

UNIVERSIDADE FEDERAL DE MINAS GERAIS  
INSTITUTO DE CIÊNCIAS BIOLÓGICAS  
DEPARTAMENTO DE BIOLOGIA GERAL  
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA



DISSERTAÇÃO DE MESTRADO

**Caracterização *in silico*, clonagem e expressão heteróloga em *Escherichia coli* de cinco proteínas de *Corynebacterium pseudotuberculosis* possivelmente implicadas na virulência deste microrganismo**

**Karina Talita de Oliveira Santana**

**ORIENTADOR: Prof. Dr. Vasco Azevedo**

BELO HORIZONTE

AGOSTO – 2014

**KARINA TALITA DE OLIVEIRA SANTANA**

**Caracterização *in silico*, clonagem e expressão heteróloga em *Escherichia coli* de cinco proteínas de *Corynebacterium pseudotuberculosis* possivelmente implicadas na virulência deste microrganismo**

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre pelo programa de Pós-Graduação em Genética, Departamento de Biologia Geral, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais.

**ORIENTADOR: Prof. Dr. Vasco Azevedo**

**BELO HORIZONTE**

**AGOSTO - 2014**

*Sem Deus, o universo não é explicável satisfatoriamente.*

*(Albert Einstein)*

*Cuide para que todo o seu foco na vida esteja nas  
soluções, não nos problemas.*

*(Anthony Robbins)*

## **COLABORAÇÕES**

Este trabalho é fruto da colaboração entre o Laboratório de Genética Celular e Molecular do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais (UFMG), Prof. Roberto Meyer e Prof. Ricardo Wagner Dias Portela (Laboratório de Imunologia e Biologia Molecular da Universidade Federal da Bahia) e Prof. Dr. Raghuvir Krishnaswamy Arni e Ricardo Barros Mariutti (Laboratório de Biologia Molecular da Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP).

## AGRADECIMENTOS

À Deus, por estar ao meu lado a cada momento e por tudo que realiza em minha vida desde sempre e para sempre.

À Universidade Federal de Minas Gerais e ao Instituto de Ciências Biológicas.

Ao Programa de Pós-Graduação em Genética da UFMG e seus professores, funcionários e colegas.

À CAPES pelo apoio financeiro.

Ao Professor Doutor Vasco Azevedo, pelos ensinamentos e pela oportunidade e confiança ao ter aberto as portas do Laboratório de Genética Celular e Molecular e possibilitado este crescimento pessoal e profissional.

Ao Professor Doutor Anderson Myoshi, pela disponibilidade e ensinamentos.

Aos membros da banca, pela disponibilidade e valiosas contribuições.

Ao Professor Dr. Arni Raghuvir Krishnaswamy Arni e ao Ricardo Mariutti do Laboratório de Biologia Molecular da Universidade Estadual Paulista “Júlio de Mesquita Filho” – Unesp pela colaboração.

Ao Professor Dr. Roberto Meyer pela colaboração.

À galera do LGCM por ter compartilhado experiências, conhecimentos, alegrias e aflições: Thiago Castro, Dayana Ribeiro, Camila Prósperi, Pâmela Mancha, Bianca Souza, Flávia Rocha, Kátia Moraes, Rachid Aref, Alfonso Gala-García, Tessália Saraiva, Wanderson Marques, Mariana Santana, Roseliane, Anne Pinto, Sara Silva, Priscilla Bagano, Camila Azevedo, Meritxell Zurita, Marcela Pacheco, Fernanda Dorella, Rodrigo Dias, Tatiane Preisser, Vanessa Bastos, Letícia de Castro, Luis Guimarães, Carlos Diniz, Alberto Junior, Pammella Teixeira, Edgar Aguiar, Lucas Amorim, Diego Mariano, Thiago Sousa, Siomar Soares, Edson Folador, Flávia Figueira, Ulisses Pereira, Leandro Benevides, Marcus Canário, Syed Shah Hassan, Sandeep Tiwari, Syed Babar Jamal.

À Fernanda Magalhães por toda dedicação e por ser uma pessoa tão especial.

Aos que compartilharam comigo não apenas a bancada, mas a solicitude, amizade, parceria e ensinamentos.

À Anne queridíssima, por todo apoio, ajuda, ensinamentos e conselhos. Muito obrigada!

À Priscilla, por toda a disponibilidade e por ter me socorrido sempre que precisei. Muito obrigada!

À Tatiana Silveira e Miriam Franco do Laboratório de Imunologia de Doenças Infecciosas, sempre solícitas, obrigada!

À Ana Cris Hostt do Laboratório de Imunobiologia pela ajuda e disponibilidade.

Ao Edgar e ao Alberto pelas preciosas ajudas “bioinformáticas”, por serem tão solícitos e queridos.

Às minhas amigas queridas Natayme Tartaglia e Renata Silva pela amizade, parceria, conselhos, ombro amigo e apoio no lab e fora dele. Vocês são um presente!

À minha mãe Marlene, meus irmãos Cássia e Vinícius e ao meu pai Edilson por todo amor, carinho, dedicação, orações, amparo e força. Por estarem presentes a cada momento e me motivarem a alcançar os meus sonhos. Muito obrigada! Amo vocês!!! Não é possível expressar o quanto sou grata a vocês!

Ao Bruno Jorge, pelo amor, apoio, compreensão, paciência e carinho. Por sonhar junto comigo sempre! Eu te amo muito!

Aos amigos que não foram citados aqui mas que estão sempre no coração.

## SUMÁRIO

LISTA DE FIGURAS .....	I
LISTA DE TABELAS E QUADROS .....	IV
LISTA DE ABREVIATURAS .....	VI
RESUMO .....	11
ABSTRACT .....	12
1 INTRODUÇÃO .....	13
1.1 O microrganismo <i>Corynebacterium pseudotuberculosis</i> .....	13
1.2 A Linfadenite Caseosa .....	14
1.2.1 Características gerais .....	14
1.2.2 Epidemiologia .....	14
1.2.3 Transmissão .....	15
1.2.4 Patogênese .....	15
1.2.5 Diagnóstico .....	16
1.2.6 Tratamento, profilaxia e vacina .....	17
1.3 Fatores de virulência de <i>C. pseudotuberculosis</i> .....	19
1.3.1.1 Fosfolipases .....	20
1.3.1.2 Fosfolipase D em <i>C. pseudotuberculosis</i> .....	22
1.3.2 Proteínas potencialmente envolvidas na virulência de <i>C. pseudotuberculosis</i> .....	24
1.3.2.1 Proteína Quinase G – PknG .....	25
1.3.2.2 Adesina – SpaC .....	29
1.3.2.3 Superóxido-dismutase C – SodC ou Cu/Zn-SOD .....	31
1.3.2.4 Neuraminidase H (NanH) .....	33
1.4 Expressão heteróloga de proteínas em <i>Escherichia coli</i> .....	34
2 JUSTIFICATIVA .....	37
3 OBJETIVOS .....	38
3.1 Objetivos Gerais .....	38
3.2 Objetivos Específicos .....	38
4 MATERIAIS E MÉTODOS .....	39
4.1 Caracterização <i>in silico</i> das ORF's e proteínas .....	39
4.1.1 Predição de parâmetros físico-químicos .....	39
4.1.2 Predição de peptídeo sinal .....	40
4.1.3 Predição de localização subcelular .....	40
4.1.4 Busca por domínios conservados .....	40
4.1.5 Alinhamento de sequências e busca por similaridade .....	41

4.1.6	Predição de epítomos .....	42
4.2	Clonagem molecular e expressão das proteínas recombinantes .....	43
4.2.1	Vetor de expressão pD444-NH e construção das ORF's sintéticas das proteínas selecionadas .....	43
4.2.2	Preparação do DNA plasmidiano .....	44
4.2.3	Transformação bacteriana .....	45
4.2.3.1	Preparo de células eletrocompetentes .....	45
4.2.3.2	Transformação de linhagens de <i>E. coli</i> com os plasmídeos pD444-NH: <i>pld</i> , pD444-NH: <i>pknG</i> , pD444-NH: <i>spac</i> , pD444-NH: <i>sodC</i> e pD444-NH: <i>nanh</i> .....	46
4.2.4	Extração em pequena escala de DNA plasmidiano .....	46
4.2.5	Indução da expressão das proteínas recombinantes em <i>E. coli</i> .....	47
4.2.5.1	Indução da expressão das proteínas recombinantes em diferentes linhagens de expressão .....	47
4.2.5.2	Resolução eletroforética de proteínas em gel de poliacrilamida sob condições desnaturantes – SDS-PAGE .....	48
4.2.5.3	Avaliação da cinética de expressão das proteínas recombinantes ....	49
4.2.5.4	Teste de solubilidade das proteínas recombinantes .....	49
4.2.6	Purificação da proteína recombinante .....	50
4.2.7	<i>Western Blot</i> .....	51
5	RESULTADOS .....	52
5.1	Caracterização <i>in silico</i> das ORFs e proteínas .....	52
5.1.1	Predição de parâmetros físico-químicos .....	52
5.1.2	Predição de peptídeo sinal` .....	52
5.1.3	Predição de localização subcelular .....	53
5.1.4	Predição de domínios conservados .....	54
5.1.5	Alinhamento de sequências e busca por similaridade .....	56
5.1.5.1	PLD .....	57
5.1.5.2	PknG .....	59
5.1.5.3	SpaC .....	65
5.1.5.4	SodC .....	67
5.1.5.5	NanH .....	71
5.1.6	Predição de epítomos .....	75
5.2	Expressão heteróloga em <i>E. coli</i> das proteínas recombinantes rPLD, rPknG, rSpaC, rSodC e rNanH de <i>C. pseudotuberculosis</i> FRC41 .....	80
5.2.1	Construção dos vetores de expressão contendo as ORF's Sintéticas .....	80



5.2.2	Avaliação da produção das proteínas rPLD, rPknG, rSpaC, rSodC e rNanH em diferentes linhagens de expressão .....	80
5.2.3	Avaliação da cinética de expressão das proteínas rPLD, rPknG, rSpaC, rSodC e rNanH .....	84
5.2.4	Teste de solubilidade das proteínas recombinantes .....	88
5.2.5	Purificação da proteína PknG .....	89
5.2.6	Confirmação da expressão da proteína rPknG por <i>Western blot</i> .....	89
6	DISCUSSÃO .....	90
7	CONCLUSÕES .....	101
8	PERSPECTIVAS .....	101
9	REFERÊNCIAS BIBLIOGRÁFICAS .....	102
10	ANEXOS .....	117
	Anexo 1 - Sequências das <i>ORF's pld, pkg, spac, sodc e nanh</i> otimizadas com códons preferenciais de <i>E. coli</i> .....	118
	Anexo 2 - Predição de peptídeo sinal das proteínas SpaC, SodC e NanH .....	121
	Anexo 3 - Alinhamentos de sequências .....	124
	Anexo 4 - Mapa dos plasmídeos pD444-NH: <i>pkg</i> , pD444-NH: <i>spac</i> , pD444-NH: <i>sodc</i> e pD444-NH: <i>nan</i> .....	144

## LISTA DE FIGURAS

<b>Figura 1:</b> Estrutura de um glicerofosfolípideo e o sítio de ação das fosfolipases.....	21
<b>Figura 2.</b> Estrutura tridimensional da PknG de <i>Mycobacterium tuberculosis</i> .....	27
<b>Figura 3.</b> Modelo de mediação por <i>pili</i> de adesão e patogênese.....	30
<b>Figura 4.</b> <i>Clusters</i> gênicos codificadores de <i>pili</i> adesivos de <i>C. pseudotuberculosis</i> FRC41.....	31
<b>Figura 5.</b> Representação esquemática do vetor pD444-NH contendo o inserto <i>pld</i> (pD444-NH: <i>pld</i> ).....	43
<b>Figura 6.</b> A) Predição de peptídeo sinal da proteína PLD através do SignalP 4.0. B) Predição de peptídeo sinal da proteína PknG através do SignalP 4.0.....	53
<b>Figura 7.</b> Representação esquemática de domínio conservado da proteína PLD de <i>C. pseudotuberculosis</i> FRC41.....	54
<b>Figura 8.</b> Representação esquemática de domínio conservado da proteína PknG de <i>C. pseudotuberculosis</i> FRC41.....	55
<b>Figura 9.</b> Representação esquemática de domínio conservado da proteína SpaC de <i>C. pseudotuberculosis</i> FRC41.....	55
<b>Figura 10.</b> Representação esquemática de domínio conservado da proteína SodC de <i>C. pseudotuberculosis</i> FRC41.....	56
<b>Figura 11.</b> Representação esquemática de domínio conservado da proteína NanH de <i>C. pseudotuberculosis</i> FRC41.....	56
<b>Figura 12.</b> Alinhamento de sequências de proteínas de microrganismos do grupo CMNR que apresentaram identidade com a proteína PLD de <i>C. pseudotuberculosis</i> FRC41.....	58

