expressão da proteína DCRB em *E.coli* B121 (DE3) utilizando o vetor pET37b.** (1) marcador de massa molecular; (2) proteínas totais bacterianas, sem indução; (3) indução após 1 hora; (4) indução após 2 horas; (5) indução após 3 horas; (6) indução após 4 horas.

A análise do teste de solubilidade via choque osmótico demonstrou que a proteína DCRB se encontra insolúvel após expressão em fusão com a CBD (Figura 12)

Assim, apesar de suas características vantajosas quanto formação das pontes dissulfeto e o correto enovelamento das proteínas que contém essas pontes (Murby et al., 1996 e Jonasson et al., 2002) e a facilitação dos processos de purificação, este vetor de expressão não foi útil para a expressão da DCRB, já que na forma solúvel inviabiliza a purificação da proteína em condições nativas. Os mesmos testes foram realizados com a indução feita a 20°C e os mesmos resultados foram obtidos.

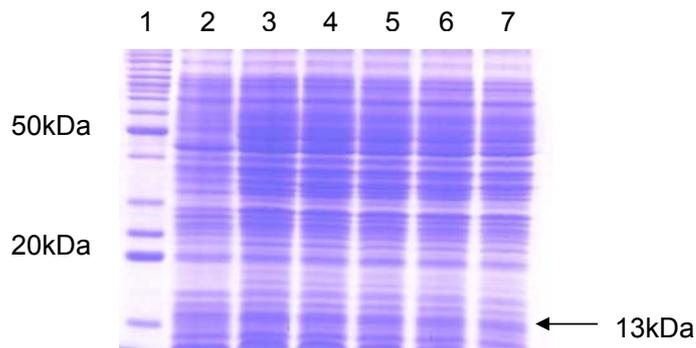


**Figura 12: Análise do teste de solubilidade da proteína pET37bDCRB expressa em *E. coli*. BL21 (DE3).** (1) marcador de massa molecular; (2) proteínas totais bacterianas, sem indução; (3) indução após 4 horas; (4 e 5) fração periplasmática; (6 e 7) fração solúvel e (8 e 9) fração insolúvel.

#### 4.6.3 Expressão da DCRB no vetor pBADTOPO

O sistema com o promotor araBAD proporciona uma expressão variável e é dose dependente do indutor (Stevens, 2000). Estas características poderiam contribuir na expressão da DCRB em condições solúveis.

Após a indução da expressão em células *E. coli* LMG 194, alíquotas de 1 ml foram retiradas e analisadas em gel SDS-PAGE 15% corado com azul de Coomassie e os resultados demonstrados na figura 13 indicam que não houve expressão da DCRB que pudesse ser detectada através desse método.



**Figura 13: Análise dos testes de indução da expressão da proteína DCRB clonada no vetor pBad.** (1) marcador de massa molecular; (2) proteínas totais bacterianas, sem indução; (3) indução com 0,00002% de L-arabinose; (4) indução com 0,0002% de L-arabinose; (5) indução com 0,002% de L-arabinose; (6) indução com 0,02% de L-arabinose; (7) indução com 0,2% de L-arabinose.

A análise do sequenciamento do vetor pBADDCRB demonstrou a integridade do gene, isto é, ausência de mutação e a junção vetor-inserto em fase de leitura o que descarta a possibilidade de algum erro ter impossibilitado a expressão gênica.

Sabe-se que outros fatores podem influenciar a expressão de proteínas recombinantes. Entre eles estão a toxicidade da proteína para a célula hospedeira, a instabilidade da proteína, a perda do plasmídeo durante o crescimento da bactéria, o processamento impróprio ou modificações pós-traducionais e tradução ineficiente (QIAexpressionist, 2003).

Outros fatores que podem comprometer a produção de proteínas é a rápida degradação do RNAm devido a sua instabilidade e a presença de códons que são raros em bactérias e que frequentemente estão presentes em genes heterólogos (Hannig e Makrides, 1998) (ver seção 4.15).

#### 4.6.4 Expressão da DCRB no vetor pCX-TOPO

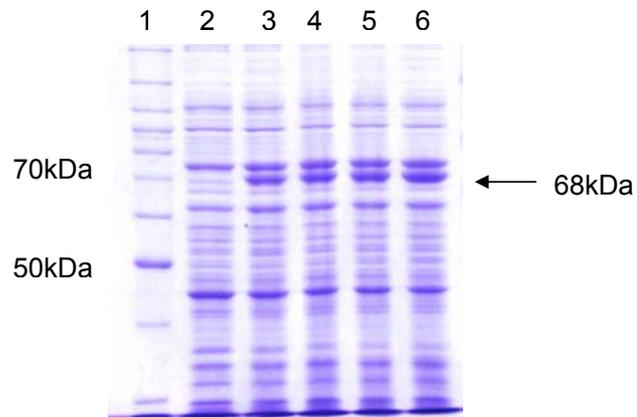
Após a transferência das colônias para o meio M11 as células de *Caulobacter crescentus* (B5 BAC) foram incubadas durante 5 dias, período sob o qual não foi observado nenhum agregado protéico e portanto, concluiu-se que não houve expressão da proteína DCRB em fusão com a proteína RsaA, visto que foi realizado um controle de expressão e neste caso observou-se o aparecimento de agregados protéicos de coloração rósea.

Após o sequenciamento foi descartada qualquer possibilidade de erro que permitisse a ausência de expressão da proteína em questão. Conclui-se então, que tal fato pode ter ocorrido por outros fatores, entre eles o fato da existência de códons raros que a maquinaria enzimática da bactéria (*Caulobacter*) não poderia traduzir ou a toxicidade DCRB para este sistema contribuindo para sua degradação prematura e impossibilitando a sua detecção a olho nu.

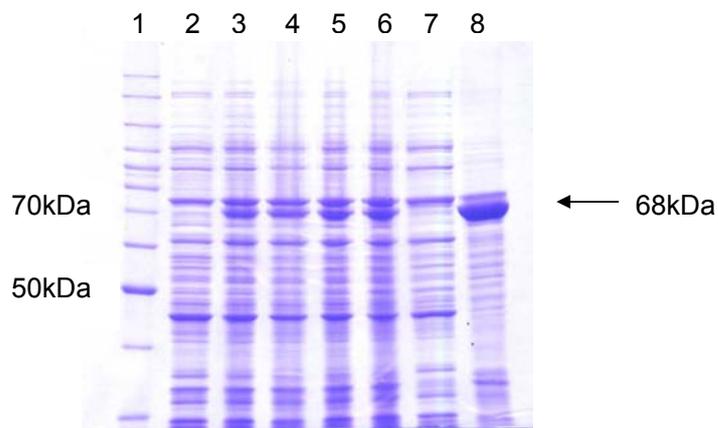
#### 4.6.5 Expressão da DCRB no vetor pTYB2

A expressão da proteína DCRB em *E. coli* ER2566 foi analisada em SDS-PAGE 10% corado com azul de Coomassie e os resultados são apresentados na figura 14 na qual pode-se verificar que após 1 hora de indução ocorreu a expressão da proteína DCRB em fusão com a inteína com um tamanho esperado de aproximadamente 68 kDa.

Assim, prosseguiu-se as análises realizando um teste de solubilidade onde amostras do sobrenadante e precipitado foram analisados em gel SDS-PAGE 10% corados com azul de Coomassie. Os resultados demonstram que a proteína DCRB em fusão com a inteína também se encontra na forma insolúvel (Figura 15). Outros testes foram realizados com a indução feita a 20°C e os mesmos resultados foram obtidos.



**Figura 14: Análise do teste de expressão da proteína DCRB no plasmídeo pTYB2.** (1) marcador de massa molecular; (2) proteínas totais bacterianas, sem indução; (3) indução após 1 hora; (4) indução após 2 horas; (5) indução após 3 horas; (6) indução após 4 horas.



**Figura 15: Análise do teste de solubilidade da proteína DCRB em fusão com a inteína.** (1) marcador de massa molecular; (2) proteínas totais bacterianas, sem indução; (3) indução após 1 hora; (4) indução após 2 horas; (5) indução após 3 horas; (6) indução após 4 horas; (7) sobrenadante e (8) precipitado.

Pelo fato da proteína estar insolúvel, para se realizar estudos estruturais foi necessário recuperar a proteína ativa expressa em corpos de inclusão. Assim, a proteína que é insolúvel tem que ser solubilizada com agentes desnaturantes que posteriormente são retirados para que ocorra a renaturação da proteína estudada. Apesar deste tratamento parecer bastante simples, ele não garante que a proteína renaturada terá sua conformação nativa, o que dificulta análises estruturais (Yasukawa et al., 1995).

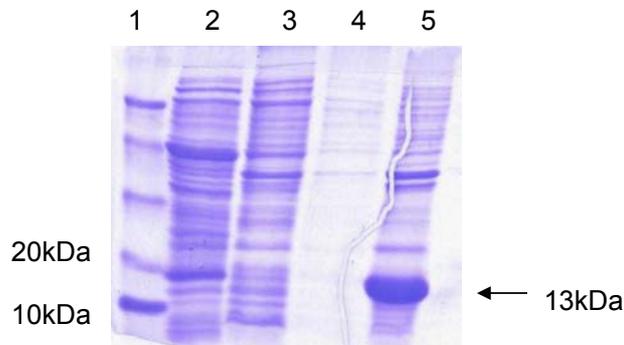
#### **4.7 Purificação da proteína DCRB em condições desnaturantes**

A fim de se obter proteína solúvel, análises de expressão e solubilidade foram realizadas utilizando-se diferentes vetores, hospedeiros, temperatura de indução da expressão e quantidade de indutor, porém em todos eles a proteína estudada estava na forma insolúvel. Assim, optou-se por fazer uma purificação em condições desnaturantes na qual, apesar da proteína não apresentar atividade biológica evidente, a purificação é simples e realizada em um único passo.

Assim, após a indução da proteína DCRB no vetor pET28a durante 4 horas, as células foram centrifugadas (10.000 rpm, 4°C por 10 minutos) e ressuspensas em tampão PBS (1X) contendo 1% de Triton X-100, um detergente que solubiliza a maioria das proteínas, com exceção dos corpos de inclusão nos quais estão localizados a DCRB (Figura 16).

As proteínas expressas em *E. coli* normalmente formam agregados protéicos insolúveis denominados corpos de inclusão. Embora esta seja uma das barreiras para estudos estruturais (Yasukawa et al., 1995) a purificação de proteínas recombinantes a partir de corpos de inclusão é relativamente fácil já que 80 a 90% destes são formados pela proteína de interesse ([www.hamptonresearch.com](http://www.hamptonresearch.com)). Porém, as condições necessárias para o

renovelamento são proteína-dependente e requerem várias tentativas já que nem sempre, após a retirada do desnaturante, a proteína alvo adquire sua conformação nativa (Tsumoto et al., 2003.)



**Figura 16: Caracterização dos corpos de inclusão produzidos através da expressão da DCRB no vetor pET28a.** (1) marcador de massa molecular; (2) precipitado após sonicação; (3) precipitado após primeira diluição; (4) precipitado após segunda diluição; (5) precipitado (corpos de inclusão).

Após a confirmação que a proteína era expressa como corpos de inclusão, um protocolo de renovelamento protéico foi aplicado e foram realizadas medidas de CD conforme descrito anteriormente, a fim de monitorar a formação de estruturas secundárias.

O espectro de CD da proteína DCRB “renovelada” mostrou um mínimo um pouco mais pronunciado ao redor de 206 nm e um mínimo discreto em torno de 222 nm, sugerindo a presença (ou aquisição) de estrutura secundária do tipo  $\alpha$ -hélice. Esse resultado demonstra que o renovelamento protéico foi realizado de maneira satisfatória (veja seção 4.11).

#### 4.8 Mutação sítio dirigida do gene DCRB

Em resultados iniciais da expressão e solubilidade da DCRB pode-se observar uma forte banda de expressão de aproximadamente 13 kDa, conforme o esperado, e que grande parte da proteína estava na forma solúvel, entretanto, após o sequenciamento verificou-se a presença de uma mutação espontânea o que inviabilizava a produção da proteína de interesse

Após a obtenção do gene DCRB de forma íntegra, este foi clonado em vários plasmídeos de expressão e a DCRB submetida aos testes de solubilidade (como descrito anteriormente), nos quais se verificou que a proteína expressa estava na forma de corpos de inclusão.