<b>Figura 13.</b> Alinhamento de porção N-terminal de sequências de serina/treonina proteínas quinases de microrganismos representantes do grupo CMNR.....	61
<b>Figura 14.</b> Alinhamento de porção central de sequências de serina/treonina proteínas quinases de microrganismos representantes do grupo CMNR que apresentaram identidade com a proteína PknG de <i>C. pseudotuberculosis</i> FRC41.....	62
<b>Figura 15.</b> Alinhamento de porção N-terminal de sequências proteicas de organismos eucariotos que apresentaram identidade com a proteína PknG de <i>C. pseudotuberculosis</i> FRC41.....	64
<b>Figura 16.</b> Alinhamento de porção N-terminal de sequências proteicas de microrganismos representantes do grupo CMNR que apresentaram identidade com a proteína SpaC de <i>C. pseudotuberculosis</i> FRC41.....	66
<b>Figura 17.</b> Alinhamento de porção N-terminal de sequências proteicas de microrganismos representantes do grupo CMNR que apresentaram identidade com a proteína SodC de <i>C. pseudotuberculosis</i> FRC41.....	68
<b>Figura 18.</b> Alinhamento de sequências proteicas de organismos eucariotos que apresentaram identidade com a proteína SodC de <i>C. pseudotuberculosis</i> FRC41.....	70
<b>Figura 19.</b> Alinhamento de porções das proteínas de microrganismos representantes do grupo CMNR que apresentaram identidade com a proteína NanH de <i>C. pseudotuberculosis</i> FRC41.....	72
<b>Figura 20.</b> Alinhamento de sequências proteicas de organismos eucariotos que apresentaram identidade com a proteína NanH de <i>C. pseudotuberculosis</i> FRC41.....	74
<b>Figura 21.</b> Avaliação da expressão da proteína recombinante rPLD.....	82
<b>Figura 22.</b> Avaliação da expressão da proteína recombinante rPknG.....	83
<b>Figura 23:</b> Avaliação da expressão da proteína recombinante rSpaC.....	83
<b>Figura 24:</b> Avaliação da expressão da proteína recombinante rSodC.....	83
<b>Figura 25:</b> Avaliação da expressão da proteína recombinante rNanH.....	84

<b>Figura 26.</b> Avaliação da cinética de expressão da proteína recombinante rPLD na linhagem BL21 Star™(DE3).....	85
<b>Figura 27.</b> Avaliação da cinética de expressão da proteína recombinante rPknG na linhagem BL21 Star™(DE3).....	86
<b>Figura 28.</b> Avaliação da cinética de expressão da proteína recombinante rSpaC na linhagem OverExpress™ C41(DE3).....	86
<b>Figura 29.</b> Avaliação da cinética de expressão da proteína recombinante rSodC na linhagem OverExpress™ C43(DE3)pLysS.....	87
<b>Figura 30.</b> Avaliação da cinética de expressão da proteína recombinante rNanH na linhagem OverExpress™ C43(DE3)pLysS.....	87
<b>Figura 31.</b> Teste de solubilidade das proteínas recombinantes.....	88
<b>Figura 32.</b> Purificação da proteína rPknG a partir da expressão na linhagem BL21 Star™(DE3) e confirmação por <i>Western blot</i> .....	89

## LISTA DE TABELAS E QUADROS

<b>Tabela 1.</b> Nomenclatura e identificadores dos genes e proteínas de <i>C. pseudotuberculosis</i> FRC41 empregados no estudo.....	39
<b>Tabela 2.</b> Plasmídeos gerados a partir da clonagem das <i>ORF's</i> em vetor de expressão pD444-NH.....	44
<b>Tabela 3.</b> Linhagens bacterianas utilizadas para clonagem molecular e expressão das proteínas recombinantes.....	45
<b>Tabela 4.</b> Peso molecular e concentração do gel de resolução (SDS-PAGE) das proteínas recombinantes.....	49
<b>Tabela 5.</b> Parâmetros físico-químicos preditos através do ProtParam.....	52
<b>Tabela 6.</b> Predição de localização subcelular pela ferramenta PSORTb.....	53
<b>Tabela 7.</b> Proteínas de microrganismos do grupo CMNR que apresentaram identidade com a proteína PLD de <i>C. pseudotuberculosis</i> FRC41.....	57
<b>Tabela 8.</b> Proteínas de microrganismos do grupo CMNR que apresentaram identidade com a proteína PknG de <i>C. pseudotuberculosis</i> FRC41.....	59
<b>Tabela 9.</b> Proteínas de organismos eucariotos que apresentaram identidade com a proteína PknG de <i>C. pseudotuberculosis</i> FRC41.....	63
<b>Tabela 10.</b> Proteínas de microrganismos do grupo CMNR que apresentaram identidade com a proteína SpaC de <i>C. pseudotuberculosis</i> FRC41.....	65
<b>Tabela 11.</b> Proteínas de microrganismos do grupo CMNR que apresentaram identidade com a proteína SodC de <i>C. pseudotuberculosis</i> FRC41.....	67
<b>Tabela 12.</b> Proteínas de organismos eucariotos que apresentaram identidade com a proteína SodC de <i>C. pseudotuberculosis</i> FRC41.....	69

<b>Tabela 13.</b> Proteínas de microrganismos do grupo CMNR que apresentaram identidade com a proteína NanH de <i>C. pseudotuberculosis</i> FRC41.....	71
<b>Tabela 14.</b> Proteínas de organismos eucariotos que apresentaram identidade com a proteína NanH de <i>C. pseudotuberculosis</i> FRC41.....	73
<b>Tabela 15.</b> Predição de densidade de epítomos de MHCI nas proteínas PLD, PknG, SpaC, SodC e NanH de <i>C. pseudotuberculosis</i> FRC41 através da ferramenta MED.....	75
<b>Tabela 16.</b> Predição de densidade de epítomos de MHCI nas regiões de menor conservação das proteínas PknG, SodC e NanH através da ferramenta MED.....	76
<b>Tabela 17.</b> Predição quantitativa de epítomos de MHCII nas proteínas PLD, PknG, SpaC, SodC e NanH de <i>C. pseudotuberculosis</i> FRC41 através da ferramenta NetMHCII.....	76
<b>Tabela 18.</b> Predição quantitativa de epítomos de MHCII nas regiões de menor conservação das proteínas PknG, SodC e NanH através da ferramenta NetMHCII.....	77
<b>Tabela 19.</b> Linhagens de <i>E.coli</i> selecionadas para a expressão das proteínas recombinantes.....	85
<b>Quadro 1.</b> Predição de epítomos lineares de células B na proteína PLD através da ferramenta BcePred.....	78

## LISTA DE ABREVIATURAS

- Amp** - Ampicilina
- BLAST** - *Basic Local Alignment Search Tool*
- ATP** – Trifosfato de adenosina
- C** - Citosina
- CAMK** – Proteína quinase Cálcio/calmodulina dependente
- CMNR** - *Corynebacterium, Mycobacterium, Nocardia e Rhodococcus*
- CMP** – Citidina monofosfato
- CTLA-4** – Linfócito T citotóxico associado a proteína 4
- CWSS** – *Cell wall sorting signal*
- DNA** - Ácido Desoxirribonucléico
- DO** - Densidade Óptica
- EBDA** – Empresa Baiana de Desenvolvimento Agrícola
- EDTA** - Ácido Etilenodiaminotetracético
- ELISA** - Ensaio de Imunoabsorção Ligado a Enzima
- g** - Grama
- G** – Guanina
- Gly** - Glicina
- GTP** – Trifosfato de guanosina
- h** - Hora
- H<sub>2</sub>O** - Água
- H<sub>2</sub>O<sub>2</sub>** – Peróxido de hidrogênio
- HCl** – Ácido clorídrico
- His** – Histidina
- HLA** - Antígeno leucocitário humano (*human leukocyte antigen*)
- I** - Induzido
- IFN** – Interferon
- IL** – Interleucina
- Ig** - Imunoglobulina
- IgG** – Imunoglobulina G
- IPTG** - Isopropil-β-D-tiogalactopiranosídeo
- KAc** – Acetato de potássio
- Kb** - Quilobases
- KCl** - Cloreto de Potássio
- KD** – Domínio quinase

**kDa** – Quilo daltons  
**KDN** - Ácido 5-hidroxineuramínico  
**KH<sub>2</sub>PO<sub>4</sub>** - Fosfato monopotássico  
**KHZ** - Kiloherztz  
**KV** – Kilovolt  
**W** - Watts  
**L** - Litro  
**lacI** – Repressor Lac  
**lacO** – Operador Lac  
**LB** - Meio de Cultura Luria-Bertani  
**LC** – Linfadenite caseosa  
**LPS** – *Lipopolissacarídeo*  
**M** – Molar  
**MED** – *Mature Epitope Density*  
**mg** - Miligrama  
**MHC** - Complexo principal de histocompatibilidade (major histocompatibility complex)  
**min** - minutos  
**mL** - Mililitro  
**mm** – milímetro  
**mM** - Milimolar  
**mRNA** – RNA mensageiro  
**N** – Normal  
**Na<sub>2</sub>HPO<sub>4</sub>** - Fosfato dissódico  
**NaCl** - Cloreto de Sódio  
**NanH** – Neuraminidase H  
**NaOH** - Hidróxido de Sódio  
**NCBI** - *National Center for Biotechnology Information*  
**Neu** – Ácido neuramínico  
**Neu5Ac** - Ácido N-acetilneuramínico  
**Neu5Gc** - Ácido N-glicosilneuramínico  
**ng** - Nanograma  
**NI** – Não induzido  
**ODH** - 2-oxoglutarato-desidrogenase  
**ORF** - fase de leitura aberta (*open reading frame*)  
**pb** - Pares de Bases  
**PBS** – Tampão fosfato salina  
**PCR** – Reação em cadeia da polimerase



**PCR** - Reação em Cadeia da Polimerase  
**pH** - Potencial Hidrogeniônico  
**pI** - Ponto isoelétrico  
**PknG** – Proteína quinase G  
**PLA** – Fosfolipase A  
**PLB** – Fosfolipase B  
**PLC** – Fosfolipase C  
**PLD** – Fosfolipase D  
**PMSF** - Fenil-metil sulfonil fluoreto  
**PVDF** - Fluoreto de polivinilideno  
**RD** – Domínio rubredoxina  
**ROS** – Espécies reativas de oxigênio  
**rpm** - Rotação por Minuto  
***rpob*** – gene subunidade beta da RNA polimerase  
**rRNA** – Ácido ribonucléico ribossômico  
**s** - segundos  
**SDS** - Dodecilsulfato de Sódio  
**SDS-PAGE** - *Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis*  
**Ser** – Serina  
**Spa** – Proteína ancorada à superfície (*surface-anchored protein*)  
**SN** – Substituição nucleofílica  
**SOD** – Superóxido dimutase  
**SodA** – Manganês superóxido dismutase  
**SodB** – Ferro Superóxido dismutase  
**SodC** – Cobre/Zinco superóxido dismutase  
**SodN** – Níquel superóxido dismutase  
**SrtA-F** – Sortases A-F  
**STP** – Serina/Treonina fosfatase  
**STPK** – Serina/Treonina quinase  
**TBE** - Tris-Borato-EDTA  
**TCA** - Ciclo do ácido cítrico  
**TCR** – Receptor de célula T  
**Th** – *T helper*  
**Thr** - Treonina  
**TIH** – Teste intradérmico de hipersensibilidade  
**TNF** – Fator de necrose tumoral  
**TPP** – Partição em três fases (*Three-phase partitioning*)

**TPR** – Tetratricopeptídeo  
**Tris** - Tris(hidroximetil)aminometano  
**tRNA** – RNA transportador  
**Tyr** – Tirosina  
**V** - Volts  
**μF** - Microfaraday (capacitância)  
**μg** - Micrograma  
**μL** - Microlitro  
**Ω** - Ohm (resistência)  
**°C** - Grau Celsius

## RESUMO

*Corynebacterium pseudotuberculosis* é o agente causador da Linfadenite caseosa (LC), doença que acomete pequenos ruminantes e acarreta diminuição de produtividade e prejuízos no agronegócio. Diferentes abordagens são estudadas visando o controle desta doença através da identificação de novas drogas ou desenvolvimento de testes diagnóstico e vacina. No entanto, ainda não foi alcançada eficácia satisfatória para o controle da LC. O sequenciamento de genomas de *Corynebacterium pseudotuberculosis* tem possibilitado a identificação de novos fatores de virulência deste microrganismo o que contribui para o melhor entendimento de sua biologia e das relações patógeno-hospedeiro. Neste trabalho, foram selecionadas como alvo de estudo, com base em análises *in silico* e busca na literatura, as proteínas PLD, PknG, SpaC, SodC e NanH. A PLD representa o principal fator de virulência de *C. pseudotuberculosis* e as demais são indicadas como potenciais fatores de virulência. Foram realizadas as caracterizações *in silico* destas proteínas a partir de predições de parâmetros físico-químicos, peptídeo sinal, domínios conservados, avaliação da sua conservação em eucariotos e epítomos. Além disso, as proteínas foram expressas utilizando sistema de expressão heteróloga em *Escherichia coli* e foram testadas cinco diferentes linhagens de expressão. A partir das linhagens que apresentaram o melhor nível de expressão de cada uma das proteínas, avaliou-se a cinética de expressão e a solubilidade da fração protéica. A proteína PknG foi purificada após otimização do protocolo de expressão para obtenção da proteína na fração solúvel e a purificação foi confirmada por *Western blot*. A partir da purificação das demais proteínas, esperamos avaliá-las quanto à sua antigenicidade e proteção contra infecção. Uma vez que fatores de virulência são importantes candidatos a alvos de drogas, testes diagnóstico ou, ainda, alvos vacinais, a obtenção destas proteínas é importante por viabilizar diversos estudos que podem contribuir para o desenvolvimento de uma estratégia de controle da LC.

## ABSTRACT

*Corynebacterium pseudotuberculosis* is the causative agent of Caseous lymphadenitis (CL), a disease that affects small ruminants and results in decreased productivity and loss in the agribusiness. Different approaches aim the control of this disease by identifying new drugs and developing new vaccines and diagnostic tests. However, there is not yet an satisfactory and efficient method for the control of LC; nevertheless, the sequencing of the genomes of *Corynebacterium pseudotuberculosis* has enabled the identification of new virulence factors of this microorganism that may contribute to a better understanding of its biology and host-pathogen relationships. In this work, proteins PLD, PknG, SpaC, SodC and NanH were selected as targets of the study based on *in silico* analysis and literature data mining. PLD is the major virulence factor of *C. pseudotuberculosis* whereas the others are pointed as potential virulence factors. *In silico* characterization of these proteins were performed from predictions of physicochemical parameters, signal peptide, conserved domains, assessment of their conservation in eukaryotes and epitopes. In addition, the proteins were also expressed using a heterologous expression system in *Escherichia coli* and five different strains were tested for expression. By using the strains that showed the highest level of expression of each protein, we assessed the expression kinetics and solubility of the recombinant proteins. The PknG protein was successfully purified after an optimization step of the expression protocol that aimed to obtain the protein in the soluble fraction. Its purification was confirmed by Western blot. Following the purification of the other proteins we will now evaluate them regarding their antigenicity and protection against infection. Since virulence factors are important candidates for drug targets, diagnostic tests or even vaccine targets, obtaining these proteins is important to perform studies that may contribute to the development of a control strategy against LC.

## 1. INTRODUÇÃO

### 1.1 O microrganismo *Corynebacterium pseudotuberculosis*

*Corynebacterium pseudotuberculosis* pertence ao gênero *Corynebacterium*, que está inserido na classe das actinobactérias e integra o grupo denominado CMNR, do qual também fazem parte os gêneros *Mycobacterium*, *Nocardia* e *Rhodococcus* (Dorella *et al.*, 2006). Os membros deste grupo apresentam uma ultraestrutura de envelope celular semelhante com complexos poliméricos compostos por peptidoglicanos, arabinogalactano e ácidos micólicos. Além disso, apresentam alto conteúdo G+C no genoma (47-74%). Várias espécies do grupo apresentam importância médica, veterinária ou biotecnológica (Dorella *et al.*, 2006). Dentre essas espécies, encontra-se a *Corynebacterium diphtheriae*, patógeno humano causador da difteria; *Corynebacterium glutamicum*, microrganismo de importância biotecnológica e industrial amplamente utilizado na produção de aminoácidos em larga escala; e *C. pseudotuberculosis*, patógeno intracelular facultativo de importância veterinária, sendo o agente etiológico da linfadenite caseosa (LC) que acomete principalmente caprinos e ovinos acarretando perdas econômicas em áreas de produção (Dorella *et al.*, 2006a). Linhagens de *C. pseudotuberculosis* pertencentes ao biovar *equi* acometem, preferencialmente, equinos, enquanto que linhagens pertencentes ao biovar *ovis* acometem pequenos ruminantes (Arsenault *et al.*, 2003; Baird, 2003; Batey, 1986; Belchior *et al.*, 2006).

*C. pseudotuberculosis* é uma bactéria Gram-positiva que exibe formas pleomórficas variando de cocóides a bastões filamentosos e pode apresentar um tamanho entre 0,5 - 0,6µm por 1 - 3µm. É uma bactéria imóvel que não esporula, não possui cápsula e possui fímbria (Jones & Collins, 1986; Baird & Fontaine, 2007). Sua condição ótima de crescimento é a 37°C em um pH de 7,0 a 7,2. Na superfície do ágar, o crescimento se dá inicialmente de forma esparsa evoluindo para uma organização em aglomerados ou paliçadas. A colônia exibe uma coloração creme alaranjado, com aparência seca e opaca e crescimento concêntrico. Em meio líquido, desenvolve-se como depósitos granulares com uma película superficial (Dorella *et al.*, 2006).

Quanto às suas propriedades bioquímicas, existem variações principalmente no que diz respeito à sua habilidade de fermentação. As linhagens de *C. pseudotuberculosis* tem a capacidade de produzir ácido, sem produzir gás, a partir de uma variedade de fontes de carbono, incluindo glicose, frutose, maltose, manose e sacarose (Merchant & Packer, 1967; Buxton & Fraser, 1977; Holt *et al.*, 1994). Quanto aos testes bioquímicos, *C. pseudotuberculosis* é fosfolipase D (PLD) positiva, catalase positiva, oxidase negativa e beta-hemolítica (Buxton & Fraser, 1977). A produção de nitrato redutase pode ser utilizada

para distinguir entre seus dois biovares, *equi* (redução de nitrato positiva - isolado de equinos e bovinos) e *ovis* (redução de nitrato negativa – isolado de ovinos e caprinos) (Biberstein *et al.*, 1971).

## **1.2 A Linfadenite Caseosa**

### **1.2.1 Características gerais**

A LC é uma doença infecto-contagiosa que se caracteriza principalmente pelo aparecimento de granulomas que evoluem para necrose nos linfonodos superficiais e internos podendo atingir algumas vísceras, como pulmões e fígado (Ayers, 1977; Menzies, 1998; Arsenault *et al.*, 2003). A forma mais frequente da doença é a externa, caracterizada pela formação de abscessos nos linfonodos superficiais e tecidos subcutâneos. Porém, os abscessos podem se formar também em órgãos internos como pulmão, rim, fígado e baço (Dorella *et al.*, 2006<sup>a</sup>; Guimarães *et al.*, 2009).

### **1.2.2 Epidemiologia**

A LC encontra-se distribuída principalmente em países com extensas criações de ovinos e caprinos, como Austrália, Nova Zelândia, África do Sul, Brasil, Canadá e Estados Unidos (Williamson, 2001; Arsenault *et al.*, 2003; Paton *et al.*, 2003; Dorella *et al.*, 2006; Guimarães *et al.*, 2009). Esta doença leva a perdas econômicas em áreas de criação devido a fatores como capacidade reprodutiva reduzida, redução de peso, diminuição da produção de leite e lã, desvalorização da pele do animal e condenação da carcaça e couro.

No Brasil, o estado da Bahia é o que apresenta o maior rebanho nacional de caprinos. Nesta região, o prejuízo causado pela incidência da LC chega a aproximadamente 4 milhões de reais ao ano (Oliveira, 2007).

A incidência da doença em rebanhos caprinos e ovinos apresenta na literatura dados discrepantes havendo diferenças na incidência percentual quando se estuda animais de abatedouros ou animais de criação extensiva. Um estudo com caprinos do semiárido brasileiro realizado por Unanian e colaboradores (1985) revelou incidência de 12% avaliando-se abscessos internos, e de 42% quando avaliados abscessos superficiais. A incidência de geral de abscessos internos e externos foi de 28.8%.

Um estudo de Meyer (2003), no qual foi realizada uma investigação sorológica na zona semiárida do estado da Bahia, revelou variação na prevalência de LC com variação de índices de 9,2% a 72,2%.

Já no estado de Minas Gerais, o rebanho é reduzido, no entanto, 84,3% dos produtores relataram já terem sofrido consequências da LC em seus rebanhos. Recentes estudos buscando avaliar a prevalência de LC em Minas Gerais a partir de testes sorológicos com uma amostra constituída por 205 propriedades revelaram soropositividade de 78,9% em caprinos e 75,8% em ovinos (Guimarães *et al.*, 2009; Seyffert *et al.*, 2010; Guimarães *et al.*, 2011).

### **1.2.3 Transmissão**

A transmissão de *C. pseudotuberculosis* entre os animais se dá principalmente devido a ferimentos superficiais causados por procedimentos do manejo ou eventos traumáticos que causem injúria ao corpo do animal e possibilitem contato com as secreções e/ou instrumentos contaminados com material caseoso de linfonodos, ou ainda pela ingestão de alimentos contaminados (Alves *et al.*, 1997; Williamson, 2001; Dorella *et al.*, 2006). *C. pseudotuberculosis* pode infectar diferentes animais como camelos, lhamas, búfalos, cabras, ovelhas, cavalos, alpacas, antílopes e primatas (Peel *et al.*, 1977; Baird & Fontaine, 2007; Selim, 2001). A infecção de humanos por *C. pseudotuberculosis* é um evento raro, porém demonstra seu potencial zoonótico estando geralmente relacionado a atividades ocupacionais (Liu *et al.*, 2005; Peel *et al.*, 1977). A sintomatologia em humanos envolve principalmente linfadenopatia (Peel *et al.*, 1997).

### **1.2.4 Patogênese**

Após a invasão do hospedeiro, *C. pseudotuberculosis* é fagocitada, porém, é resistente ao fagolisossomo e continua a se multiplicar levando a lise destas células de defesa e liberação das bactérias (Bogdan *et al.*, 1997). A resistência ao fagolisossomo pode ser devida, entre outros fatores, a lipídeos de superfície, que estão envolvidos na ação patogênica deste microrganismo (Hard, 1972). Além disso, a camada lipídica fornece uma proteção contra a ação de enzimas degradativas do fagolisossomo das células fagocíticas, permite melhor aderência das bactérias e causa uma toxicidade local pela indução de necrose hemorrágica (Alves & Pinheiro, 1997). Com a propagação da *C. pseudotuberculosis* e a morte das células do sistema imune, o microrganismo compromete o linfonodo mais próximo causando necrose (Jones *et al.*, 2000). A bactéria pode então replicar-se, colonizar os linfonodos e atingir outros órgãos e células. A fosfolipase D (PLD) também está envolvida na disseminação de *C. pseudotuberculosis* e é considerada um importante fator de virulência desta bactéria (Airello *et al.*, 2001). O papel da toxina PLD em *C. pseudotuberculosis* será discutido adiante neste trabalho.

No intuito do hospedeiro conter e eliminar a *C. pseudotuberculosis*, ocorre formação de granulomas. Estes são formados pelo aparecimento de necrose caseosa em células epitelióides que posteriormente são flanqueados por células epiteliais, linfócitos e tecido conjuntivo fibroso. Isto ocorre sucessivas vezes sobre o tecido conjuntivo fibroso, resultando no nódulo final formado por uma massa esférica laminada com aspecto de cebola, característico da LC (Jones *et al.*, 2000). Estes granulomas são formados devido à resposta imune adaptativa, que pode ser tanto humoral como celular (Paule, 2003). Estudos de Walker e colaboradores (1991) demonstraram a presença de células CD4 e CD8 distribuídas no tecido linfóide e alta concentração de células T CD8+ no interior das lesões de linfadenite caseosa. Estas características podem contribuir para um mecanismo antibacteriano impedindo a disseminação (Modlin *et al.*, 1988). Em lesões iniciais, células T CD4+ são predominantes, enquanto que em lesões maduras predominam células T CD8+ (Pepin *et al.*, 1994).

As citocinas IFN- $\gamma$  e TNF- $\alpha$  estão envolvidas na resposta a *C. pseudotuberculosis* (Simmons *et al.*, 1997). Elas combatem a infecção primária estando envolvidas no mecanismo de resistência do hospedeiro. Já na infecção secundária, podem estar envolvidas na ativação de macrófagos levando a uma imunidade protetora (Denis, 1991; Nakane *et al.*, 1991). As células Th1 podem estar envolvidas no desenvolvimento de resistência do hospedeiro, como sugerido pelo trabalho de Lan e colaboradores (1999), no qual não foi detectada IL-4 após uma infecção secundária. Há uma mudança do perfil de citocinas durante a infecção, e, além disso, há uma alta produção de anticorpos contra antígenos bacterianos como, por exemplo, a PLD (Dorella *et al.*, 2009).

### 1.2.5 Diagnóstico

O diagnóstico de LC é baseado em achados clínico-epidemiológicos apoiados por exames laboratoriais (Aleman & Spier, 2001). Na rotina, o diagnóstico é firmado a partir de isolamento do microrganismo dos abscessos via punção aspirativa ou extirpação cirúrgica (Ribeiro *et al.*, 2001; Nokazi *et al.*, 2000; Williamson, 2001). O microrganismo isolado é cultivado e analisado via coloração de Gram, Giemsa e Panóptico e passa por testes bioquímicos, como o sistema API Coryne (API-bioMérieux, Inc., La Balme lês Grottes, France), que consiste de uma bateria de 21 testes bioquímicos (Dorella *et al.*, 2006). Porém, estas abordagens não permitem o diagnóstico sub-clínico da doença, o que seria de grande importância para o seu controle.