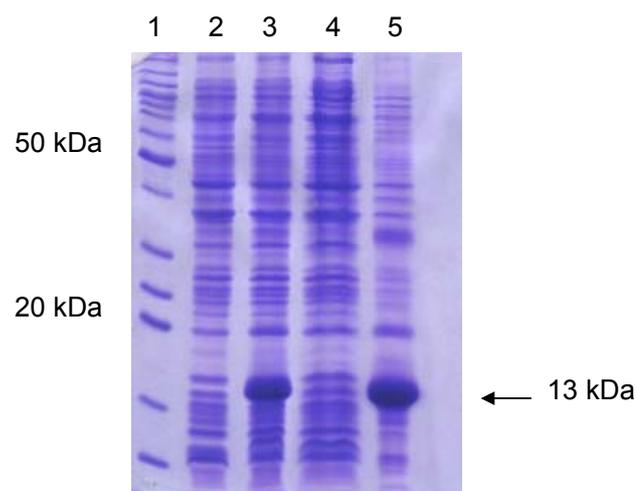
Analisando a seqüência de aminoácidos da proteína DCRB íntegra pode-se observar a existência de três cisteínas na sua composição (Figura 17). Esta seqüência foi comparada a seqüência mutante e verificou-se que a proteína resultante continha apenas duas cisteínas, sendo a última substituída por outro aminoácido. A insolubilidade da proteína DCRB íntegra poderia ser consequência da formação de pontes dissulfeto entre as cadeias polipeptídicas.

1	<u>MSLIILTRDD EPRIFTPDSD AASPALHST SPLPDPASAS REEKILPKVC</u>
51	<u>NIVSCLSSLP ASTTDSGLASPTIITREGQQFWAK</u> * CLIWKYQLYLHGLHKK
101	SDGRRDKQIS ASPST

**Figura 17: Seqüência de aminoácidos codificados pela DCRB.** Em vermelho os três aminoácidos cisteína presentes na seqüência total. A região sublinhada corresponde à proteína denominada DCRB 2/3. Na DCRB mutada a última cisteína (\*) é substituída pelo aminoácido valina (C85V).

Assim, considerando a possibilidade de que a cisteína C85 da DCRB pudesse estar contribuindo para sua insolubilidade através da formação de ligações dissulfeto, resolvemos mutar o códon respectivo por um códon correspondente ao aminoácido valina e verificar se este aminoácido realmente estaria contribuindo para a insolubilidade da proteína estudada.

Após a realização da mutação sítio dirigida, 20 clones foram analisados e sequenciados, verificando-se que 8 deles eram mutantes (Figura 19). Em seguida procedeu-se a análise da expressão gênica e da solubilidade da proteína expressa na qual pode-se verificar uma banda de expressão da proteína de interesse com aproximadamente 13 kDa como o esperado. Porém, após o teste de solubilidade, observou-se que a proteína expressa estava na forma insolúvel (Figura 18) indicando que a última cisteína não é a responsável pela insolubilidade da DCRB.



**Figura 18: Análise da expressão e solubilidade da proteína DCRB mutada em *E.coli* (B121-DE3) em SDS-PAGE 15%.** (1) marcador de peso molecular; (2) proteínas totais bacterianas, sem indução; (3) indução após 4 horas; (4) sobrenadante e (5) precipitado.

```

111 120 130 140 150 160 170 180 190 200 210 220
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
dcrbmut1 GCAGGAAATTCGATTCGAGAGGTCGTCATATGTCGTTAATCATCTTGACGAGAGATGATGAAACCCGGATATTTACCCAGACAGTGTGCCGCTTCACACGATTGCAC
dcrbcv   ATGTCGTTAATCATCTTGACGAGAGATGATGAAACCCGGATATTTACCCAGACAGTGTGCCGCTTCACACGATTGCAC
Consensus .....ATGTCGTTAATCATCTTGACGAGAGATGATGAAACCCGGATATTTACCCAGACAGTGTGCCGCTTCACACGATTGCAC

221 230 240 250 260 270 280 290 300 310 320 330
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
dcrbmut1 CTACTTCCCGCTTCTGATCTGCTCAGCTTCTCCTCCACAGAGAGAGAAAATTTGCTAAGTCTGCACATCGTTTCTGCTGAGTTTCAGCTGCCAGCT
dcrbcv   CTACTTCCCGCTTCTGATCTGCTCAGCTTCTCCTCCACAGAGAGAGAAAATTTGCTAAGTCTGCACATCGTTTCTGCTGAGTTTCAGCTGCCAGCT
Consensus CTACTTCCCGCTTCTGATCTGCTCAGCTTCTCCTCCACAGAGAGAGAAAATTTGCTAAGTCTGCACATCGTTTCTGCTGAGTTTCAGCTGCCAGCT

331 340 350 360 370 380 390 400 410 420 430 440
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
dcrbmut1 TCTCTACGGATTCTGGACTTGGCAGCCCCACATCATARCCAGAGAGGGGCGAGCAATTTGGCCAAAGTTCTGATTTGGAAATACCACTTTACCTCCATGGGCTCCA
dcrbcv   TCTCTACGGATTCTGGACTTGGCAGCCCCACATCATARCCAGAGAGGGGCGAGCAATTTGGCCAAAGTTCTGATTTGGAAATACCACTTTACCTCCATGGGCTCCA
Consensus TCTCTACGGATTCTGGACTTGGCAGCCCCACATCATARCCAGAGAGGGGCGAGCAATTTGGCCAAAGTTCTGATTTGGAAATACCACTTTACCTCCATGGGCTCCA

441 450 460 470 480 490 500 510 520 530 540 550
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
dcrbmut1 CAGGAATCAGATGGGAGAGGGACAGCAGATAGCGCAGGCCATCACCCTGAATTCATCAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCCGGTACCAGCT
dcrbcv   CAGGAATCAGATGGGAGAGGGACAGCAGATAGCGCAGGCCATCACCCTGA
Consensus CAGGAATCAGATGGGAGAGGGACAGCAGATAGCGCAGGCCATCACCCTGA.....

```

**Figura 19: Análise do sequenciamento de um clone mutante do DCRB.**

Em destaque (azul) pode-se observar a substituição da base desejada, isto é, TG por GT.

#### 4.9 Deleção de parte da proteína DCRB

Além da mutação sítio dirigida, realizamos também a clonagem da região N-terminal da proteína DCRB que é comum a todos os clones e que possivelmente esteja relacionada à solubilidade já que a proteína expressa inicialmente (com mutação) estava na forma solúvel.

Analisando a seqüência de aminoácidos que codifica a proteína DCRB observa-se que há uma equivalência entre a proporção de aminoácidos hidrofóbicos (49%) e hidrofílicos (51%) que se mantém quando é deletada a região C-terminal correspondente a 32 aminoácidos sendo 11 hidrofóbicos, 18 hidrofílicos e 3 não carregados.

Após a clonagem, testes de expressão da proteína denominada por nós de DCRB 2/3 (Figura 17) foram realizados e pode-se observar em gel SDS-PAGE 15% uma banda de expressão com um tamanho aproximado de 10 kDa como esperado e, após os testes de

solubilidade, verificou-se que parte desta proteína está na forma solúvel quando a expressão é induzida a 20°C.

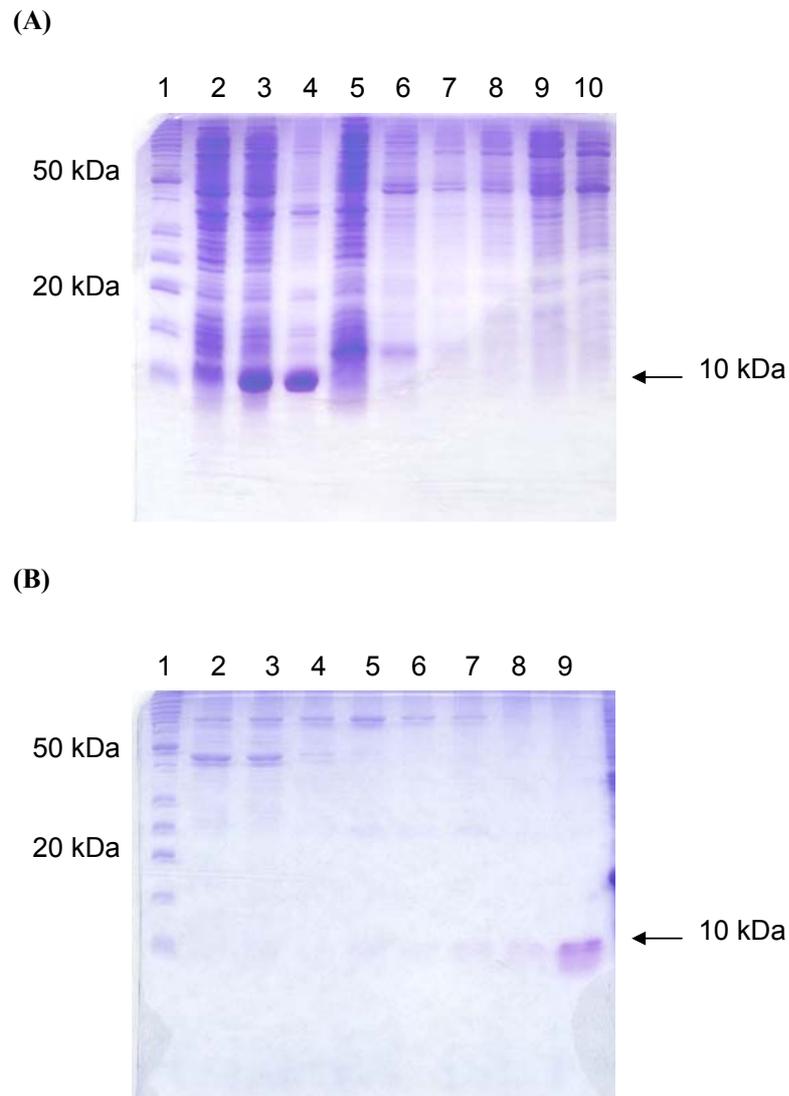
Em seguida foram realizados testes de purificação da DCRB 2/3 em resina de níquel e pode-se observar que a proteína de interesse foi purificada e eluída em tampão de sonicação com 250 mM de imidazol (Figura 20A e 20B).

Assim, após a diálise, a amostra da DCRB 2/3 purificada foi submetida à análise de dicroísmo circular (CD) e o espectro mostra estrutura secundária organizada como pode ser visto na figura 24 (Seção 4.11).

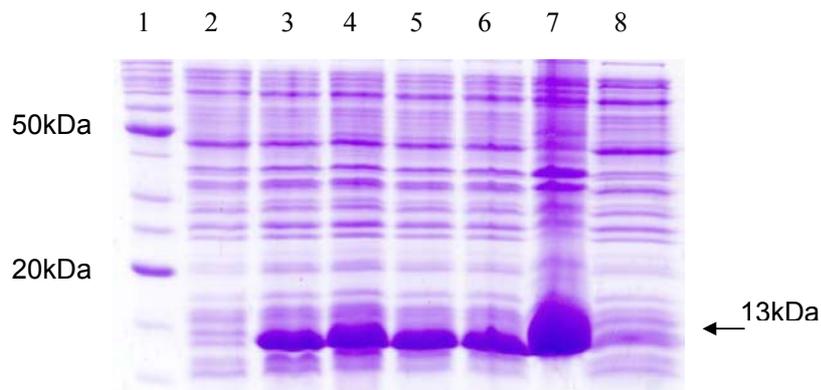
#### **4.10 Expressão da DCRB utilizando-se a cepa Rosetta (DE3)**

Após várias tentativas de expressão da DCRB em condições solúveis, foi gentilmente disponibilizada pela Profª Dra Ana Paula Ullian de Araújo (USP-São Carlos) a cepa denominada Rosetta (DE3) (Novagen) uma linhagem derivada que permite baixos níveis de expressão em todas as células possibilitando o aumento de solubilidade e da estabilidade da proteína de interesse. Permite a expressão de proteínas que contém códons AGG, AGA, AUA, CUA, CCC e GGA, raros em *E. coli*.

Assim uma nova tentativa de expressão foi realizada utilizando-se o plasmídeo recombinante denominado pET28aDCRB que foi introduzido em bactérias de expressão Rosetta (DE3) e induzidas com IPTG a uma concentração final de 0,4 mM por 4 horas. As análises realizadas através de SDS-PAGE 15% (Laemmli, 1970) corado com azul de Coomassie revelam a superexpressão da proteína DCRB com um tamanho aproximado de 13kDa conforme esperado. Além disso, os testes de solubilidade revelam que uma pequena porção da proteína resultante se encontrava na forma solúvel (Figura 21).



**Figura 20: Análise da expressão da proteína DCRB 2/3 em *E.coli* (BI21-DE3). SDS-PAGE 15% mostrando teste de expressão e purificação da proteína. (A): (1) marcador de massa molecular; (2) proteínas totais bacterianas, sem indução; (3) indução após 4 horas; (4) eluato; (5-10) início das eluições. (B) (1) marcador de peso molecular; (2-9) frações das eluições de 50 a 250 mM de imidazol.**



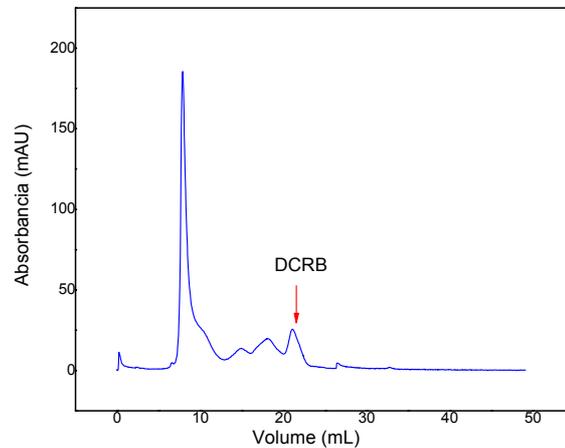
**Figura 21: Análise do teste de expressão e solubilidade da proteína DCRB na bactéria Rosetta (DE3).** (1) marcador de massa molecular; (2) proteínas totais bacterianas, sem indução ; (3) indução após 1 hora; (4) indução após 2 horas; (5) indução após 3 horas; (6) indução após 4 horas; (7) precipitado e (8) sobrenadante.

Após a obtenção de uma pequena fração da proteína DCRB na forma solúvel, o sobrenadante foi submetido a uma cromatografia de afinidade utilizando-se a resina Ni-NTA Superflow (Qiagen) como descrito em 3.8.1 e a proteína DCRB foi totalmente eluída na concentração de 250mM de imidazol, porém com contaminantes (Figura 23A).

Segundo o manual do fabricante da resina, algumas alterações na composição do tampão de lavagem da coluna como, adição de 10 a 20mM de imidazol, aumento da concentração de sal acima de 200 mM, o acréscimo de detergente e etanol/glicerol podem evitar a ligação de proteínas inespecíficas a resina. Todas essas alterações foram realizadas de acordo com as sugestões recomendadas, entretanto não trouxeram modificação no padrão de purificação.

Assim, tornou-se necessário mais um passo cromatográfico para a purificação da proteína estudada. Após a obtenção da fração solúvel da DCRB proveniente de três litros de cultura e a passagem do sobrenadante na coluna de afinidade, as frações contendo a

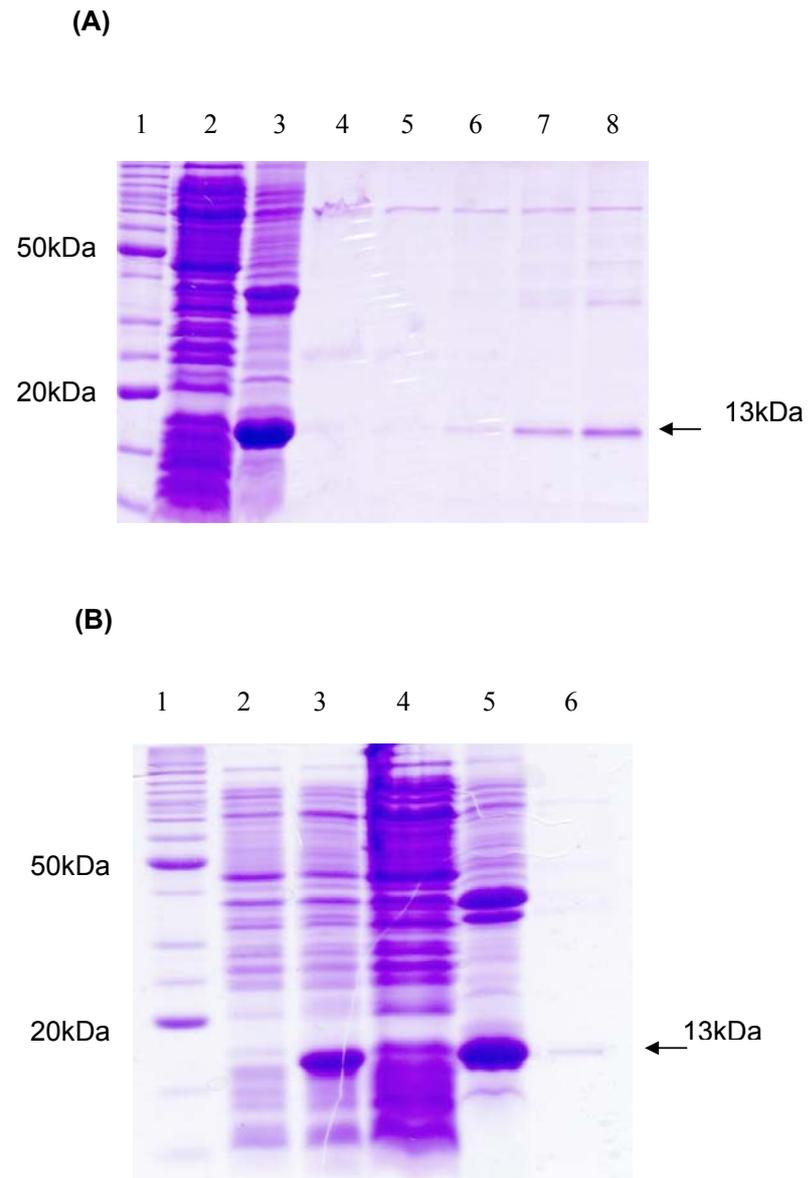
proteína DCRB e seus contaminantes, isto é, as eluídas em uma concentração de 250 mM de imidazol foram dialisadas e concentradas até 500  $\mu$ l e então aplicadas em uma coluna Superose 12 de exclusão molecular como descrito em 3.7.2 (Figura 22)



**Figura 22. Perfil cromatográfico da eluição da DCRB na coluna Superose12.** A amostra da fração mais pura da cromatografia de afinidade foi concentrada em Centricron 10 e aplicada na coluna de exclusão molecular.

A amostra correspondente ao pico de eluição foi concentrada até 0,2 mg/ml e submetida à análise de CD na qual, apesar da pequena quantidade de proteína, pode-se observar a formação de bandas indicando a presença de estruturas organizadas (Figura 24).

Esta amostra foi novamente concentrada e aplicada em um gel de poliacrilamida no qual foi possível observar a proteína DCRB totalmente purificada, porém em quantidade muito pequena (Figura 23B).



**Figura 23: Análise da purificação da proteína DCRB.** (A): Cromatografia de afinidade onde 1- marcador de massa molecular; 2-sobrenadante; 3-precipitado; 4-8: eluição nas concentrações de 75 mM (2), 100 mM (1), 100 mM(2), 250 (1) mM e 250 mM(2) de imidazol. (B): Cromatografia de exclusão molecular onde 1-marcador de peso molecular; 2- amostra não induzida, 3-após 4 horas de indução, 4-sobrenadante, 5-precipitado e 6-amostra purificada.

#### 4.11 Análise comparativa dos espectros de CD da proteína DCRB

O dicroísmo circular (CD) é uma técnica utilizada para estimar a porcentagem dos componentes da estrutura secundária das proteínas que são as hélices  $\alpha$ , folhas  $\beta$ , voltas  $\beta$  e seqüências randômicas.

A espectrometria de CD baseia-se no desvio da luz circularmente polarizada incidente em compostos quirais (ou assimétricos). Esta espectroscopia também é conhecida como quiróptica, pois o fenômeno só ocorre em compostos que têm esta propriedade. Em proteínas os grupos assimétricos envolvem as ligações peptídicas, as cadeias laterais dos aminoácidos e grupos prostéticos (Nakanish et al., 1994).

A luz circularmente polarizada desempenha um papel decisivo na espectroscopia de dicroísmo circular. Um feixe de luz circularmente polarizado consiste de dois feixes de onda plana, linearmente polarizados, ortogonais entre si e fora de fase por  $90^\circ$ . A onda polarizada linearmente pode ser decomposta em duas componentes individuais circularmente polarizadas uma à esquerda e outra à direita.

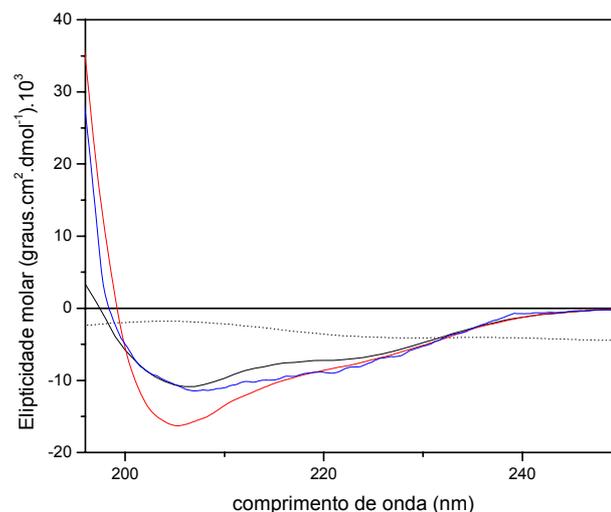
Quando esta luz passa por uma substância assimétrica (amostra) ela é absorvida, e cada componente (circular à direita e circular à esquerda) interage diferentemente devido a absorvidade intrínseca dos diferentes compostos. Assim, o dicroísmo circular é definido como a diferença entre a absorção das componentes polarizadas à esquerda e à direita.

Nos resultados apresentados na figura 24 podemos observar o espectro da proteína DCRB desnaturada (pontilhado), obtido a partir das proteínas solubilizadas dos corpos de inclusão onde nenhuma banda referente à estrutura secundária organizada é observada. Ainda nesta figura, pode-se observar que o espectro renovado (preto) adquiriu dois mínimos, embora suaves, característicos de estrutura secundária ordenada, em torno de 206 e 217 nm, podendo conter hélices alfa (Woody et al., 1994).

Quando expressa na forma solúvel (azul), o espectro resultante possui forma e posições de mínimos muito similares ao espectro da proteína renovada. Este resultado sugere que o processo de renovamento utilizado foi efetivo para solubilizar e adquirir proteínas estruturadas a partir dos corpos de inclusão.

Em vermelho vê-se o espectro da DCRB 2/3. Neste caso, verifica-se uma mudança conformacional no espectro de CD, traduzida pelo aumento no mínimo próximo de 200nm, sugerindo um aumento de estruturas randômicas (Woody et al., 1994).

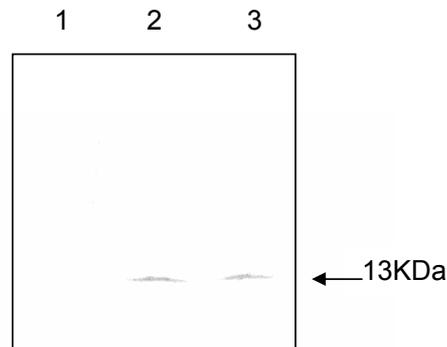
Novas análises devem ser realizadas a fim de se obter maiores quantidades de proteína purificada para que se possa realizar análises de estrutura secundária de melhor qualidade e a análise da estrutura tridimensional que são muito importantes para a definição da função das proteínas e suas possíveis utilizações.



**Figura 24: Espectros de CD.** Proteína DCRB obtida na forma solúvel (azul), quando desnaturada para posterior renovamento (pontilhado), renovada (preta) e deletada, DCRB 2/3 (vermelha). As medidas foram realizadas em um espectropolarímetro J715 (Jasco Inc.), em cubetas cilíndricas de quartzo com 1 mm de caminho ótico. Cada espectro é resultante de em média 16 varreduras.

#### 4.12 Produção e titulação de anticorpos policlonais

Conforme descrito em material e métodos, camundongos machos foram imunizados com a proteína DCRB isolada de gel de poliacrilamida e após 60 dias os soros desses camundongos foram testados através de “Western Blot” no qual pode-se verificar o reconhecimento da proteína DCRB superexpressa (Figura 25).