Na tentativa de estabelecer um diagnóstico sub-clínico, abordagens sorológicas foram implementadas. Estes testes foram baseados na soroneutralização (Carne, 1940), soro-aglutinação (Cameron & McOmie, 1940; Awad, 1960; Kestintepe, 1976; Burrel 1978;



Shigidi, 1979; Lund *et al.*, 1982), imunodifusão em gel (Shigidi, 1974; Burrel, 1980b; Nain *et al.*, 1984), hemaglutinação indireta passiva (Shigidi, 1978), fixação do complemento (Shigidi, 1979), inibição da hemólise em tubo (Burrel, 1980).

Posteriormente, foram desenvolvidos testes de ELISA para o diagnóstico da LC com o objetivo de aumentar a sensibilidade do diagnóstico (Shen *et al.*, 1982; Johnson *et al.*, 1987; Binns *et al.*, 2007; Carminatti *et al.*, 2003; Dercksen *et al.*, 2000; Paule *et al.*, 2003). Porém, as abordagens sorológicas são inviabilizadas na prática, ora por sua baixa sensibilidade ou por terem seus resultados afetados pela vacinação dos animais ou similaridade antigênica das corinebactérias, ora pelo elevado custo (Dorella *et al.*, 2006).

As estratégias moleculares, como a PCR, também foram desenvolvidas com a finalidade de aumentar a sensibilidade e especificidade do diagnóstico (Cetinkaya *et al.*, 2002; Pavan *et al.*, 2012). Porém, os testes moleculares apresentam como limitação o elevado custo de implantação e, em alguns casos, reações cruzadas entre espécies relacionadas (Aleman *et al.*, 1996; Baird & Fontaine, 2007). Pesquisas do nosso grupo levaram ao desenvolvimento de um ensaio de PCR multiplex baseado na amplificação dos genes 16SrRNA, *rpoB* e *pld*. Este ensaio tem como vantagem a possibilidade de utilização de DNA extraído de amostras clínicas, boa especificidade e sensibilidade, porém não detecta *C. pseudotuberculosis* a partir de amostras sorológicas (Pacheco *et al.*, 2007).

Testes intradérmicos de hipersensibilidade (TIH) também são estudados para o diagnóstico da LC devido à sua praticidade e possibilidade de diagnóstico subclínico no campo. Os primeiros testes foram elaborados a partir de filtrados de cultura de *C. pseudotuberculosis* (Cesari, 1930). Posteriormente, foi utilizada uma pseudotuberculina baseada na tuberculina de Koch obtendo-se baixa especificidade (Carne, 1932). Foi também utilizada suspensão de cultura bacteriana de *C. pseudotuberculosis* e *Corynebacterium pyogenes* (Costa Filho, 1978); bactérias fragmentadas por sonicação (Renshaw *et al.*, 1979; Brown *et al.*, 1986); proteína hidrossolúvel de *C. pseudotuberculosis* (Langenegger *et al.*, 1987). Porém, ainda não foi alcançada especificidade, sensibilidade e praticidade satisfatórias. Em um trabalho do nosso grupo de pesquisa, Antunes (2011), utilizando a técnica TPP (do inglês, *Three-phase partitioning*), obteve as exoproteínas totais de *C. pseudotuberculosis* e as utilizou em TIH obtendo bons resultados, porém o teste ainda precisa de otimizações.

### **1.2.6 Tratamento, profilaxia e vacina**

Antibióticos sistêmicos são empregados no tratamento da linfadenite caseosa. Porém, apesar de *C. pseudotuberculosis* ser sensível a antibióticos *in vitro*, o tratamento não tem a mesma eficácia *in vivo*, uma vez que o antibiótico não é capaz de penetrar a cápsula

dos abscessos. Dessa forma, uma alternativa é a drenagem e extirpação dos linfonodos superficiais acometidos. Porém, esta prática pode contribuir para a disseminação da doença, uma vez que leva à exposição dos microrganismos presentes no abscesso ao ambiente e instrumentos sendo que a quantidade de microrganismos liberados por um abscesso é capaz de contaminar todo o rebanho. Além disso, esse método não é capaz de eliminar totalmente os microrganismos além de não ser aplicável quando linfonodos e órgãos internos já estão acometidos (Williamson, 2001; Guimarães *et al.*, 2009).

Dessa forma, as estratégias de prevenção e controle são as mais eficientes contra a LC. A principal medida profilática contra a LC é a identificação e remoção dos animais infectados (Binns *et al.*, 2001). Além disso, deve-se ter o cuidado de higienização de baias e outros equipamentos que possam ter contato com os animais (Williamson, 2001; Dorella *et al.*, 2006). Outra medida é evitar que estábulos e canchais possam causar, de alguma forma, lesões cutâneas nos animais (Valli & Parry, 2007). Porém, essas medidas encontram dificuldades na prática do manejo dos rebanhos e faz com que a medida ideal para o controle da LC seja a prevenção através da vacinação (Dorella *et al.*, 2006). Várias abordagens vêm sendo estudadas no intuito de se alcançar o desenvolvimento de uma vacina com eficácia protetora. Entre essas abordagens, encontra-se o uso de bactérias vivas atenuadas, bactéria mortas, frações de parede celular, sobrenadante de culturas bacterianas, componentes celulares combinados a sobrenadante de culturas e vacinas de DNA (LeaMaster *et al.*, 1987; Holstald, 1989; Ellis *et al.*, 1991; Eggleton *et al.*, 1991; Costa *et al.*, 2011).

Atualmente, estão disponíveis no mercado vacinas comerciais contra a linfadenite caseosa. A maioria delas são combinadas a vacinas contra outros patógenos como *Clostridium tetani*, *Cl. perfringens*, *Cl. septicum*, *Cl. novyi* e *Cl. chauvoei* (Paton *et al.*, 1995; Piontkowski *et al.*, 1998; Stanford *et al.*, 1998; Williamson, 2001). Essas formulações são denominadas vacinas de toxóide por serem baseadas em formas inativadas da toxina PLD, importante fator de virulência que apresenta também capacidade de estimulação do sistema imune do hospedeiro, sendo por isso muito pesquisada como alvo vacinal (Dorella *et al.*, 2006). Um estudo de Paton e colaboradores (2003) demonstrou que embora 43% dos criadores estudados fizessem uso de vacinas comerciais contra a LC, apenas 12% deles o faziam de forma correta. Além disso, algumas vacinas apresentam eficácia diferenciada entre caprinos e ovinos, o que requer que o esquema de vacinação seja ajustado de acordo com o rebanho (Williamson, 2001).

Uma vacina viva atenuada utilizando a linhagem 1002 de *Corynebacterium pseudotuberculosis* em suspensão foi desenvolvida pela Empresa Baiana de Desenvolvimento Agrícola (EBDA). A linhagem 1002 apresenta baixa virulência e foi isolada pela EBDA em 1971 de um abscesso de linfonodo superficial pré-escapular de um caprino

infectado na Bahia. Em testes experimentais, esta vacina apresentou 83% de proteção em caprinos. É licenciada para o uso no Brasil desde 2002. Pode ser aplicada em animais jovens por via subcutânea, mas não é capaz de atingir os resultados satisfatórios com a aplicação de uma única dose e requer a aplicação de dose reforço. Por se tratar de uma vacina viva, requer armazenamento a baixas temperaturas e tem viabilidade por tempo limitado. No campo, esta vacina apresentou proteção menor que os 83% obtidos experimentalmente (Dorella *et al.*, 2006; EBDA-Empresa Baiana de Desenvolvimento Agrícola S/A (BR/BA), 1999). Uma versão liofilizada desta vacina também foi estudada por Meyer e colaboradores (2002) em parceria com a EBDA. Os resultados foram satisfatórios necessitando, porém, de doses maiores para se obter resultados similares à formulação em suspensão.

Uma alternativa para a geração de linhagens vivas atenuadas é a atenuação utilizando engenharia genética. Estudo de Dorella *et al.*, 2006, através de mutagênese aleatória por transposição baseado no TNFuZ, identificou 34 mutantes, dentre eles o mutante CP13 que conferiu proteção de 81% em modelo murino. No entanto, ainda são necessários estudos acerca deste mutante a fim de caracterizar o tipo de resposta imune e proteção observada. Estudos também devem ser realizados em caprinos e ovinos a fim de verificar se o nível de proteção será mantido nos animais alvo da doença.

Diante dos fatores expostos acerca da problemática de controle da LC, não havendo diagnóstico subclínico, tratamento ou vacina que combine nível adequado de proteção e praticidade de utilização, evidencia-se a necessidade de estudos que busquem a identificação precoce dos animais infectados para seu isolamento através de testes diagnósticos mais sensíveis e precisos e o desenvolvimento drogas e vacinas mais eficazes a fim de impedir a introdução permanência da LC no rebanho.

### **1.3 Fatores de virulência de *C. pseudotuberculosis***

Os fatores de virulência desempenham papel importante na adesão, invasão, colonização, disseminação no hospedeiro e evasão da resposta imune do hospedeiro frente ao patógeno (Dorella *et al.*, 2006). Para o desenvolvimento de estratégias eficazes para o controle da LC, faz-se necessária a identificação e estudo de fatores envolvidos na virulência de *C. pseudotuberculosis* que ajudem a entender melhor a biologia deste microrganismo e os mecanismos de interação parasito-hospedeiro. Estes fatores podem servir como alvos vacinais, alvos de drogas, componentes de testes-diagnóstico ou alvos de mutações para o desenvolvimento de linhagens atenuadas.

Os lipídeos de superfície de *C. pseudotuberculosis* são fatores que contribuem para a patogênese de *C. pseudotuberculosis*. Um estudo de Muckle e Gyles (1982) demonstrou

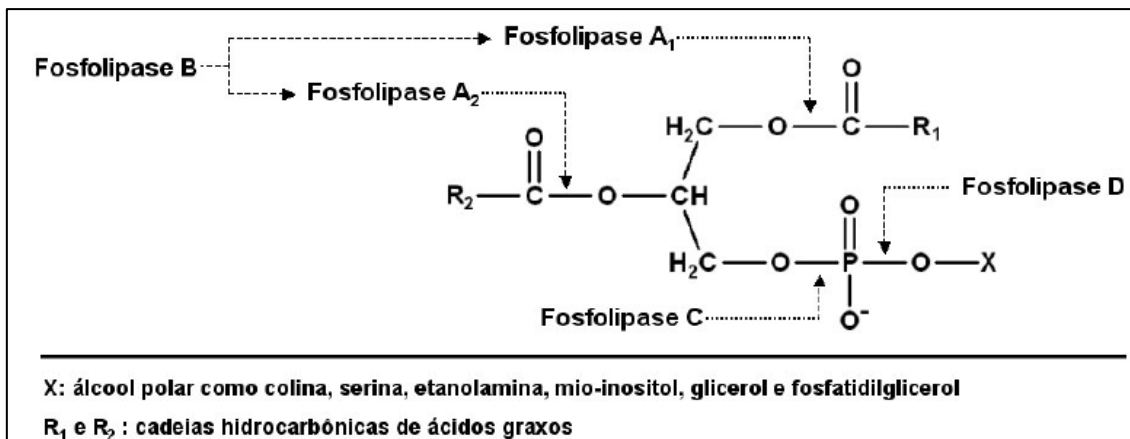
haver uma relação de proporção direta entre a quantidade de lipídeos de superfície e a formação de abscessos crônicos. A toxicidade dos lipídeos extraídos de *C. pseudotuberculosis* foi demonstrada pela indução de necrose hemorrágica após injeção intradérmica em cobaias (Jolly, 1966). Os lipídeos de superfície também constituem uma barreira física contra a defesa do hospedeiro. Além disso, estão envolvidos na aderência do microrganismo e promovem citotoxicidade local (Hard, 1969; Alves *et al.*, 1977).

Genes envolvidos na captação de ferro também foram apontados como importantes para a virulência de *C. pseudotuberculosis*. Foi identificado um *cluster* gênico constituído por quatro genes organizados em operon, (*fag*) *A*, *B*, *C* e *D*. Estes genes são importantes por possibilitarem a aquisição de ferro em ambientes hostis, como o ambiente intracelular, onde este nutriente é escasso. Mutante *fag B(C)*, no qual há uma interrupção no gene *fag (B)* e bloqueio da expressão de *fag(B)* e *fag(C)*, manteve sua capacidade de utilização de ferro *in vitro*, porém apresentou redução na sobrevivência e capacidade de formação de abscessos em animais infectados experimentalmente (Billington *et al.*, 2002).

A fosfolipase D (PLD) é considerada o principal fator de virulência de *C. pseudotuberculosis*. A PLD está envolvida com a persistência e disseminação da bactéria no hospedeiro, o que é conferido pela sua capacidade de hidrolisar ligações éster de esfingomiélna na membrana das células do hospedeiro. Além disso, PLD apresenta atividade citotóxica, dermonecrótica e leucotóxica (Carne & Onon, 1978; Coyle & Lipsky, 1982; Tashijan Campbell, 1983; McManara *et al.*, 1994; McNamara *et al.*, 1995; Songer, 1997). A seguir, serão apresentadas informações relacionadas às fosfolipases, em especial, a PLD.

### 1.3.1.1 Fosfolipases

As fosfolipases constituem uma superfamília de esterase com a função de hidrólise de fosfolipídios em ligações éster específicas (Kini, 1997). Essas enzimas são encontradas em organismos diversos como vírus, bactérias, fungos, plantas e animais (Exton, 1990; Jenkins & Frohman, 2005; Wang & Kolko, 2010; Richmond & Smith, 2011). Nesses diferentes organismos, as fosfolipases podem desempenhar importantes papéis de relevância fisiológica: enzimas digestivas, como fosfolipases presentes em veneno de serpentes; manutenção e remodelamento de membranas celulares; sinalização celular através da geração de lipídeos bioativos (Richmond & Smith, 2011). As fosfolipases são classificadas, de acordo com a ligação clivada em seu substrato, em cinco famílias: fosfolipases A<sub>1</sub> (PLA<sub>1</sub>), fosfolipases A<sub>2</sub> (PLA<sub>2</sub>), fosfolipases B (PLB), fosfolipases C (PLC) e fosfolipases D (PLD) (Figura 1) (Kini, 1997).



**Figura 1:** Estrutura de um glicerofosfolípido e o sítio de ação das fosfolipases. \*Adaptado de Kini, 1997.

As PLA<sub>1</sub> catalisam a clivagem de ácidos graxos na posição sn-1 de fosfolípídeos resultando na formação de 2-acilfosfolípídeos e ácido graxo livre. São encontradas em diversos tecidos de organismos além de apresentarem interesses industriais devido à aplicação comercial de 2-acilfosfolípídeos gerados. As isozimas de PLA<sub>1</sub> são divididas em dois grupos de acordo com a sua localização celular, o grupo 1 compreende as isozimas intracelulares e o grupo 2 compreende as extracelulares (Song & Rhee, 2001; Richmond & Smith, 2011).

As PLA<sub>2</sub> catalisam a hidrólise na posição sn-2 de fosfolípídeos de membrana levando à formação de 1-acilfosfolípídeos e ácidos graxos livres. São enzimas bem conservadas e fazem parte de uma superfamília que se subdivide em 11 grupos. Quanto à localização celular, as PLA<sub>2</sub> podem ser secretadas, citosólicas e lisossomal. Importantes compostos são gerados por intermédio das PLA<sub>2</sub> como o ácido araquidônico, ácido lisofosfatídico e lisofosfatidilcolina (Six & Dennis, 2000; Kudo, 2004). Também podem estar envolvidas na apoptose e homeostasia de Ca<sup>2+</sup> (Murakami, 2004).

Fosfolipases B (PLB) hidrolisam ligações nas posições sn-1 e sn-2 tendo ação fosfolípídica, lisofosfolípídica e transacídica. São consideradas fatores de virulência em alguns organismos patogênicos, porém ainda não se sabe qual de suas atividades são importantes para a patogenicidade (Shen *et al.*, 2004; Ganendren *et al.*, 2006)

As PLCs são responsáveis pela clivagem de fosfatidilinositóis de membrana para produzir os intermediários intracelulares 1,2-diacilglicerol e inositol 1,4,5-trifosfato. Dessa forma, as PLCs estão envolvidas em diversos mecanismos de sinalização celular em mamíferos influenciando na ação de neurotransmissores, hormônios, fatores de crescimento, entre outros (Rebecchi & Pentylala, 2000).

PLD hidrolisa ligações fosfodiéster de fosfolípídeos levando a formação de ácido fosfatídico e um aminoálcool. PLD também pode catalisar reações de transfosfatidilação pela substituição da cadeia hidroxilada na presença de um álcool acceptor apropriado formando um novo fosfatidil-álcool (Ulbrich-Hofmann *et al.*, 2005;). PLDs são encontradas em vírus, bactérias, fungos, plantas e animais. Em mamíferos, estão envolvidas em regulação do citoesqueleto, sinalização celular e tráfego de vesículas (Jenkins & Frohman, 2005).

#### **1.3.1.2 Fosfolipase D em *C. pseudotuberculosis***

Em linhagens de *C. pseudotuberculosis*, a PLD é encontrada com massa molecular de 31,2 kDa. Está presente em sobrenadante de cultura de *C. pseudotuberculosis*, sendo uma proteína secretada, e é reconhecida por anticorpos de soro de animais naturalmente infectados (Linder & Bernheimer, 1978; Songer *et al.*, 1990). Possui uma sequência correspondente ao peptídeo sinal anterior à sequência da proteína madura (Hodgson *et al.*, 1990). É uma toxina que enfraquece a membrana celular de células do hospedeiro por quebrar a esfingomiélin presente nas membranas celulares em colina livre e ceramida-fosfato (Carne & Onon, 1978). Dessa forma, a PLD favorece a disseminação do microrganismo (McNamara, 1994). Conforme descrito por Baird e Fontaine (2007), a PLD atua sobre o endotélio vascular causando um aumento da permeabilidade o que resulta na disseminação do microrganismo até os linfonodos causando necrose dérmica e agregação plaquetária podendo levar à falência renal. A enzima induz ainda alteração na membrana de eritrócitos, formação de vesículas (Brodgen *et al.*, 1990), redução de viabilidade e quimiotaxia de macrófagos no local da infecção (Yozwiak & Songer, 1993).

Os anticorpos contra a PLD podem ser detectados já nos 30 primeiros dias após a infecção por *western blot* tendo a intensidade de reação aumentada após o 6º mês de infecção (Vale *et al.*, 2003). Por ser capaz de funcionar como imunógeno e por ser considerada o principal fator de virulência de *C. pseudotuberculosis*, a PLD é bastante estudada para o seu uso em vacinas.

A letalidade da PLD foi comprovada por um estudo de Hsu e colaboradores (1985), no qual a injeção parenteral em pequenos ruminantes de PLD purificada ou combinada levou à morte destes animais em 48 horas. Com o objetivo de produzir uma PLD enzimaticamente inativa, Haynes e colaboradores (1992) conduziram um estudo no qual foi induzida mutagênese no gene *pld* pela exposição ao ácido fórmico e então foi induzida a expressão em *E. coli*. Um dos clones gerados produziu PLD enzimaticamente inativa (toxóide), porém com a antigenicidade conservada. A análise da sequência desta proteína

revelou que a mutação foi uma troca de uma única base levando à substituição do códon de His<sup>20</sup> por Tyr, sugerindo que o aminoácido His<sup>20</sup> faz parte do sítio ativo da PLD.

Hodgson e colaboradores (1992), construíram um mutante a partir da deleção do gene *pld* em *C. pseudotuberculosis* (Toxminus). Este estudo foi realizado com o objetivo de gerar um mutante atenuado que pudesse ser utilizado como vacina. Esta linhagem avirulenta foi utilizada em experimentos de vacinação de ovelhas por via subcutânea e levou a uma diminuição do número e severidade dos abscessos, quando comparado ao grupo de animais não vacinados, sugerindo uma resposta protetora. Porém, apesar da linhagem Toxminus estimular resposta humoral e celular, o nível dessa resposta foi menor que o induzido pela linhagem selvagem não atingindo eficácia protetora. Estes resultados evidenciam que outros antígenos além da PLD estão envolvidos na proteção contra LC. O estudo evidenciou também o papel da PLD como fator de virulência, uma vez que a linhagem Toxminus não é capaz de levar ao desenvolvimento da LC. Os resultados obtidos com a linhagem Toxminus indicam seu potencial para ser utilizada como veículo de entrega de outros antígenos a fim de melhorar a proteção conferida (Hodgson *et al.*, 1992).

Formas atóxicas da PLD totalmente ou parcialmente purificadas ofereceram proteção significativa contra LC em ovinos. Porém, estas formulações não são administradas em dose única (Burrell, 1983; Eggleton *et al.*, 1991a; Eggleton *et al.*, 1991b). No estudo de Hodgson e colaboradores (1994), a utilização da linhagem Toxminus como veículo expressando uma forma atóxica da proteína PLD foi vista como alternativa para o desenvolvimento de uma vacina que levasse a um nível de proteção adequado sendo administrada em uma dose única. Para isso, a linhagem Toxminus foi transformada com o plasmídeo pTB253 carregando o gene codificador da PLD atóxica desenvolvida por Haynes e colaboradores (1992), a qual apresenta uma substituição de His<sup>20</sup> por Tyr, e utilizada para imunização de ovinos por via oral. Foi observada diminuição no nível de secreção da PLD mutante. Não houve resposta imune humoral satisfatória com a utilização do toxóide combinado à linhagem Toxminus. A linhagem Toxminus sozinha quando administrada por via oral também não alcançou eficácia protetora.

Mutações no sítio catalítico da PLD podem causar diminuição na capacidade de secreção da proteína (Tachedijan *et al.*, 1995). Hodgson e colaboradores (1999) utilizaram uma estratégia de mutagênese dirigida para o desenvolvimento de uma forma atóxica da PLD na tentativa de melhorar a atividade e nível de secreção da proteína mutada. A PLD foi modificada pela substituição de His<sup>20</sup> por Ser. O índice de proteção alcançado foi inferior ao alcançado com o toxóide clássico inativado por ácido fórmico.

A PLD foi utilizada também como vacina de DNA. De Rose e colaboradores (2002) inocularam um plasmídeo contendo uma forma inativada da PLD combinada a Ig de CTLA-4 em ovinos por via intramuscular, injeção subcutânea e *gene gun*. O melhor resultado foi

obtido pela administração por via intramuscular atingindo a proteção em 45% dos animais desafiados. Houve uma elevação dos níveis séricos de IgG 2 nos animais imunizados, o que sugere uma resposta do tipo Th1 levando a uma imunidade protetora contra *C. pseudotuberculosis*.

A PLD é vista, então, como um potencial adjuvante em novas formulações vacinais. Além disso, sua estrutura, atividade enzimática e mecanismo de ação no hospedeiro precisam ainda ser melhor estudados.

### **1.3.2 Proteínas potencialmente envolvidas na virulência de *C. pseudotuberculosis***

Nosso grupo de pesquisa realizou o sequenciamento pioneiro do genoma da linhagem *C. pseudotuberculosis* FRC41, isolada de linfonodo inguinal de uma paciente francesa de 12 anos com linfadenite necrosante. A partir da análise do genoma, utilizando abordagens *in silico* juntamente com o conhecimento prévio a partir de dados experimentais, alguns genes possivelmente envolvidos na virulência de *C. pseudotuberculosis* foram identificados (Trost *et al.*, 2010). Além do gene *pld*, foram sugeridos outros, como o gene *spaC*, que codifica a proteína SpaC, componente de uma adesina, sendo as adesinas bacterianas consideradas importante fator de virulência (Petri *et al.*, 2010; Trost *et al.*, 2010). Outro gene identificado foi *sodC* que codifica a superóxido-dismutase dependente de cobre e zinco (Trost *et al.*, 2010). As superóxido-dismutases estão envolvidas nos mecanismos de proteção contra a geração de intermediários reativos de oxigênio e nitrogênio, que por sua vez apresentam atividade antimicrobiana. Essas enzimas atuam na conversão de ânions superóxido em oxigênio molecular e H<sub>2</sub>O<sub>2</sub>, o qual, por sua vez, é quebrado em H<sub>2</sub>O pela enzima catalase (Fridovich, 1995). O gene *nanH* (neuraminidase H) também foi ressaltado como codificador de um provável fator de virulência. NanH é uma neuraminidase extracelular que pode estar envolvida com o reconhecimento de ácido siálico na superfície de células do hospedeiro (Trost *et al.*, 2010; Vimr *et al.*, 2004).

Corrêa (2014) realizou um estudo no qual foram comparados os níveis transcricionais dos genes *pld*, *spaC*, *sodC* e *nanH* em meio de cultura e em amostras extraída de abscessos caseosos através de qRT-PCR. Além disso, a expressão do gene *pkng* foi investigada em cultura de macrófagos de camundongos das linhagens Balb/c e C57BL/6. Os genes *pld*, *sodC*, *spaC* e *nanH* tiveram os níveis transcricionais aumentados de 6 a 8 vezes nos abscessos se comparados às culturas *in vitro*. O gene *pkng* apresentou maior magnitude de expressão em cultura de macrófagos na primeira hora após infecção.

Além dos fatores de virulência levantados por estudos comparativos a partir da linhagem FRC41 (Trost *et al.*, 2010), uma classe de moléculas que merece atenção são as quinases bacterianas. Essas moléculas fazem parte de importantes mecanismos de



sinalização celular e podem estar envolvidas, entre outras coisas, na modulação da expressão de fatores de virulência. Uma dessas quinases bacterianas é a PknG, uma serina-treonina proteína quinase, que é bastante conservada em actinomicetos, como *Mycobacterium* e *Corynebacterium* (Niebish, 2006; Corrêa, 2014). Em micobactérias, PknG parece estar envolvida com a maturação do fagolisossomo, além de influenciar o metabolismo de glutamina em *C. glutamicum* e *Mycobacterium tuberculosis* (O'Hare *et al.*, 2008; Cowley *et al.*, 2004; Walburguer *et al.*, 2004).