**Figura 25: Análise do reconhecimento da DCRB pelo antissoro contra esta proteína.** (1) controle negativo; (2 e 3) reconhecimento da proteína DCRB pelo anticorpo produzido.

A titulação desse anticorpo foi realizada através do teste de ELISA e a leitura foi realizada em espectrofotômetro a 490 nm e os resultados são demonstrados na Tabela 4.

Esses resultados indicam que o título do anticorpo produzido contra a DCRB é de 1:64.000 o que é bastante elevado.

A produção desse anticorpo foi muito importante visto que ele pode ser utilizado em estudos posteriores para a monitoração da expressão da proteína DCRB em células de mamíferos e ainda poderá ser utilizado em futuros experimentos como ensaios de imunoprecipitação e de localização subcelular utilizando-se o fracionamento celular.

**Tabela 4: Valores correspondentes à leitura da placa de titulação**

Amostras	Concentração	ABS (490 nm)
Controle	1:50	0
C1	1:50	0.095
C2	1:100	0.096
C3	1:500	0.091
C4	1:1000	0.085
C5	1:2000	0.077
C6	1:4000	0.059
C7	1:8000	0.055
C8	1:16000	0.032
C9	1:32000	0.009
C10	1:64000	0.002
C11	1:96000	-0.012
C12	1:128000	-0.013
C13	1:192000	-0.017
C14	1:256000	-0.017
C15	1:512000	-0.027

Os resultados da titulação também podem ser observados através da intensidade da cor em cada poço da placa (Figura 26).



**Figura 26: Visualização da coloração observada no teste de titulação do anticorpo contra a DCRB.** No poço C10 observa-se a diluição 1:64.000.

#### 4.13 Expressão da DCRB em células de mamíferos

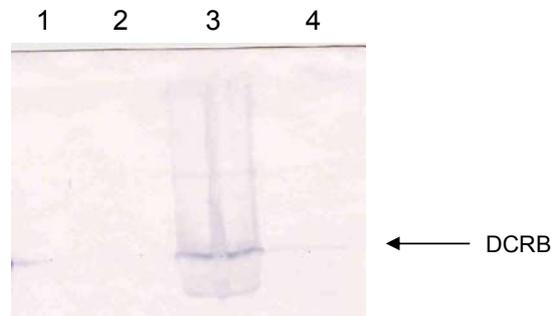
Células de mamíferos são frequentemente os melhores hospedeiros para a expressão de proteínas recombinantes de vertebrados. Isso se deve ao fato do sistema de expressão em células de mamíferos produzir as mesmas modificações pós-traducionais e reconhecer os mesmos sinais para a síntese, processamento e secreção utilizados nos organismos do qual a seqüência foi originalmente derivada (QIAexpressionist, 2003).

Existem dois tipos de expressão em células de mamíferos: o primeiro é o chamado transiente, no qual ocorre a expressão da proteína de interesse de 1 a 4 dias após a transfecção e o segundo é a expressão estável no qual ocorre a integração do gene de interesse ao DNA cromossômico.

Existe uma grande variedade de vetores disponíveis comercialmente para a realização da expressão de proteínas em células de mamíferos que possuem um promotor forte para o início da transcrição, sinais para o processamento do RNAm como seqüências de clivagem e poliadenilação, marcadores para seleção em células de mamíferos quando a expressão é estável e também uma seqüência que permite a propagação em células bacterianas (QIAexpressionist, 2003).

Devido às dificuldades de expressar em *E. coli* a DCRB na forma solúvel, fizemos uma tentativa de expressão desta proteína em células de mamíferos denominadas COS-7 (“SV African Green Monkey Kidney”). Após a transfecção com o plasmídeo pcDNA3DCRB e seleção do clone recombinante, as células COS-7 foram mantidas em meio seletivo até que tivesse uma quantidade suficiente para que pudessemos analisar a expressão e solubilidade da proteína DCRB. Após a separação do sobrenadante e do precipitado, o sobrenadante foi dialisado durante a noite e então submetido a uma cromatografia de afinidade utilizando uma resina de níquel com descrito em 3.8.1.

A DCRB expressa em células COS-7 foi submetida à confirmação pela análise de “western blot” utilizando-se o anticorpo contra a DCRB produzido neste trabalho (como descrito na seção 4.12). Os resultados podem ser observados na figura 27.



**Figura 27:** Análise por “western blot” da expressão e purificação da DCRB em células de mamíferos COS-7. (1) células COS-7 não transfectadas, (2) sobrenadante, (3) precipitado e (4) eluato.

Através dessas análises verifica-se que a DCRB é expressa em células COS-7 não transfectadas (Figura 27). Este resultado torna-se especialmente importante, pois confirma a existência da proteína DCRB demonstrando mais uma vez que as conclusões de Toyoda et al. (2002) estão equivocadas.

Uma pequena fração na forma solúvel da proteína DCRB foi detectada nas células COS-7 transfectadas, entretanto a maioria da proteína expressa ainda permaneceu no precipitado (Figura 27). Segundo Geisse et al. (1996) o sucesso do sistema eucarioto varia dependendo da natureza da proteína a ser expressa e a sensibilidade desta à degradação proteolítica. O segundo evento pode ter ocorrido com a DCRB que, após ser degradada foi precipitada o que fez com que fosse detectada em maior parte na fração insolúvel.

Apesar da detecção da proteína DCRB na forma solúvel, a quantidade observada é muito pequena e, quando submetida à cromatografia de afinidade, parte da proteína de interesse não se ligou a resina e parte foi diluída e não detectada pelo anticorpo.

Um grande número de células transfectadas é necessário para a produção da proteína recombinante seja feita em quantidade significativa o que torna este método bastante demorado e caro.

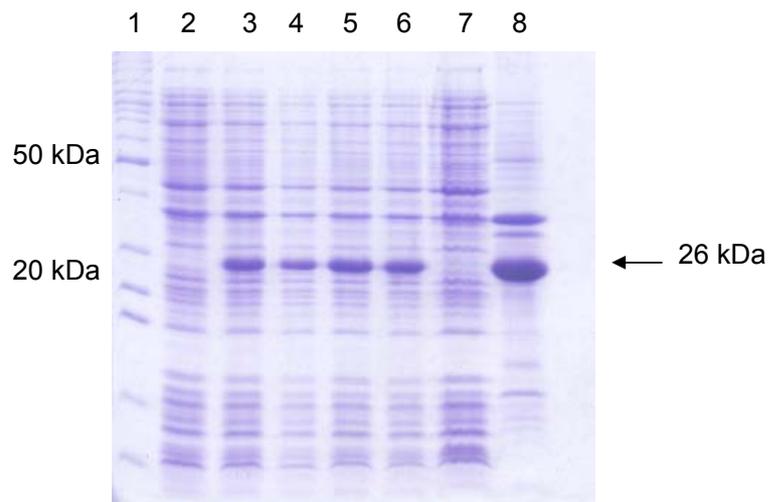
#### **4.14 Amplificação, clonagem e expressão de cinco ORFs do cromossomo 21**

Das nove ORFs do cromossomo 21 localizadas na região crítica da Síndrome de Down selecionadas e submetidas a reação de PCR utilizando-se como molde uma biblioteca de cérebro fetal humano e um painel de cDNA de vários tecidos humanos, em cinco pôde-se observar, através da análise em gel de agarose 1-1,5%, uma banda de amplificação com os respectivos tamanhos esperados, sendo elas as c21orf 45 (701 pb), c21orf 59 (872 pb), c21orf 83-2 (536 pb), c21orf 95 (464 pb) e c21orf 101 (377 pb). Nas demais ORFs (c21orf18; c21orf63; c21orf83-1 e c21orf102) não foram observadas bandas de amplificação.

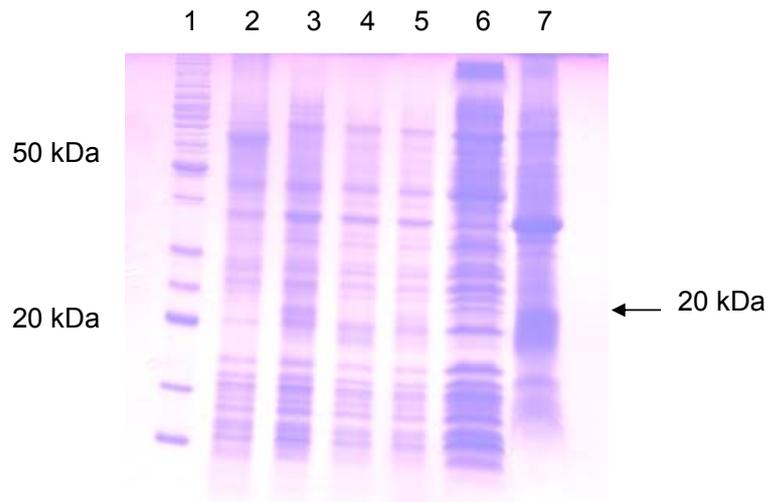
Após a amplificação as ORFs foram clonadas (pBSIIKS) e subclonadas (pET28a) como descrito anteriormente. Em seguida foram realizados testes de expressão em células de *E. coli*, nos quais se verificou bandas de expressão significativa das proteínas codificadas pelas ORFs c21orf 45 e c21orf 83-2 com aproximadamente 26 e 20 kDa respectivamente (Figura 28 e 29). As demais ORFs denominadas c21orf59, c21orf95 e c21orf101 não revelaram nenhuma banda de expressão protéica.

A ausência de bandas de expressão pode ser decorrente de toxicidade ou instabilidade da proteína, do processamento inadequado, tradução ineficiente ou então à necessidade de modificações pós-traducionais (veja seção 4.15) (QIAexpressionist, 2003).

De posse das proteínas expressas, realizou-se então os testes de solubilidade no qual se verificou que tanto a c21orf 45 quanto a c21orf 83-2 estão insolúveis. (Figura 28 e 29 respectivamente). Uma pequena quantidade de proteína solúvel foi detectada quando a c21orf45 e a c21orf59 foram expressas na cepa Rosetta (DE3) (dados não mostrados).



**Figura 28: Análise da expressão da ORF45 em *E.coli* BL21(DE3). SDS-Page 15% mostrando teste de expressão e solubilidade da proteína ORF 45. (1) marcador de massa molecular; (2) proteínas totais bacterianas, sem indução; (3) indução após 1 hora; (4) indução após 2 horas; (5) após 3 horas; (6) indução após 4 horas (7) sobrenadante e (8) precipitado**



**Figura 29: Análise da expressão da ORF 83-2 em *E.coli* BL21(DE3). SDS-PAGE 15% mostrando teste de expressão e solubilidade da proteína ORF 83-2. (1) marcador de massa molecular; (2) proteínas totais bacterianas, sem indução; (3) indução após 1 hora; (4) indução após 2 horas; (5) após 3 horas; (6) sobrenadante e (7) precipitado.**

#### **4.15 Expressão e solubilidade das proteínas heterólogas em *E. coli***

Atualmente existem vários sistemas de expressão de proteínas recombinantes como em bactérias, leveduras, insetos, plantas e células de mamíferos. O sistema bacteriano continua sendo o mais utilizado principalmente devido aos elevados níveis de expressão, baixo custo e a possibilidade de resultados positivos em um curto período de tempo, porém não possui a maquinaria enzimática para realizar as modificações pós-traducionais o que pode inviabilizar a expressão de algumas proteínas.

A escolha do sistema de expressão mais adequado depende de alguns fatores como a necessidade de grandes quantidades de proteína, a necessidade de processamento pós-traducional, a facilidade de purificação e a toxicidade da proteína alvo para a célula

hospedeira, entretanto esta tarefa tem sido bastante facilitada já que hoje existe uma variedade muito grande de vetores disponíveis comercialmente.

A produção de proteínas heterólogas depende da clonagem do gene, da indução da expressão e da purificação da proteína de interesse. A clonagem de genes é um procedimento bastante simples e atualmente realizado com bastante sucesso. Porém, a expressão desses genes depende de vários fatores e deve ser otimizada de acordo com a proteína estudada, já que cada uma delas possui propriedades físicas e químicas particulares (Braun et al., 2002).

Em alguns casos, não há a expressão da proteína de interesse como o que ocorreu no presente trabalho para a DCRB quando foram usados os plasmídeos pBADTOPO e pCXTOPO e para as ORFs c21orf59, c21orf95 e c21orf101 quando foi utilizado o plasmídeo pET28a.

A primeira análise a ser realizada é o sequenciamento dos vetores para a verificação da integridade dos genes, isto é, ausência de mutação e a junção vetor-inserto em fase de leitura o que poderia impossibilitar a expressão gênica. Essa hipótese foi descartada por nós visto que através do sequenciamento verificou-se o sucesso das clonagens.

Segundo Hannig e Makrides (1998) outros fatores podem comprometer a produção de proteínas heterólogas entre eles a rápida degradação do RNAm devido a sua instabilidade e a presença de códons que são raros em bactérias e que frequentemente estão presentes em genes heterólogos. Além disso, a toxicidade da proteína para a célula hospedeira, a instabilidade da proteína, a perda do plasmídeo durante o crescimento da bactéria, o processamento impróprio ou modificações pós-traducionais e tradução ineficiente são fatores que também devem ser considerados (QIAexpressionist, 2003).

Entre os fatores descritos acima, os mais comuns são a toxicidade e/ou a instabilidade da proteína e a perda do plasmídeo durante o crescimento da bactéria. A toxicidade da proteína para o hospedeiro pode ser avaliada através da comparação entre a eficiência de transformação das bactérias pelos plasmídeos controle e que possuem o gene de interesse. Desse modo, a eficiência de transformação observada utilizando-se o plasmídeo controle é bem maior do que com o plasmídeo recombinante o que não foi observado em nossos experimentos. Assim, a ausência de expressão das proteínas estudadas neste trabalho se deu por outros motivos que segundo Wang et al. (2003) podem não ser óbvios.

A ausência de expressão da DCRB quando utilizados os vetores pBADTOPO e pCXTOPO também pode ser justificada respectivamente pela presença de um promotor fraco e pela utilização de outro hospedeiro (*Caulobacter*) que não *E. coli*.

Porém, a expressão da DCRB foi realizada com sucesso quando esta foi clonada nos vetores pET28a, pET37b e pTYB2. O mesmo foi observado para as ORFs c21orf45 e c21orf83-2 quando clonadas no pET28a porém, quando submetidas ao teste de solubilidade, todas elas apresentaram-se insolúveis.

Segundo Sati et al. (2002) a solubilidade de proteínas não ligantes de membrana é um complexo fenômeno bioquímico, e geralmente é considerado que estas proteínas quando enoveladas adequadamente são razoavelmente solúveis em soluções aquosas.

Entre as proteínas estudadas neste trabalho, nenhuma delas possui predição de localização em membrana (veja seção 4.16), entretanto quando expressas em sistema procarioto, em todas elas foram observados agregados protéicos insolúveis denominados corpos de inclusão. Isto ocorre devido ao mau enovelamento das proteínas superexpressas.

O sistema de expressão utilizando o vetor da linha pET (Novagen), especialmente o pET28a, foi escolhido devido a algumas características tais como: fusão com seqüência

codificadora da His-Tag que facilita a purificação da proteína expressa e de sítios de clivagem para protease como a trombina. Além disso, para estudos iniciais de estrutura da proteína, não é necessário a retirada da pequena cauda de histidinas que interfere pouco nas análises preliminares de estrutura e função de proteínas, mas pode interferir na atividade da proteína se ela for uma enzima (Araújo et al., 2000).

Cada proteína é expressa em uma condição específica sendo que esta deve ser descoberta de modo empírico. Assim, tentou-se a expressão da proteína DCRB na forma solúvel em vários vetores, cada um com uma característica particular.

Dentre eles o pET37b que possui uma seqüência que direciona a proteína alvo para o espaço periplasmático; que seria um ambiente que favorece o correto enovelamento de proteínas que contém resíduos de cisteína como a DCRB que possui três aminoácidos deste tipo (Murby et al., 1996 e Jonasson et al., 2002) ou o vetor pCX-TOPO que possui a proteína de fusão RsaA que é secretada para o meio de cultura em forma de agregados protéicos e facilmente purificada através de filtração. Já o vetor pBADTOPO possui um promotor fraco e conseqüentemente uma expressão fraca que poderia interferir na solubilidade da proteína em questão (Stevens, 2000).

Outra estratégia utilizada foi a expressão da DCRB utilizando-se o vetor pTYB2 que possui a proteína de fusão denominada inteína que além de solúvel possui afinidade por CBD e o poder de autoclivagem na presença de determinados compostos o que facilita muito os passos de purificação da proteína de interesse. Segundo Davis et al. (1999), proteínas que são altamente expressas e em condições solúveis podem ser usadas como parceiras na expressão de proteínas inicialmente insolúveis proporcionando o aumento da solubilidade da proteína alvo.

Porém, em nossos estudos as proteínas superexpressas estão na forma insolúvel em todas as condições testadas, o que é muito comum acontecer em proteínas heterólogas

expressas em procariotos. Estimativas realizadas para genes humanos expressos em *E. coli* indicam que apenas de 15-20% das proteínas são solúveis, 20-40% formam corpos de inclusão e o restante não tem expressão significativa ou são degradados (Fernandez e Hoeffler, 1998, Stevens, 2000).

Considerando que a obtenção de proteína na forma nativa é o primeiro passo para estudos estruturais, várias estratégias tem sido utilizadas para que as proteínas recombinantes sejam expressas na forma solúvel.

Uma das alternativas para expressar proteínas solúveis é a fusão da proteína de interesse com outras proteínas já expressas em *E. coli*. As mais usadas são a GST (*Schistosoma japonicum* glutathione S-transferase) (Smith e Johnson, 1988, in Davis et al., 1999), MBP (*E. coli* maltose-binding protein) (di Guan et al., 1988, in Davis et al., 1999) e Tiorredoxina (*E. coli* thioredoxin) (LaVallie et al., 1993, in Davis et al., 1999). A GST e MBP foram escolhidas pelos autores porque combinam altos níveis de expressão com facilidade de purificação enquanto a tiorredoxina aumenta o nível de expressão de proteínas na forma solúvel.

Davis et al. (1999), usando um modelo estatístico de solubilidade de proteínas em *E. coli* selecionaram três proteínas, NusA, GrpE e BFR e verificaram que a produção de proteínas insolúveis em fusão com a NusA pode ser vantajosa, não só pelas características de solubilização da NusA devido a sua composição de aminoácidos, como seu alto nível de expressão.

Segundo Hammarström et al. (2002) apesar da utilização de proteínas parceiras ser uma estratégia reconhecidamente eficiente no aumento da solubilidade de proteínas heterólogas ainda não há um estudo comparativo dos efeitos das diferentes proteínas de fusão o que contribui para que a parceira mais adequada seja encontrada de modo empírico.

Outra alternativa utilizada para a produção de proteínas na forma solúvel é a co-expressão com proteínas que tornem a proteína quimérica solúvel. Neste sentido, Yasukawa et al. (1995) examinaram o efeito da co-expressão de tiorredoxina (Trx) de *E. coli* e chaperonina GroESL de *E. coli* em oito proteínas de vertebrados e concluíram que a solubilidade de todas as oito foi aumentada em co-expressão com a tiorredoxina (Trx) enquanto que quando co-expressas com a chaperonina GroESL aumentou a solubilidade de apenas quatro das oito proteínas estudadas. As proteínas Trx e GroESL estão relacionadas ao enovelamento correto e consequentemente na solubilidade das proteínas.

Considerando as vantagens mencionadas anteriormente e as dificuldades de aumentar a solubilidade de uma das proteínas estudadas neste trabalho, a DCRB foi co-expressa com a GroESL e os testes de solubilidade revelam que uma pequena parte da proteína resultante se encontra na forma solúvel, entretanto em quantidades insuficientes para a purificação da DCRB nativa, tornando inviável a produção da proteína de interesse através da co-expressão com esta chaperonina.

Outros fatores também podem influenciar a solubilidade das proteínas como os vetores e hospedeiros utilizados, a concentração do indutor, a temperatura em que a indução é realizada e a velocidade de crescimento da bactéria. Essas variáveis foram testadas em nosso trabalho, porém não houve alteração em relação à solubilidade das proteínas estudadas.

Uma pequena alteração na solubilidade da DCRB foi observada quando esta proteína foi expressa na cepa denominada Rosetta (DE3) (Novagen) uma linhagem derivada que permite baixos níveis de expressão em todas as células possibilitando o aumento de solubilidade e da estabilidade da proteína de interesse. Além disso, permite a expressão de proteínas que contém códons AGG, AGA, AUA, CUA, CCC e GGA, raramente codificados em *E. coli*.

Outra característica que influencia na solubilidade das proteínas é a sua composição de aminoácidos. Segundo o modelo estatístico de Wilkinson-Harrison, cinco parâmetros baseados nos aminoácidos foram relacionados com a formação de corpos de inclusão. Entretanto, tem sido discutido que apenas dois parâmetros são críticos na distinção de expressão de proteínas solúveis ou insolúveis. Esses parâmetros seriam a média aproximada de cargas que leva em conta as diferenças no número de resíduos de ácido aspártico mais ácido glutâmico versus lisina mais arginina comparado ao número total de resíduos de asparagina, glicina, prolina e serina (Davis et al., 1999).

Análises da predição da solubilidade, que leva em consideração a relação entre os aminoácidos presentes na seqüência de proteínas expressas em *E. coli*, revelaram que a DCRB tem uma probabilidade de 89% de ser insolúvel, predição que foi confirmada em todas as análises realizadas neste trabalho. A mesma predição foi realizada para a proteína DCRB 2/3 e como resultado obtivemos que a probabilidade desta proteína ser solúvel é de 68% o que foi verificado experimentalmente.

Por outro lado, quando realizamos a mutação sítio dirigida substituindo uma cisteína por uma valina o que diminui a probabilidade de formação de pontes de sulfeto de maneira errônea e conseqüentemente a precipitação da proteína expressa, verificou-se que a proteína continua na forma de corpos de inclusão.

Assim, podemos sugerir que o fator determinante para a insolubilidade da proteína DCRB expressa em *E. coli* é a composição de aminoácidos da proteína estudada. Estas mesmas análises foram realizadas para as orfs expressas neste trabalho e indicam que tanto a ORF45 como a ORF83-2 são insolúveis com uma probabilidade de 96% e 92 % respectivamente.

Essas observações demonstram que a composição de aminoácidos das proteínas estudadas é de fundamental importância para o sucesso da expressão e solubilidade em *E. coli*.

#### **4.16 Localização subcelular das proteínas estudadas neste trabalho**

Nos últimos 10 anos, a GFP (Green Fluorescent Protein) tem sido usada com sucesso como marcador da expressão gênica e processos intracelulares em organismos que vão desde bactérias a ratos transgênicos. Entre as aplicações que tem obtido mais sucesso está o uso da GFP para monitorar a dinâmica da localização de proteínas e o destino de proteínas fusionadas na vida celular e em organismos como um todo (Sambrook e Russell, 2001).

A localização celular é uma ferramenta muito importante para inferir na função das proteínas e isso se torna especialmente importante para proteínas como a DCRB e as codificadas pelas ORFs aqui estudadas visto que não há muitas informações a respeito dessas proteínas.

Assim, foram realizadas predições de localização dessas proteínas utilizando um programa de computador denominado Predict NLS (<http://maple.bioc.columbia.edu/predictNLS/>) que verifica a ocorrência de sinais de localização nuclear (NLS), que enviam a proteína sintetizada no citoplasma para o núcleo via NPC (complexo do poro nuclear) ou outras seqüências de sinalização para organelas e também os programas PSORT II (<http://psort.nibb.ac.jp/form2.html>) e TargetP (<http://www.cbs.dtu.dk/services/TargetP/>) que fazem a predição de localização celular de proteínas.

As proteínas estudadas no presente trabalho variam de 10 a 32 KDa sendo que todas possuem equivalentes proporções de aminoácidos hidrofóbicos e hidrofílicos exceto as orfs c21orf83-2 e c21orf95.

Os algoritmos de predição de localização celular de proteínas utilizam basicamente os seguintes métodos: (1) busca de sinais de endereçamento para locais específicos da célula como o NLS, (2) comparação da seqüência de aminoácidos da proteína alvo com outras proteínas de localização previamente conhecida e (3) a relação existente entre a composição total de aminoácidos da proteína e a localização subcelular (Andrade et al., 1998).