### 1.3.2.1 Proteína Quinase G – PknG

A fosforilação reversível de proteínas é a modificação pós-traducional mais importante na transdução de sinais. Os sistemas de transdução de sinais, por sua vez, estão envolvidos nos mecanismos celulares de monitoramento, resposta, adaptação e sobrevivência em diferentes ambientes (Raggiaschi *et al.*, 2005). Estes mecanismos mediados por fosforilação protéica estão presentes tanto em eucariotos como em procariotos. Os eventos de fosforilação são mediados por proteínas denominadas quinases. As proteínas quinases catalisam a fosforilação através da transferência de um grupo fosforila de ATP ou GTP (menos frequentemente), para treonina ou serina (Ser/Thr-quinases) ou resíduos de tirosina (Tyr-quinases) (Burnside & Rajagopal, 2011).

As quinases são chamadas de “interruptores” moleculares que existem no estado “on”- ativado ou “off” – inativado (Doran *et al.*, 2003). A transição de um estado a outro é controlada por diversos mecanismos que incluem a ligação de fatores alostéricos e localização celular (Burnside & Rajagopal, 2011). A fosforilação de proteínas também está associada à virulência de diversos microrganismos patogênicos, uma vez que essas proteínas fosforiláveis encontram diferentes substratos na célula do hospedeiro (Cozzone, 2005; Whitmore & Lamont, 2012; Canova & Molle, 2014; Fujita *et al.*, 2014).

Inúmeras serina-treonina quinases (*Serine-Threonine Protein Kinase* - STPKs) e suas respectivas fosfatases (*Serine-Threonine Phosphatase* - STPs) foram identificadas em procariotos, entre eles: *Myxococcus xanthus*, *Streptomyces spp*, *Anabaena*, *Cyanobacterium*, *Bacillus spp*, *Streptococcus spp*, *Mycobacterium spp*, *Yersinia spp*, *Listeria monocytogenes*, *Pseudomonas aeruginosa*, *Enterococcus faecalis* e *Staphylococcus aureus* (Burnside & Rajagopal, 2011).

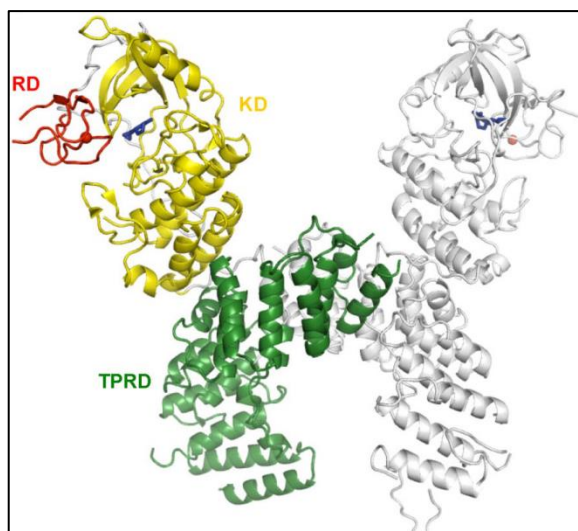
A compreensão das consequências da fosforilação de serina/treonina em patógenos Gram-positivos é importante para o desenvolvimento de drogas, bem como avaliação de seu potencial como alvo vacinal. Imunização com STPK de *Streptococcus pneumoniae*, por exemplo, conferiu proteção em modelo murino contra sepse e pneumonia (Burnside & Rajagopal, 2011; Giefing *et al.*, 2008; Giefing *et al.*, 2010)

As STPKs e STPs procarióticas apresentam em torno de 35% de identidade aminoacídica com relação às suas equivalentes eucarióticas. Procariotos contem, em geral, uma única cópia de STPK e STP no genoma e esses genes estão localizados em operon sendo, dessa forma, cotranscritos. A maioria das STPKs eucarióticas são associadas à membrana e as STPs correspondentes são citoplasmáticas. O domínio catalítico de STPKs está na região N-terminal, é intracelular e possui 11 subdomínios conservados (domínios de Hanks) que formam a estrutura catalítica conservada. A porção extracelular, por sua vez, apresenta alto grau de diversidade de sequência aminoacídica (Hanks *et al.*, 1988; Shi *et al.*, 1998; Burnside & Rajagopal, 2011).

Diferentes STPKs já foram estudadas em microrganismos pertencentes ao grupo CMNR, como *M. tuberculosis* e *C. glutamicum*.

No genoma de *M. tuberculosis* H37Rv foram identificados onze genes codificadores de STPKs e quatro codificadores de fosfatases (Av-Gay & Everett, 2000; Koul *et al.*, 2001; Prsic *et al.*, 2010). Uma dessas onze STPKs é a Proteína Quinase G (PknG), uma proteína de 82kDa solúvel que é secretada durante a infecção por *M. tuberculosis* no citosol de macrófagos infectados. Esta proteína contribui para a sobrevivência intracelular do microrganismo através do bloqueio da fusão fagossomo-lisossomo. Acredita-se que a secreção de PknG intra-macróforo se dê por vias alternativas, uma vez que esta proteína não apresenta sinal de exportação. Esta modulação da fusão fagossomo-lisossomo por PknG pode se dar via fosforilação de proteínas do hospedeiro após secreção de PknG no interior dos macrófagos (Walburger *et al.*, 2004). Dessa forma, PknG desempenha importante papel na virulência de *M. tuberculosis* sendo um alvo atrativo para o desenvolvimento de drogas (Scherr *et al.*, 2009). São necessários estudos a fim de entender os mecanismos pelos quais a PknG está envolvida na inibição da maturação do fagolisossomo e consequente inibição da eliminação do microrganismo pelo hospedeiro.

PknG de *M. tuberculosis* apresenta três domínios, sendo eles rubredoxina aminoterminal (RD), quinase (KD) e tetratricopeptídico (TPR) C-terminal (Figura 2). O domínio rubredoxina é essencial para a atividade de PknG, está envolvido no transporte de elétrons e controla a atividade quinase (Chao *et al.*, 2010). O domínio TPR está envolvido com interação proteína-proteína e montagem de complexos protéicos (Blatch & Lässle, 1999). No seu sítio de ligação a ATP, PknG possui um aminoácido lisina, característica típica de STPKs (Av-Gay & Everett, 2000).



**Figura 2.** Estrutura tridimensional da PknG de *Mycobacterium tuberculosis*.

RD – Domínio Rubredoxina; KD – Domínio Quinase; TPRD – Domínio Tetratricopeptídico.

\*Adaptado de Scherr *et al.*, 2007.

Estudos de Scherr e colaboradores (2009) demonstraram que a autofosforilação de resíduos de treonina N-terminais é essencial para impedir a fusão fagossomo-lisossomo para eliminação de *M. tuberculosis*. PknG é também a única quinase solúvel mantida no genoma de *M. smegmatis*, que acredita-se ter mantido o conjunto mínimo de genes para sua virulência (Cole *et al.*, 2001).

O gene codificador de PknG em micobactérias é o gene terminal de um operon que codifica *glnh*, supostamente envolvido com a captação de glutamina, o que indicaria que PknG também poderia estar envolvida no metabolismo deste aminoácido (Av-Gay & Everett, 2000). Cowley e colaboradores (2004), utilizando linhagem mutante *M. tuberculosis* deficiente para o gene *pknG*, demonstrou que a falta deste gene implica em redução da viabilidade celular *in vitro* e após infecção em camundongos Balb/C. O mutante também acumulou glutamato e glutamina e houve redução de 50% na síntese “*de novo*” de glutamina. Dessa forma, PknG foi implicada com a transferência de sinal durante o estresse nutricional em *M. tuberculosis* e adaptação metabólica.

O trabalho de Cowley e colaboradores (2004) indica que PknG não tem ação restrita ao modo de vida intracelular. Genes homólogos de PknG estão presentes em todos os demais membros da subordem *Corynebacterineae* com genomas sequenciados que incluem espécies do gênero *Corynebacterium*, *Mycobacterium*, *Nocardia*, *Rhodococcus* e *Streptomyces* (Niebisch *et al.*, 2006).

Em *C. glutamicum*, o gene *pknG* está situado em operon que inclui também os genes *glnX* e *glnH* codificadores de uma integrina hipotética de membrana e lipoproteína perioplasmática ligadora de glutamina, respectivamente (Kalinowski *et al.*, 2003).

Niebisch e colaboradores (2006) estudaram a função de PknG em *C. glutamicum*, um microrganismo não patogênico de interesse biotecnológico que tem sido utilizada para o entendimento de genes homólogos de *M. tuberculosis*. Neste estudo, demonstrou-se que PknG foi indispensável para utilização de glutamina em *C. glutamicum* uma vez que o mutante deficiente para o gene *pknG* ( $\Delta pknG$ ) apresentou redução drástica de crescimento em meio mínimo contendo glutamina como única fonte de carbono e nitrogênio. Também foram estudados mutantes deficientes em *glnx* e *glnh*. Em meio líquido mínimo contendo glutamina, o mutante  $\Delta glnx$  apresentou fenótipo semelhante ao mutante  $\Delta pknG$  não havendo crescimento nas primeiras 50h de incubação. Já o mutante  $\Delta glnh$  foi capaz de replicar em meio líquido mínimo contendo glutamina. Medidas da concentração intracelular de glutamato e glutamina indicaram que a deficiência de utilização de glutamina por  $\Delta pknG$  não é devido a dificuldade na captação de glutamina, mas sim devido a deficiência no catabolismo do glutamato. Neste mesmo estudo, análises proteômicas demonstraram que *odhI*, gene essencial para utilização de glutamina, é o substrato de PknG *in vivo*. OdhI inibe o complexo 2-oxoglutarato-desidrogenase (ODH) do ciclo do ácido cítrico (TCA). ODH catalisa a conversão NAD<sup>+</sup>- dependente de  $\alpha$ -cetoglutarato em succinil-CoA. Na ausência de PknG, OdhI é ativa e inibe ODH. A fosforilação por PknG inativa OdhI que passa a não inativar ODH (Chao *et al.*, 2010). OdhI é homóloga a GarA de *M. smegmatis* (Belanger & Hatfull, 1999). O mutante deficiente em *odhI* foi capaz de replicar em glutamina e uma deleção adicional de *odhI* sobre o mutante  $\Delta pknG$  restaurou a capacidade de crescimento em glutamina (Niebisch *et al.*, 2006).

O gene *pknG* foi identificado em *C. pseudotuberculosis* (Trost *et al.*, 2010), porém ainda não existem dados experimentais acerca da função deste gene em *C. pseudotuberculosis*. Considerando que PknG é uma proteína conservada em micobactérias e corinebactérias, é possível que esta proteína também esteja envolvida na capacidade de sobrevivência de *C. pseudotuberculosis* no macrófago do hospedeiro e/ou no metabolismo de glutamato. Estudos são necessários no sentido de entender a função deste gene em *C. pseudotuberculosis*, elucidar a estrutura da proteína a avaliar seu potencial como alvo vacinal ou no desenvolvimento de drogas contra a LC.

### 1.3.2.2 Adesina - SpaC

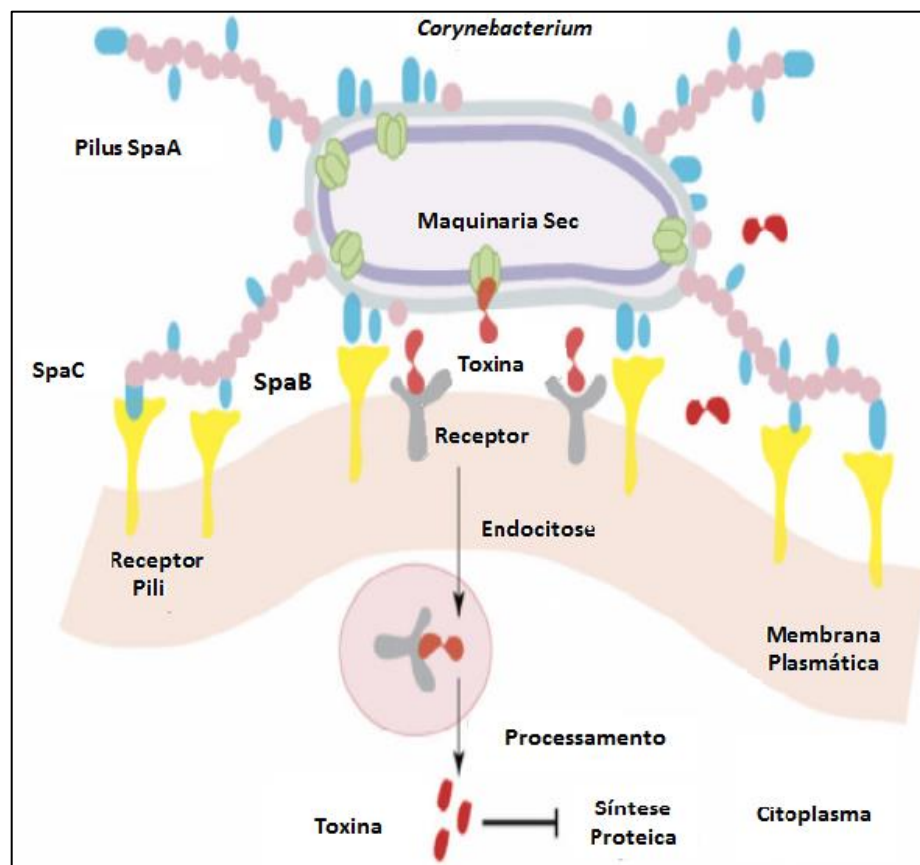
Adesão e invasão de tecidos específicos do hospedeiro são os primeiros estágios de uma infecção bacteriana (Foster & Hook, 1998). Estes processos acontecem por intermédio de filamentos proteicos presentes nas superfícies dos microrganismos, chamados de *pili* ou fímbrias, que possuem funções adesivas durante a colonização. Estes *pili* também desempenham papel na evasão do sistema imune e formação de biofilme (Mandlik *et al.*, 2008). A presença de *pili* já foi observada na superfície de várias espécies de corinebactérias (Kuamazawa & Yanagawa, 1972; Yanagawa & Honda, 1976).