Segundo Cedano et al. (1997), proteínas nucleares são geralmente pobres em resíduos de aminoácidos hidrofóbicos e ricas em resíduos com carga. Além disso, elas possuem grande quantidade de serina, treonina, prolina, asparagina e glutamina. Já, nas proteínas citoplasmáticas existe um balanço entre resíduos de aminoácidos básicos e ácidos (Andrade et al., 1998).

Quanto à composição de aminoácidos as proteínas DCRB, c21orf45, c21orf59 e c21orf101 possuem um equilíbrio em relação à quantidade de aminoácidos hidrofóbicos e hidrofílicos. Porém, nas proteínas c21orf83-2 e c21orf95 a maioria dos aminoácidos são hidrofóbicos, 60% e 62% respectivamente (Tabela 5).

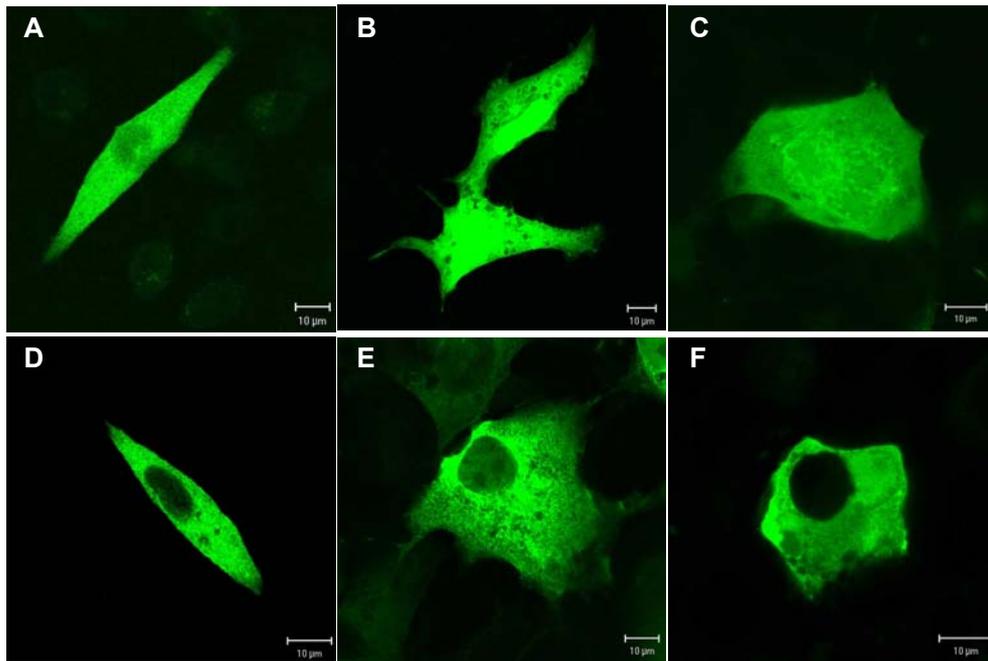
A similaridade entre proporções de aminoácidos hidrofóbicos e hidrofílicos é uma característica que aumenta a possibilidade de erro na predição de localização celular (Cedano et al., 1997).

De acordo com PSORT II, DCRB tem uma probabilidade de 52% de ser nuclear, c21orf45 tem igual probabilidade de ser nuclear ou citoplasmática (39%) e c21orf59 possui 61% de probabilidade de ser citoplasmática. Já as orfs c21orf95 e c21orf101 possuem 61% e 35% de probabilidade de ser nuclear, respectivamente. Entre as proteínas estudadas,

somente a c21orf83-2 apresentou maior probabilidade de ter a sua localização em uma organela específica do compartimento celular com uma probabilidade de 39% de ser mitocondrial. Embora, quatro das proteínas analisadas tenham sua predição de localização nuclear, apenas para a c21orf101 foi predito um sinal do tipo KKRRKK no c-terminal quando essas seqüências foram analisadas pelo PredictNLS.

Analisando a seqüência de aminoácidos da proteína DCRB observa-se uma predição de localização nuclear com 52% de probabilidade, porém não foi previsto um NLS clássico. Dados experimentais demonstram que a proteína GFP-DCRB foi distribuída no citoplasma em todas as células testadas (Figura 30). A detecção da DCRB preferencialmente no citoplasma nos levou a procurar um sinal de exportação nuclear (NES) ou um mecanismo de retenção citoplasmática. Análises da composição de aminoácidos da DCRB indica um possível sinal de NES (xxxLxxxLxx) localizado entre os aminoácidos 55 e 65. Entretanto, a localização subcelular das formas truncadas da DCRB mostraram o mesmo padrão de distribuição GFP-DCRB, sugerindo que estes 9 aminoácidos não atuam como um NES funcional (Figura 31). Entretanto, o alinhamento desta seqüência usando o NCBI BLAST revelou similaridade com uma região da Drebrina (número de acesso 1610233A), que é uma proteína citoplasmática relacionada a neurodegeneração e presente em portadores da Síndrome de Down e Mal de Alzheimer (Shim and Lubec, 2002).

Em anexo



**Figura 30: Análise da localização intracelular da proteína DCRB.** Imagens obtidas por microscopia confocal evidenciando a expressão de pEGFP-C1 (A, B, C) e pEGFP-DCRB (D, E, F) nas linhagens CHO-1 (A e D); COS-7 (B e E) e HEK293 (C e F).

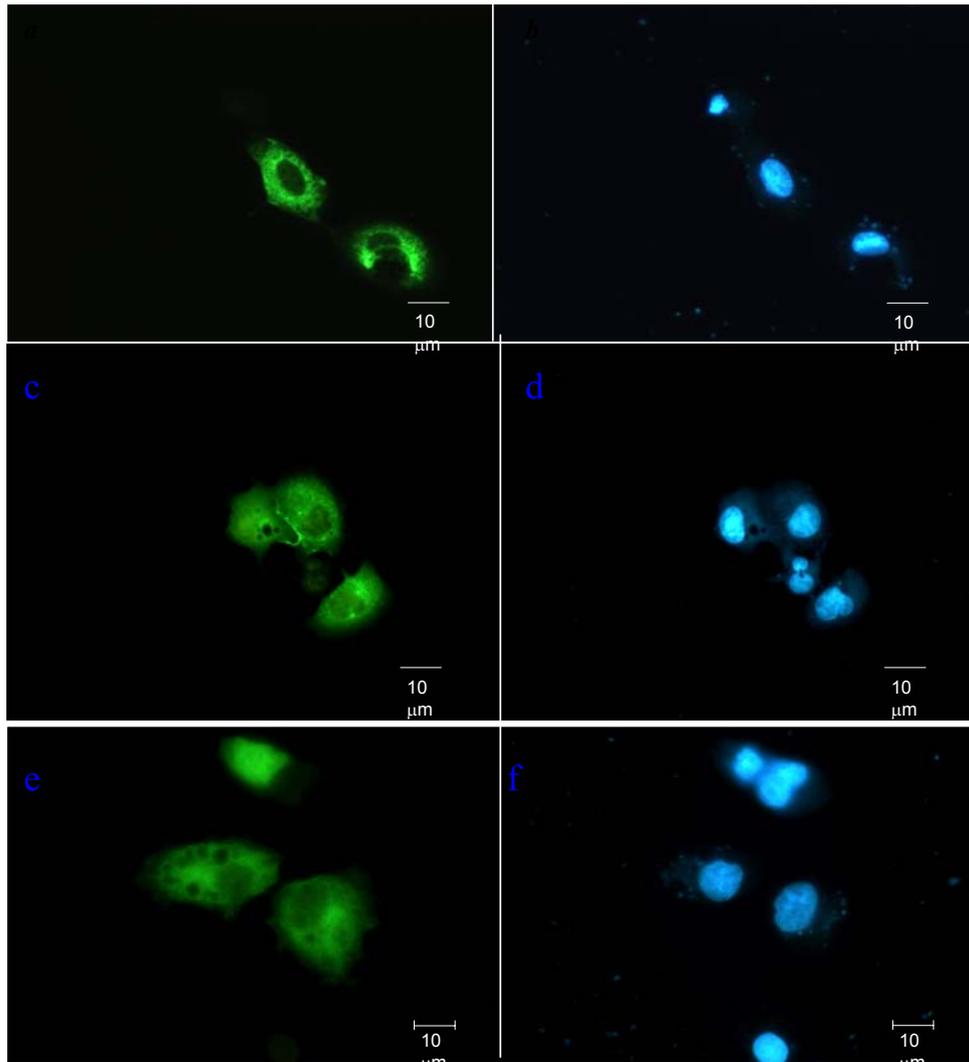
Células CHO-1, COS-7 e HEK 293 foram transfectadas com o plasmídeo pEGFP-C1, onde pode-se observar que a GFP se distribuiu uniformemente por toda a célula (Figura 30A,B, C e 32A).

A observação da localização da proteína GFP-ORF45 no citoplasma de todas as linhagens de células analisadas confirmam os resultados apresentados pelos programas de predição de sinal de localização nuclear (PredictNLS) e análise de predição de localização subcelular de proteínas (PSORT II) (Figura 32). A seqüência de aminoácidos da c21orf45 apresenta 96% de similaridade com a proteína FAPP1. Esta proteína possui um domínio

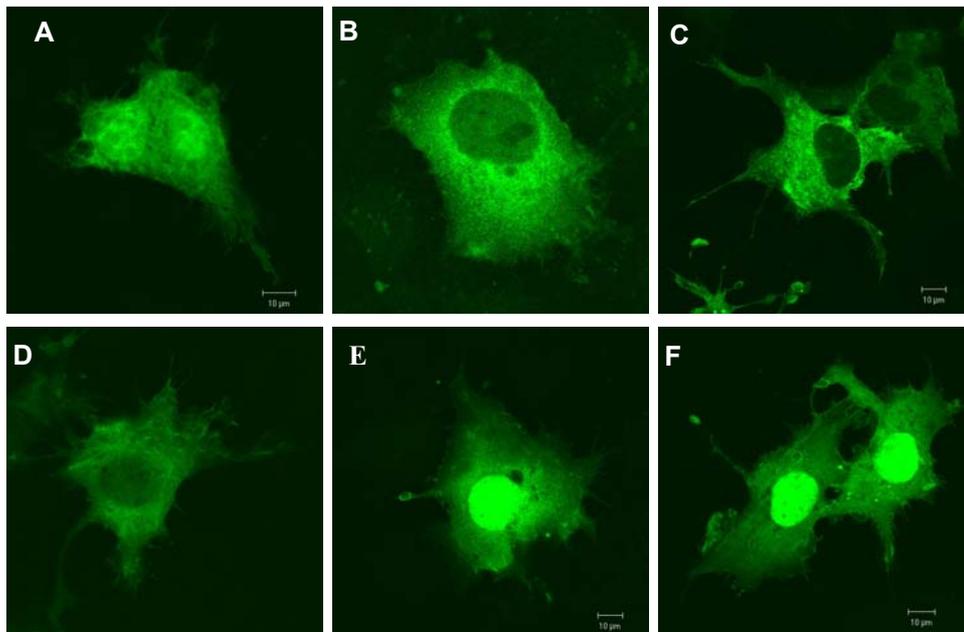
“pleckstrin homology” (PH) que pode estar envolvido com interações proteína-proteína que ocorrem na sinalização celular. A proteína FAPP1 poderia estar participando deste evento possivelmente como uma molécula adaptadora (Dowler et al., 2000). Formas mutantes de proteínas com domínio PH estão relacionadas a câncer e desordens relacionadas ao desenvolvimento (Shaw, 1996).

Nenhuma das pesquisas realizadas pelo Blast detectou qualquer similaridade com proteínas conhecidas ou presença de domínio conservado para as proteínas c21orf59 e c21orf83-2. As predições “in silico” para a proteína c21orf83-2 indicam uma localização mitocondrial e citoplasmática respectivamente com 39% e 22% de probabilidade (Tabela 5), porém dados experimentais demonstram que GFP-ORF83-2 não é uma proteína mitocondrial e sim preferencialmente citoplasmática em todas as linhagens de células estudadas (Figura 32). O mesmo padrão foi observado para a fusão GFP-ORF59 confirmando a predição de localização realizada pelo algoritmo PSORT II com 61% de probabilidade e ausência de NLS típico (Figura 32).

Chatterjee e Fischer (2000), verificaram que resíduos de leucina são chaves para a localização citoplasmática das proteínas RGS4 e RGS16 estudadas por eles. Na proteína DCRB esses aminoácidos representam 11% da seqüência total, na c21orf45 são 13% e na ORF83-2 são 9%, o que poderia estar contribuindo para localização citoplasmática dessas proteínas.



**Figura 31: Análise da localização intracelular da proteína DCRB e das suas formas truncadas.** Imagens obtidas por microscopia de epifluorescência convencional evidenciando a localização intracelular das formas estudadas na linhagem celular denominada COS-7. (A e B) pEGFPDCRB; (C e D) pEGFPDCRB-primeira parte e (E e F) pEGFPDCRB-segunda parte. As imagens B, D e F são da coloração com DAPI, evidenciando a integridade da membrana nuclear.



**Figura 32: Análise da localização intracelular das proteínas estudadas.** Visualização por microscopia confocal de células COS-7 transfectadas com pEGFP-C1(A), GFP-ORF45 (B), GFP-ORF83-2 (C), GFP-ORF59 (D), GFP-ORF 95 (E) e GFP-ORF101 (F).

A c21orf95 possui 100% identidade com a proteína CYYR1 uma proteína que possui um domínio central rico em cisteína e tirosina o qual é altamente conservado entre os vertebrados desde os peixes até humanos, apresenta 81% de identidade com o ortólogo CYYR1 de rato e pode estar relacionada a tumores do sistema neuroendócrino (Vitale et al., 2002). Todas as proteínas desta família são preditas conter um provável domínio transmembrana, juntamente com um peptídeo sinal sugerindo uma possível localização na membrana celular, entretanto, em nossas análises, a GFP-ORF95 revela uma localização preferencialmente nuclear em todas as linhagens de células estudadas confirmando a predição de localização nuclear de 61% de probabilidade, porém sem a presença de um NLS típico (Figura 32). Além disso, a c21orf95 possui uma predominância de aminoácidos

prolina (13%) e 10% de aminoácidos básicos em sua seqüência que são características comuns em proteínas nucleares (Cedano et al., 1997 e Andrade et al., 1998).

A seqüência de aminoácidos da c21orf101 revelou um sinal de localização nuclear do tipo KKRRK (aminoácidos 121 a 125) no c-terminal quando analisada pelo algoritmo PredictNLS. A mesma predição foi feita pelo PSORT II embora com uma probabilidade de apenas 35%. Os resultados foram confirmados através da análise de células transientemente transfectadas com GFP-ORF101 que apresentaram uma localização preferencialmente nuclear em todas as linhagens estudadas confirmando a predição de localização celular (Figura 32).

A proteína c21orf101 apresenta um domínio conservado de proteína ribossômica mitocondrial (MRP) e 100% de identidade com a proteína ribossômica mitocondrial S6 (MRPS6). Todas as MRPs de mamíferos são codificadas por genes nucleares, sintetizadas por ribossomos no citoplasma e importadas para as mitocôndrias (Kenmochi et al., 2001) entretanto este padrão não foi verificado em nosso trabalho.

Os dados a respeito da localização subcelular de proteínas que não possuem qualquer domínio conservado ou similaridade com proteínas conhecidas são muito importantes e servirão de base para experimentos futuros já que são as primeiras informações a respeito dessas proteínas. Considerando a escassez de informações a respeito da função das proteínas potencialmente relacionadas à Síndrome de Down, o presente trabalho vem contribuir com análises da expressão, purificação e as localizações subcelulares de proteínas localizadas na DSCR bem como sugerir as funções dessas proteínas com base na presença de domínios conservados presentes em proteínas com funções já descritas. Assim, é provável que estas proteínas estejam relacionadas ao transporte e sinalização celular, desenvolvimento de tumores, problemas cardíacos, Mal de

Alzheimer e doenças neuronais que são as principais características presentes em portadores da Síndrome de Down.

## 5. CONCLUSÕES

- 1- As ORF DCRB, c21orf45, c21orf59, c21orf83-2, c21orf95 e c21orf101 foram isoladas com sucesso.
- 2- O gene DCRB de *Homo sapiens* possui a mesma organização genômica que a de *Pan troglodytes*.
- 3- O gene DCRB possui similaridade com vários ESTs humanos depositados em banco de dados.
- 4- Além da expressão em placenta, a expressão da DCRB também foi detectada em linhagens de células tumorais de pulmão e melanoma.
- 5- Dentre as proteínas analisadas apenas a DCRB, a c21orf45 e a c21orf83-2 foram expressas na forma solúvel quando utilizou-se a cepa Rosetta (DE3), porém em quantidades muito pequenas.
- 6- Baseado nas análises de deleção e mutação sítio dirigida, pôde-se concluir que a insolubilidade da proteína DCRB provavelmente é devido a sua composição de aminoácidos e não decorrente de formação de pontes dissulfeto entre as cadeias polipeptídicas.
- 7- O renovelamento da proteína DCRB foi realizado de forma satisfatória.
- 8- A produção do antissoro da DCRB foi obtida com sucesso, reconhecendo a proteína recombinante bem como a proteína endógena em células COS-7.
- 9- As proteínas DCRB, c21orf45, c21orf59 e c21orf83-2 são citoplasmáticas enquanto que a c21orf95 e a c21orf101 são predominantemente nucleares.

## 6. REFERÊNCIAS BIBLIOGRÁFICAS

- ABRÃO POSSIK, P. et al. DSCR2, a Down syndrome critical region protein, is localized to the endoplasmic reticulum of mammalian cells. **European Journal of Histochemistry**: 49 (3), 267-272, 2004.
- AMANO, K. et al. Dosage-dependent over-expression of genes in the trisomic region of Ts1Cje mouse model for Down syndrome. **Human Molecular Genetics**. 13 (13): 1333-1340, 2004.
- ANDRADE, M. A.; O'DONOGHUE, S. I. e ROST, B. Adaptation of Protein Surfaces to Subcellular Location. **J Mol Biol** 276: 517-525, 1998.
- ANTONARAKIS, S. E. 10 Years of *Genomics*, Chromosome 21, and Down Syndrome. **Genomics** 51: 1-16, 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.
- ARAÚJO, A. P. U. et al. Influence of the histidine tail on the structure and activity of recombinant chlorocatechol 1,2-dioxygenase. **Biochem. Biophys. Res. Com.** 272(2): 480-484, 2000.
- 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. Proteoma-scale purification of human proteins from bacteria. **Proc Natl Acad Sci U S A** 99:2654-2659, 2002.
- CAPONI, 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 Location of Proteins. **J. Mol. Biol.** 266: 594-600, 1997.
- CHATTERJEE, T. K e FISCHER, R.A. Cytoplasmic, Nuclear, and Golgi Localization of RGS Protein. **The Journal of Biological Chemistry** 275 (31): 24013-24021, 2000.
- CHEON, M. S. et al. Protein levels of genes encoded on chromosome 21 in fetal Down syndrome brain: Challenging the gene dosage effect hypothesis (Part I). **Amino Acids** 24: 111-117, 2003a.
- CHEON, M. S. et al. Protein levels of genes encoded on chromosome 21 in fetal Down syndrome brain: Challenging the gene dosage effect hypothesis (Part II). **Amino Acids** 24: 119-125, 2003b.

- CHEON, M. S. et al. Protein levels of genes encoded on chromosome 21 in fetal Down syndrome brain: Challenging the gene dosage effect hypothesis (Part III). **Amino Acids** 24: 127-134, 2003c.
- CHEON, M. S. et al. Protein levels of genes encoded on chromosome 21 in fetal Down syndrome brain: Challenging the gene dosage effect hypothesis (Part IV). **Amino Acids** 25: 41-47, 2003d.
- CORPET, F. Multiple sequence alignment with hierarchical clustering. **Nucl. Acids Res.** 16: 10.881-10890, 1988.
- COKOL, M.; NAIR, R. e ROST, B. Finding Nuclear localization signals. **EMBO Reports** 1: 411-415, 2000.
- CORRÊA, E. M. “Expressão heteróloga, purificação e caracterização das proteínas humanas DCRA (Down Syndrome Critical Region Gene A) e DSCR8 (Down Syndrome Critical Region Gene 8). Tese de doutorado. Universidade Federal de São Carlos, 2004.
- CRAWFORD, D.R. et al. Hamster *adapt78* mRNA is a Down Syndrome Critical Region Homologue that is Inducible by Oxidative Stress. **Arch. Biochem. Biophys.** 342 (1): 6-12, 1997.
- 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.
- 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.
- DOWLER, S. et al. Identification of pleckstrin-homology-domain-containing proteins with novel phosphoinositide-binding specificities. **Biochem J** 351: 19-31, 2000.
- DOWN, J.L.H. Observations on an Ethnic Classification of idiots. **London Hosp. Clin. Lect. Rep.** 3: 259, 1866.
- ENGIDAWORK, E. et al. Expression of hypothetical proteins in human fetal brain: increased expression of hypothetical protein 28.5 KDa in Down Syndrome, a clue for its tentative role. **Molecular Genetics and Metabolism** 78: 295-301, 2003.
- EPSTEIN, C. The Consequences of Chromosome Imbalance: Principles, Mechanisms and Models. Cambridge University Press, New York. 1986.
- EPSTEIN, C. et al. Protocols to Establish Genotype-Phenotype Correlations in Down Syndrome. **Am. J. Hum. Genet.** 49: 207-235, 1991.
- FAN, C. et al. Expression Patterns of Two Murine Homologs of *Drosophila Single-Minded* Suggest Possible Roles in Embryonic Patterning and in the Pathogenesis of Down Syndrome. **Molecular and Cellular Neuroscience** 7: 1-16, 1996.

- FERNANDEZ, J. e HOEFFLER, J. P. Gene Expression Systems: using Nature for the art of expression, 1998.
- FERNANDO-MIGUEL, R.; CHEON, M. S. e LUBEC, G. Protein levels of genes encoded on chromosome 21 in fetal Down syndrome brain (Part V): Overexpression of phosphatidyl-inositol-glycan class P protein (DSCR5). **Amino Acids** 26: 255-261, 2004.
- FONG, C. T. e BRODEUR, G. M. Down's syndrome and leukemia: epidemiology, genetics, cytogenetics and mechanisms of leukemogenesis. **Cancer Genet Cytogenet.** 28: 55-76, 1987.
- FUENTES, J. J. et al. A New Human Gene from the Down Syndrome Critical Region Encodes a Proline-Rich Protein highly Expressed in Fetal Brain and Heart. **Human Mol. Genet.**( 4): 1935-1944, 1995.
- FUENTES, J. J., PRITCHARD, M. A. e ESTIVILL, X. Genomic Organization, Alternative Splicing, and Expression Patterns of the *dscr1* (Down Syndrome Candidate Region 1) Gene. **Genomics.** 44 (3): 358-361, 1997.
- FUENTES, J. J. et al. DSCR1, overexpressed in Down syndrome, is an inhibitor of calcineurin-mediated signaling pathways. **Human Molecular Genetics** 9:(11): 1681-1690, 2000.
- GARDINER, K. et al. Annotation of human chromosome 21 for relevance to down syndrome: gene structure and expression analysis. **Genomics** 79(6): 833-843, 2002.
- GEISSE, S. et al. Eukaryotic expression systems: a comparison. **Protein Expression and Purification.** 8(3): 271-282, 1996.
- GUIMERA, J. et al. A human homologue of *Drosophila minibrain* (MNB) is expressed in the neural regions affected in Down syndrome and maps to the critical region. **Human Molecular Genetics** 5 (9): 1305-1310, 1996.
- GUIMERA, J. et al. Cosmid Contig and Transcriptional Map of Three Regions of Human Chromosome 21q22: Identification of 37 Novel Transcripts by Direct Selection. **Genomis** 45: 59-67, 1997
- GUIMERA, J. et al. Human *Minibrain Homologue* (*MNBH/DYRK1*): Characterization, Alternative Splicing, Differential Tissue Expression, and Overexpression in Down syndrome. **Genomics** 57: 407-418, 1999.
- 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.
- HANNIG, G. e MARKRIDES, S. C. Strategies for optimizing heterologous protein expression in *Escherichia coli*. **Trends in Biotechnology** 16: 54-60, 1998.