Os *pili* de bactérias Gram-positivas são constituídos de polímeros de ligação cruzada ligados covalentemente a peptidoglicanos de parede celular (Ton-That & Schneewind, 2004; Mandlik *et al.*, 2008). O mecanismo de montagem destes *pili* requer a ação da sortase (Ton-That & Schneewind, 2004). A sortase é uma transpeptidase cisteína conservada em bactérias Gram-positivas (Mazmanian *et al.*, 1999; Ilangovan *et al.*, 2001; Comfort and Clubb, 2004; Dramsi *et al.*, 2005; Marraffini *et al.*, 2006). Esta transpeptidase cliva o precursor de pilina reconhecendo o motivo LPXTG presente no *cell wall sorting signal* (CWSS) de uma proteína precursora e clivando o polipeptídeo entre o aminoácido treonina e glicina. Então, a sortase liga a treonina N-terminal do polipeptídeo clivado ao peptideoglicano da parede celular (Mazmanian *et al.*, 2001; Ton-That & Schneewind, 2004). Esta reação leva à formação das fibras do *pilus* na superfície bacteriana (Ton-That & Schneewind, 2004).

Em *Corynebacterium diphtheriae* NCTC13129, o *pili* é composto por três subunidades de pilina (SpaA, SpaB e SpaC). Além das pilinas, são codificadas seis sortases (SrtA-F), das quais cinco estão envolvidas com a polimerização do *pilus*. Foram identificados três *clusters* gênicos cromossômicos em *C. diphtheriae* codificando nove proteínas de superfície carregando peptídeo sinal N-terminal e *sorting signal* C-terminal além de cinco genes de sortase (*srtA-E*). O gene *srt-F* está localizado em outra região do cromossomo. Entre estes *cluster* está o codificador do *pilus* tipo Spa-A (*spaA-srtA-spaB-spaC*). Os três tipos de *pilus* codificados por cada um dos três *clusters* apresentam estrutura semelhante composta por uma pilina principal e duas pilinas menores. O *pilus* tipo Spa-A, por exemplo, é formado por ligações covalentes cruzadas entre a pilina SpaA, formando a haste do *pilus*, juntamente com a pilina menor SpaC localizada na extremidade, e outra pilina menor, SpaB. SpaA é produzida em maior quantidade formando longos *pili* contendo SpaC na extremidade e SpaB interespaçada. As pilinas menores também são expostas na superfície celular. As pilinas SpaC são as que carregam o CWSS C-terminal. A polimerização de SpaA é catalisada pela sortase SrtA (Ton-That & Schneewind, 2003). As pilinas menores, SpaC e SpaB podem ser

chamadas de adesinas devido ao seu envolvimento com a aderência a células do hospedeiro.

Mandlik e colaboradores (2008) propõem um modelo de adesão e patogênese mediado por *pilus* baseado em *C. diphtheriae* que é demonstrado na Figura 3. Neste modelo, a expressão de adesinas contribui para o contato inicial do microrganismo com a célula do hospedeiro e, então, interações adicionais com as células do hospedeiro através de pilinas menores ancoradas a parede celular da bactérias formam uma zona de adesão entre as células. Esta adesão contribuiria para que houvesse uma maior proximidade entre essas células facilitando interações ligante-receptor envolvendo *pili* não adesivos. A zona de adesão favoreceria também a entrega de fatores de virulência. Além disso, a adesão possibilitaria a invasão de células do hospedeiro por microrganismos Gram-positivos (Mandlik *et al.*, 2008).

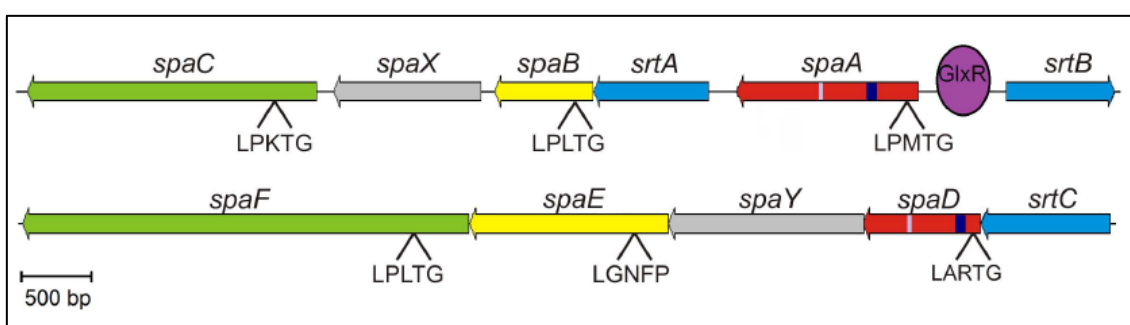


**Figura 3.** Modelo de mediação por *pili* de adesão e patogênese.

Adesinas fazem o contato inicial com os receptores das células do hospedeiro e pilinas ligadas a parede celular medeiam a formação da zona de adesão, o que possibilita interações adicionais ligante-receptor, entrega de fatores de virulência e invasão intracelular de patógenos.

\*Adaptado de Mandlik *et al.*, 2008.

O sequenciamento de *C.pseudotuberculosis* FRC41 e a procura por proteínas que possuíssem o motivo LPXTG, revelou seis proteínas que haviam sido anotadas como hipotéticas sendo que seis delas apresentaram similaridade a subunidades de *pili* adesivos de *C. diphtheriae* NCTC 13129. Essas regiões codificadoras estão organizadas em dois *clusters* gênicos como mostrado na Figura 4. Os *pili* adesivos de *C. pseudotuberculosis* consistem de duas subunidades de pilina principais (SpaA e SpaD), duas subunidades de pilinas menores (SpaB e SpaE) e duas proteínas de extremidade (SpaC e SpaF). São codificadas também duas proteínas hipotéticas, SpaX e SpaY, de função desconhecida (Trost *et al.*, 2010).



**Figura 4.** Clusters gênicos codificadores de *pili* adesivos de *C. pseudotuberculosis* FRC41.

Os *clusters* gênicos codificam sortases, necessárias para a montagem do *pilus* (azul), pilinas principais (vermelho), pilinas menores (amarelo), pilinas de extremidade (verde) e proteínas de função desconhecida (cinza).

\*Adaptado de Trost *et al.*, 2010.

É possível que as adesinas de *C. pseudotuberculosis* desempenhem papel semelhante às adesinas de *C. diphtheriae*. Sendo a SpaC uma proteína de extremidade de *pili* envolvida no contato inicial do microrganismo com a célula do hospedeiro, ela despertou interesse no sentido de ser melhor compreendido e estudado o seu papel em *C. pseudotuberculosis*. São necessários estudos com relação à sua ação, estrutura e também imunogenicidade.

### 1.3.2.3 Superóxido-dismutase C - SodC ou Cu/Zn-SOD

Superóxido-dismutases (SOD's) são metaloenzimas que detoxificam radicais de oxigênio pela reação de superóxido ( $O_2^-$ ) com íons hidrogênio ( $H^+$ ) gerando oxigênio molecular ( $O_2$ ) e peróxido de hidrogênio  $H_2O_2$ , o qual é reduzido a água e oxigênio pela enzima catalase ou peroxidase (Fridovich, 1995). As espécies reativas de oxigênio (ROS)

como  $O_2^-$ ,  $OH^-$  e  $H_2O_2$  causam danos aos ácidos nucleicos, proteínas e lipídeos das células (Imlay & Linn, 1988; Storz *et al.*, 1990).

As SODs estão presentes em procariotos e eucariotos e protegem as células dos ânions superóxido gerados durante a aerobiose. Além de proteger dos superóxidos gerados endogenamente, as SODs estão envolvidas nos mecanismos de patogenicidade de agentes etiológicos que acometem plantas e animais (Dhar *et al.*, 2013). Patógenos bacterianos enfrentam estresse oxidativo após o surto respiratório seguinte à fagocitose, e as SODs podem estar envolvidas também na detoxificação de ROS gerados dessa forma (Smith *et al.*, 1996; Lynch & Kuramitsu, 2000; Leclere *et al.*, 2001; Lefebvre & Valvano, 2001; Santos *et al.*, 2001; Dhar *et al.*, 2013). As SODs são classificadas em quatro tipos de acordo com o seu cofator metálico: cofator manganês (Mn-SOD ou SodA), cofator ferro (Fe-SOD ou SodB), cofator cobre e zinco (Cu/Zn-SOD ou SodC) e níquel cofator (Ni-SOD ou SodN) (Fridovich, 1995; Youn *et al.*, 1996). SodA e SodB procarióticas são encontradas no citoplasma da célula enquanto que SodC é periplasmática em Gram-negativas ou extracelularmente expostas em Gram-positivas (Fridovich, 1995). Em *Streptomyces* SodN é encontrada no citoplasma da célula (Youn *et al.*, 1996). Células que não produzem SODs apresentam, em geral, aumento do dano causado por estresse oxidativo (Dhar *et al.*, 2013).

A enzima SodC apresenta atividade de proteção associada à virulência em diversos patógenos. Em *Mycobacterium tuberculosis*, contribui para a resistência do microrganismo contra o estresse oxidativo gerado por macrófagos ativados (Dussurget *et al.*, 2001; Piddington *et al.*, 2001). SodC foi investigada em *Salmonella choleraesuis*, na qual foi associada com a sensibilidade a íons superóxido e óxido nítrico *in vitro*, capacidade de sobrevivência em macrófagos murinos e persistência em modelo murino de infecção (Sansone *et al.*, 2002). Em *Neisseria meningitidis*, SodC influencia no processo de fagocitose protegendo as células opsonizadas (Dunn *et al.*, 2003). Keith & Valvano (2007), demonstraram a participação da enzima na virulência de *Burkholderia cenocepacia* e confirmaram a localização periplasmática da enzima neste microrganismo. Em 2009, Ibarra & Steele-Mortimer relacionaram SodC à virulência de *Salmonella enterica* sorovar Thyphimurium. A enzima também foi apontada como fator de virulência em *Salmonella enterica* sorovar Typhi e Paratyphi (Sanjay *et al.*, 2010), *Bacillus anthracis* (Tonello & Zornetta, 2012), e *Staphylococcus aureus* (Gaupp *et al.*, 2012).

O sequenciamento do genoma de *C. pseudotuberculosis* FRC41 identificou dois tipos de SOD nesta linhagem a Mn-SOD (SodA) citoplasmática e a Cu/Zn-SOD (SodC) secretada que pode estar ancorada a membrana celular. Esta localização extracelular de SodC sugere que ela pode proteger a superfície das células de *C. pseudotuberculosis* contra ânions superóxido gerados pelas células do hospedeiro (Trost *et al.*, 2010). São necessários



estudos a fim de avaliar o papel protetor de SodC em *C. pseudotuberculosis*, sua ação *in vitro* bem como elucidação da sua estrutura e avaliação de imunogenicidade.

#### 1.3.2.4 Neuraminidase H (NanH)

Neuraminidases, também chamadas de sialidases, são glicosil-hidrolases que hidrolisam ligações  $\alpha$  de ácidos siálicos terminais a partir de vários sialoglicoconjugados em diferentes organismos, como vírus, microrganismos e animais vertebrados (Taylor, 1996). Ácidos siálicos, ou ácidos neuramínicos, são uma família de mais de 40 monossacarídeos de nove carbonos encontrados na porção terminal de glicoproteínas, glicolípídeos e polissacarídeos. Existem quatro tipos de ácido siálico: ácido N-acetilneuramínico (Neu5Ac), ácido N-glicosilneuramínico (Neu5Gc) e ácido 5-hidroxineuramínico (KDN) e ácido neuramínico (Neu). Estes ácidos siálicos podem sofrer diversas modificações por substituição no grupo hidroxil de O-acetil, O-metil, O-sulfate e grupos fosfato (Varki, 1999). São essas diferenças estruturais que contribuem para a função biológica de sialoglicoconjugados de superfície celular (Kim *et al.*, 2011). Eucariotos e procaríotos sintetizam ácido siálico através de vias bioquímicas complexas, sendo a principal delas envolvendo a síntese “*de novo*” de Neu5Ac na célula e sua ativação pela adição de citidina monofosfato (CMP) formando CMP-Neu5Ac que, por sua vez, é utilizado como doador de ácido siálico para síntese de sialoglicoconjugados por sialiltransferases (Kim *et al.*, 2011).