- HASSOLD, T. e JACOBS, P. Trissomy in Man. **Annu. Rev. Genet.** 18: 69-97, 1984.
- HATTORI, M. et al. The DNA sequence of human chromosome 21. **Nature** 405: 311-319, 2000.
- HILL, D. A. et al. Mortality and cancer incidence among individuals with Down syndrome. **Arch Intern Med.** 163: 705-711, 2003.
- HOBBS, C.A. et al. Polimorphisms in Genes Involved in Folato Metabolism as Maternal Risk Factors for Down Syndrome. **Am. J. Hum. Genet.** 67: 623-630, 2000.
- HOOKE, E. B. Down Syndrome rates and relaxed selection at older maternal ages. **Am. J. Hum. Genet.** 53 (6): 1307-1313, 1983.
- HORTON, R. M. et al. Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. **Biotechniques** 8(5): 528-535, 1990.
- HIZUKA, M. et al. Down Syndrome Candidate Region 1, a Downstream Target of VEGF, Participates in Endothelial Cell Migration and Angiogenesis. **J Vasc Res** 41: 334-344, 2004.
- JONASSON, P. et al. Genetic design for facilitated production and recovery of recombinant proteins em *Escherichia coli*. **Biotechnol. Appl. Biochem.** 35: 91-105, 2002.
- KENMOCHI, N. et al. The Human Mitochondrial Ribosomal Protein Genes: Mapping of 54 Genes to the Chromosomes and Implications for Human Disorders. **Genomics** 77: 65-70, 2001.
- KORENBERG, J. R. et al. Familial Down Syndrome with Normal Karyotype: Definition of the Region. **Am. J. Hum. Genet.** 43: A 110, 1988.
- KORENBERG, J. R., BRADLEY, C. e DISTECHE, C. M. Down Syndrome: Molecular Mapping of the Congenital Heart Disease and Duodenal Stenosis. **Am. J. Hum. Genet.** 50: 294-302, 1992.
- KORENBERG, J. et al. Down Syndrome phenotypes: The Consequence of Chromosome Imbalance. **Proc. Natl. Acad. Sci.** 91: 4997-5001, 1994.
- LAEMMLI, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **Nature** 227:680-685, 1970.
- LEJEUNE, J., TURPIN, R. e GAUTIER, M. Le Mongolisme, premier exemple d'aberration autosomique humaine. **Ann. Génét.** 1: 41-49, 1959.
- LI, S. et al. Identification of a Novel Cytoplasmic Protein That Specifically Binds to Nuclear Localization Signal Motifs. **The Journal of Biological Chemistry** 273 (11): 6183-6189, 1998.

- MAO, R. et al. Global up-regulation of chromosome 21 gene expression in the developing Down syndrome brain. **Genomix** 81: 457-467, 2003.
- MATTEI, J. F. et al. Trisomy 21 for the Region 21q23: Identification by High-Resolution R Banding Patterns. **Hum. Genet.** 56: 409-411, 1981.
- MIYAZAKI, T. et al. Molecular cloning of a novel thyroid hormone-responsive gene, ZAKI-4, in human skin fibroblasts. **Journal of Biological Chemistry** 271(24): 14567-71, 1996.
- MURBY, M.; UHLÉN M. e STAHL, S. Upstream Strategies to Minimize Proteolytic Degradation upon Recombinant Production in *Escherichia coli*. **Protein Expression and Purification.** 7: 129-136, 1996.
- NAKAI, K. e HORTON, P. PSORT: a program for detecting the sorting signals of proteins and prediction their localization. **Trends Biochem Sci** 24(1):34-35, 1999.
- NAKANISH, K., BEROVA, N. e WOODY, R.W. Circular Dichroism-Principles and Applications. Nova York, V.CH Publishers, INC 1994.
- NAKAMURA, A.; HATTORI, M. e SAKAKI, Y. Isolation of a novel human gene from the Down Syndrome critical region of chromosome 21q22.2. **J. Biochem (Tokyo)**, 122(4): 872-7, 1997a.
- NAKAMURA, A.; HATTORI, M. e SAKAKI, Y. A novel gene isolated from human placenta located in Down Syndrome Critical Region on chromosome 21. **DNA Res.** 4(5): 321-4, 1997b.
- OLOF, E. et al. Prediction subcellular localization of proteins based on their N-terminal acid sequence. **J. Mol. Biol.** 3000: 1005-1016, 2000.
- PFISTER, S. M. et al. Mutational analyses of the signals involved in the subcellular location of DSCR1. **BMC Cell Biology** 3: 24, 2002.
- PRODROMOU, C. e PEARL, L. H. Recursive PCR: a novel technique for total gene synthesis. **Protein Engineering** 5 (8): 827-829, 1992.
- QIAExpressionist. A handbook for high-level expression and purification of 6xHis-tagged proteins, 2003.
- RAHMANI, Z. 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.
- REYMOND, A. et al. From PREDs and Open Reading Frames to cDNA isolation: Revisiting the Human Chromosome 21 Transcription Maps. **Genomics** 78 (1-2): 46-54, 2001.
- REYMOND, A. et al. Nineteen Additional Unpredicted Transcripts from Human Chromosome 21. **Genomics** 79 (6): 824-832, 2002.

- ROIZEN, N. J. e PATTERSON, D. Down's Syndrome. **Lancet** 361: 1281-89, 2003.
- ROTHERMEL, B. et al. A protein encoded within the Down Syndrome Critical Region is enriched in striated muscles and inhibits Calcineurin signaling. **Journal of Biological Chemistry** 275: 8719-8725, 2000.
- SAGO, H. et al. Ts1Cje, a new partial trisomy 16 mouse model for Down Syndrome, exhibits learning and behavioral abnormalities. **Proc. Natl. Acad. Science USA** 95: 6256-6261, 1998.
- SALAMOV, A. A.; NISHIKAWA, T. e SWINDELLS, M. B. Assessing protein coding region integrity in cDNA sequencing projects. **Bioinformatics** 14(5):384-90, 1998.
- SAMBROOK, J. e RUSSELL, D. W. Molecular Cloning: a laboratory manual. 3<sup>o</sup> Edition. **Cold Spring Harbor Press**, New York: Cold Spring Harbor Laboratory, 2001.
- SANGER, F., NICLEN, S. e COULSON, A. R. DNA sequencing with chain terminating inhibitors. **Proc. Natl. Acad. Science USA** 74: 5463-5467, 1977.
- SATI, S.P. et al. Extra terminal residues have a profound effect on the folding and solubility of a Plasmodium falciparum sexual stage-specific protein over-expressed in *Escherichia coli*. **Eur J Biochem** 269(21): 5259-63, 2002.
- SHAW, G. The pleckstrin homology domain: na intriguing multifunctional protein module. **Bioessays** 18(1): 35-46, 1996.
- SHIM, K. S. e LUBEC, G. Debrin, a dendritic spine protein, is manifold decreased in brains of patients with Alzheimer's disease and Down syndrome. **Neuroscience Letters** 324: 209-212, 2002.
- SHIBUYA, K. et al. Isolation of Two Novel Genes, DSCR5 and DSCR6, from Down Syndrome Critical Region on Human Chromosome 21q22.2. **Biochemical and Biophysical Research Communications** 271, 693-698, 2000.
- SILVEIRA, H. C. S. et al. "A calcineurin inhibitory protein overexpressed in Down's Syndrome interacts with the product of a ubiquitously expressed transcript (UXT)". **Brazilian Journal of Medical and Biological Research** 37: 785-789, 2004.
- SOMMER, C. A. "Expressão heteróloga e ensaios de purificação da proteína DSCR2 para a produção de anticorpos e posteriores estudos estruturais. Dissertação de Mestrado, Universidade Federal de São Carlos, 2003.
- SREERAMA, N. e WOODY, R.W. A Self-Consistent Method for the analysis of Protein Secondary Structure from Circular Dichroism. **Anal. Biochem.** 209: 32-44, 1993.

- SREERAMA, N.; VENYAMINOV, S.Y. e WOODY, R.W. Estimation of the number of a-helical and b-strand segments in proteins using CD spectroscopy. **Protein Sci.** 8: 370-380, 1999.
- SREERAMA, N. e WOODY, R.W. Estimation of protein secondary structure from CD spectra: Comparison of CONTIN, SELCON and CDSSTR methods with an expanded reference set. **Anal. Biochem.** 282: 252-260, 2000.
- STEVENS, R. C. Design of high-throughput methods of protein production for structural biology. **Structure Fold Des** 8 (9): R-177-185, 2000.
- TAKAMATSU, K. et al. Identification of two novel primate-specific genes in DSCR. **DNA Research** 30; 9(3): 89-97, 2002.
- TOYODA, A. et al. Comparative Genomic Sequence Analysis of the Human Chromosome 21 Down Syndrome Critical Region. **Genome Research** 12: 1323-1332, 2002.
- TSUMOTO, K. et al. Practical considerations in refolding proteins from inclusion bodies. **Protein Expression and Purification** 28: 1-8, 2003.
- VIALARD, F. et al. Overexpression of mSim2 gene in the zona limitans of the diencephalon of segmental trisomy 16 Ts1Cje fetuses, a mouse model for trisomy 21: a novel whole-mount based RNA hybridization study. **Developmental Brain Research** 121: 73-78, 2000.
- VIDAL-TABOADA, J. M. et al. Identification and Characterization of a New Gene from Human Chromosome 21 between Markers D21S343 and D21S268 Encoding a Leucine-Rich Protein. **Biochemical and Biophysical Research Communications** 250: 547-554, 1998.
- VIDAL-TABOADA, J. M. et al. Down Syndrome Critical Region Gene 2: Expression during Mouse Development and in Human Cells Lines Indicates a Function Related to Cell Proliferation. **Biochemical and Biophysical Research Communications** 272: 156-163, 2000.
- VITALE, L. et al. Cysteine and tyrosine-rich 1 (CYR1), a novel unpredicted gene on human chromosome 21 (21q21.2), encodes a cysteine and tyrosine-rich protein and defines a new family of highly conserved vertebrate-specific genes. **Gene** 290: 141-151, 2002.
- VOGEL, F. e MOTULSKY, A. G. Human Genetics. Problems and Approches. 3<sup>rd</sup> Edition. Springer-Verlag Berlin Heidelberg, Germany, 1997.
- YAMAMOTO, Y. et al. A Case of 21q Syndrome with Normal SOD1 Activity. **Hum. Genet.** 48: 27-31, 1986.
- YASUKAWA, T. et al. Increase of solubility of Foreign Proteins in *E. coli*. by Coproduction of Bacterial Thioredoxin. **J. Biol. Chem.** 270: 25328-25331, 1995.

- WANG, A. et al. A novel method for increasing the expression level of recombinant proteins. **Protein Expression and Purification** 30: 124-133, 2003.
- WHITE, B. A. PCR Cloning Protocols from molecular cloning to genetic engineering. **Methods in Molecular Biology**, 1997.
- WOODY R., BEROVA N. e NAKANISHI K. Circular Dichroism: Principles and applications. VCH publishers, NY, USA, 1994.

## ANEXO

Tabela 5: Análises in silico e in vivo das proteínas estudadas

Proteína/ Número de Aminoácidos	NCBI Blast	Resíduos de Aminoácidos hidrofóbicos (%)	Resíduos de Aminoácidos hidrofilicos (%)	Predominância de resíduos de Aminoácidos (%)	Predict NLS	TargetP	
DCRB (118)	47% identidade com Drebrina	49%	51%	Serina (14%) Leucina (11%) Prolina (10%)	Ausente	<b>68% outr</b> 24% secr 9% mito	
c21orf 45 (233)	96% identidade com FAPP1	47%	53%	Serina (13%) Ác. glutâmico (13%) Leucina (13%)	Ausente	<b>70% outr</b> 32% mito 4,5% secr	
c21orf 59 (290)	Ausente	47%	53%	Ác. glutâmico (11%) Lisina (10%) Alanina (7%)	Ausente	<b>87% outr</b> 13% mito 6% secr	
c21orf 83-2 (178)	Ausente	<b>60%</b>	<b>40%</b>	Serina (11%) Alanina (9%) Leucina (9%) Prolina (8%)	Ausente	31% outro <b>57% mito</b> 8,3% secr	
c21orf 95 (154)	100% identidade com CYR1	<b>62%</b>	<b>38%</b>	Prolina (13%) Glicina (9%) Alanina (8%) Isoleucina (8%)	Ausente	7% outro 3% mito <b>95% secr</b>	
c21orf 101 (125)	Domínio MRP e 100% identidade com MRPS6	49,6%	50,4%	Leucina (11%) Alanina (10%) Arginina (8%) Ác. glutâmico (8%)	<b>KKRKK</b> C-terminal	<b>68% outr</b> 14% mito 19% secr	