Em espécies animais, as neuraminidases possuem substratos autógenos, porém, em procaríotos, agem sobre moléculas do hospedeiro (Arden *et al.*, 1972). Nos eucariotos, essas enzimas estão envolvidas na modificação de sialoglicoconjugados de parede celular influenciando a regulação de interações célula-célula e célula-molécula e, dessa forma, medeiam processos de reconhecimento e adesão celular (Buschiazzo & Alzari, 2008). As neuraminidases estão presentes também em mais de 70 microrganismos das ordens Pseudomonadales e Eubacteriales, sendo que a maioria dos microrganismos que produzem neuraminidases são patogênicos ou comensais quando em contato com células do hospedeiro (Roggentin *et al.*, 1993; (Kim *et al.*, 2011). Dessa forma, as neuraminidases são apontadas como fator de virulência em diversos microrganismos. As neuraminidases bacterianas estão envolvidas no sequestro de ácidos siálicos de glicoproteínas e glicolípídeos sializados ou outros glicoconjugados para nutrição. Estes ácidos siálicos sequestrados via sialidases ou *trans*-sialidases podem ser transferidos para a superfície celular do microrganismo contribuindo para evasão do sistema imune e interação com células do hospedeiro (Vimr & Lichtensteiger, 2002; Vimr *et al.*, 2004).

As neuraminidases são divididas em dois grupos com base em seu tamanho. As neuraminidases com peso molecular em torno de 40kDa formam o grupo das

neuraminidases de pequeno tamanho como as que são encontradas em *Salmonella spp* e *Clostridium perfringens*. As com peso molecular acima de 80kDa são as neuraminidases de grande tamanho como as de *Clostridium tertium* e *Vibrio cholerae* (Mizan *et al.*, 2000).

A maioria das neuraminidases bacterianas são secretadas. Porém, algumas encontram-se ligadas a superfície como a Neuraminidase A (NanA) de *Streptococcus pneumoniae* (Dalia *et al.*, 2010), NanH de *C. diphtheriae* (Kim *et al.*, 2009) e TEDE0471 de *Treponema denticola* (Kurniyati *et al.*, 2013). A homologia de sequência aminoacídica entre as neuraminidases bacterianas é de menos de 30%. Porém, a topologia do domínio sialidase é bem conservada tendo motivos e resíduos característicos (Kim *et al.*, 2011).

Em *C. diphtheriae* NCTC1312 foram identificados dois genes codificadores de sialidases putativas, NanH (DIP0543) e NanI (DIP0330) (Cerdeño-Tárraga *et al.*, 2003). Na linhagem *C. diphtheriae* KCTC3075 foi identificada e caracterizada a Neuraminidase H (NanH) com atividade neuramidase e *trans*-sialidase (Kim *et al.*, 2009). Análises bioinformáticas de NanH de *C. diphtheriae* KCTC3075 predisseram a presença de um peptídeo sinal N-terminal de 32 aminoácidos e uma estrutura de alfa-hélice transmembrana C-terminal levando a supor que trata-se de uma proteína ligada a membrana com potencial função de fator de virulência. O peso molecular de NanH nesta linhagem é de 70 kDa, o que sugere a presença de um domínio adicional, como o domínio de ligação a carboidrato *lectin-like* (Kim *et al.*, 2009). As neuraminidases de superfície celular são objetos de estudo como alvo vacinal em várias doenças patogênicas (Taylor, 1996; Mattos-Guaraldi *et al.*, 1998; Tai, 2006).

A partir de estudos de dados do genoma da linhagem FRC41 de *C. pseudotuberculosis*, o gene *nanH* foi identificado e sugerido como provável fator de virulência de *C. pseudotuberculosis* sendo predito como codificador de uma neuraminidase extracelular (Trost *et al.*, 2010).

São necessários estudos para a caracterização da atividade de NanH em *C. pseudotuberculosis*, elucidação de sua estrutura e avaliação de seu potencial imunogênico.

#### **1.4 Expressão heteróloga de proteínas em *Escherichia coli***

A expressão heteróloga é uma importante ferramenta para o estudo, caracterização e elucidação estrutural de proteínas. Com o surgimento da tecnologia do DNA recombinante, surgiu a “segunda geração” de vacinas baseando-se no uso de proteínas recombinantes. A partir do sequenciamento de novos genomas bacterianos e a utilização de tecnologias como a bioinformática, tem sido possível ampliar o conhecimento acerca da patogênese destes microrganismos. Com isso, novos alvos tem sido identificados para

utilização como vacina recombinante de subunidade ou ainda para o desenvolvimento de vacinas vivas atenuadas (Scarselli *et al.*, 2005).

Vários antígenos vacinais já foram expressos em *E.coli* e testados quanto à sua eficácia protetora contra o desafio exibindo resultados variados. As vacinas recombinantes de subunidade apresentam inúmeras vantagens como a exclusão do patógeno no processo de desenvolvimento da vacina, otimização das formas de apresentação dos antígenos, estabilidade e facilidade de administração (Liljeqvist & Stahl, 1999).

A expressão heteróloga fornece subsídios também para a produção de proteínas em larga escala para a realização de estudos estruturais. Esta é uma abordagem bastante atrativa para o desenvolvimento de fármacos, uma vez que grande parte deles tem proteínas como alvo (Marshall *et al.*, 2003).

A expressão de proteínas recombinantes pode utilizar diversos sistemas de expressão podendo ser realizada em bactérias, leveduras, células de insetos, plantas ou mamíferos. A escolha do sistema de expressão deve ser feita levando-se em consideração a estrutura, funcionalidade e complexidade da proteína a ser produzida bem como a produtividade esperada (Demain & Vaishnav, 2009).

Um dos primeiros e mais amplamente utilizado hospedeiro para produção de proteínas heterólogas é a *Escherichia coli* (Terpe, 2006). *E. coli* apresenta como vantagens alta taxa de crescimento, rápida expressão, facilidade de cultivo, boa taxa de produção, amplo conhecimento de suas características genéticas e disponibilidade de ferramentas para sua manipulação genética (Swartz, 1996; Demain & Vaishnav, 2009). Este sistema é bastante útil para a expressão de proteínas que não passem por modificações pós-traducionais em seu organismo de origem (Demain & Vaishnav, 2009).

Para uma eficiente expressão ao se adotar *E. coli* como sistema de expressão, alguns fatores devem ser levados em consideração. O cultivo em altas densidades celulares pode levar a toxicidade pela formação de acetato. Neste caso, pode-se optar pelo controle de oxigênio e alimentação exponencial de glicose (Demain & Vaishnav, 2009). Outro fato é que a superprodução de proteínas heterólogas pode levar a formação de agregados insolúveis, os corpúsculos de inclusão (Baneyx, 1999; Schumann E Ferreira, 2004; Terpe, 2006). Estes corpúsculos de inclusão são formados provavelmente porque as chaperonas tem sua capacidade extrapolada durante a alta produção de proteínas heterólogas (Kurland E Gallant, 1996; Mergulhão *et al.*, 2005). A formação de corpúsculos de inclusão pode ser um problema, uma vez que a proteína precisará passar por processo de reenovelamento e solubilização para que seja obtida em sua forma solúvel e ativa (Baneyx, 1999; Swarts, 2001). Porém, os corpúsculos de inclusão podem apresentar algumas vantagens como a proteção do produto de degradação proteolítica e a obtenção de altos níveis de produção (Rudolph, 1996).

A escolha do plasmídeo também é muito importante para obter uma produção adequada da proteína heteróloga. Deve-se levar em consideração na escolha do vetor, o promotor e sua regulação, região de terminação e iniciação, origem de replicação e número de cópias, marcador de seleção, códons, estabilidade do mRNA, estabilidade da proteína e a linhagem hospedeira a ser utilizada (Makrides, 1996; Sorensen & Mortensen, 2005; Jana & Deb, 2005). É desejável que o promotor seja forte, finamente regulado com indução simples e independente dos outros constituintes do meio (Terpe, 2006).

Para maximização da produção da proteína de interesse, deve-se aumentar a densidade celular no cultivo manipulando variáveis como temperatura, composição do meio e aeração (Hanning & Makrides, 1998; Baneyx, 1999; Jana & Deb, 2005; Sorensen & Mortensen, 2005; Peti & Page, 2007).

Como exposto, o sistema de expressão em *E. coli* representa uma ferramenta interessante e versátil para a produção de proteínas heterólogas. Desta forma, este sistema foi utilizado neste trabalho para a obtenção das proteínas PLD, PknG, SodC, SpaC e NanH. Estas proteínas, por sua vez, se caracterizam como importantes alvos de estudo com relação à sua estrutura, função e antigenicidade em *C. pseudotuberculosis*.

## 2. JUSTIFICATIVA

A LC encontra-se distribuída principalmente em países onde há extensas criações de ovinos e caprinos. No Brasil, em estudo realizado por Ribeiro e colaboradores (2001) a estimativa era que 30% dos rebanhos estivessem acometidos pela LC (Ribeiro *et al.*, 2001). Em Minas Gerais, estudos indicaram que os índices podem chegar a aproximadamente 80% de animais infectados (Seyffert *et al.*, 2010; Guimarães *et al.*, 2009). A LC é responsável por significativas perdas econômicas principalmente na caprinovinocultura (Baird & Fontaine, 2007). O conhecimento acerca dos fatores de virulência envolvidos na LC abre caminho para novas pesquisas que visem tanto o melhor entendimento da patogenicidade de *Corynebacterium pseudotuberculosis* quanto o desenvolvimento de novas abordagens para o seu controle. Neste sentido, várias pesquisas foram conduzidas baseadas no fator de virulência mais conhecido e estudado de *C. pseudotuberculosis*, a PLD, para o estudo de diferentes formulações e combinações vacinais (Haynes *et al.*, 1992; Hodgson *et al.*, 1992; Walker *et al.*, 1994; Hodgson *et al.*, 1994; Hodgson *et al.*, 1999; De Rose *et al.*, 2002; Fontaine *et al.*, 2006)

Além disso, outras proteínas com características relacionadas à virulência do microrganismo também são importantes alvos de estudo. Entre estas proteínas, encontram-se as extracelulares que, por sua vez, são candidatas vacinais em potencial por fazerem parte da interface parasito-hospedeiro sendo possivelmente antigênicas (Hansson *et al.*, 2000). As proteínas SpaC (Adesina), SodC (Superóxido-dismutase C) e NanH (Neuraminidase H) fazem parte deste grupo de proteínas sendo prováveis fatores de virulência e expostas extracelularmente. Proteínas moduladoras da expressão de virulência (Burnside & Rajagopal, 2011) também constituem alvos interessantes para o desenvolvimento de vacinas. Como demonstrado pela proteína StkP de *S. pneumoniae* que foi capaz de conferir proteção após desafio em modelos murinos de sepse e pneumonia (Giefing *et al.*, 2008; Giefing *et al.*, 2010). StkP pertence ao grupo das STPKs, assim como a proteína PknG, que por sua vez pode estar envolvida no metabolismo de glutamato, bem como na modulação da fusão de fagossomo e lisossomo (Walburger *et al.*, 2004; Av-Gay & Everett, 2000).

A obtenção de proteínas em escala suficiente para serem estudadas quanto aos seus diversos aspectos é um passo importante para o desenvolvimento de uma nova alternativa no controle da LC. Sugerimos como alvos de estudo as proteínas PLD, PknG, SpaC, SodC e NanH, uma vez que são possíveis fatores envolvidos na virulência de *C. pseudotuberculosis*. Neste sentido, a expressão heteróloga destas proteínas apresenta-se como uma importante ferramenta de estudo.

### 3. OBJETIVOS

#### 3.1. Objetivos Gerais

Caracterização *in silico* e obtenção das proteínas PlD, PknG, SpaC, SodC e NanH de *Corynebacterium pseudotuberculosis* via sistema de expressão heteróloga em *E.coli* e, assim, possibilitar a caracterização estrutural e determinação do potencial imunogênico destas proteínas.

#### 3.2. Objetivos Específicos

- Caracterização *in silico* das proteínas PlD, PknG, SpaC, SodC e NanH, através de predição dos parâmetros físico-químicos, peptídeo sinal, localização subcelular, presença de domínios conservados, conservação da sequência, epítomos.
- Síntese das *ORFs* (*Open Reading Frame*) codificadoras das proteínas PlD, PknG, SpaC, SodC e NanH modificadas com códons preferenciais de *E. coli* (© 2013 DNA2.0, Inc.).
- Avaliação da expressão de cada uma das proteínas por cinco diferentes linhagens de *Escherichia coli*.
- Avaliação da cinética de expressão de cada uma das proteínas na linhagem de *E.coli* escolhida.
- Expressão em maior escala e purificação de rPknG.

## 4. MATERIAIS E MÉTODOS

### 4.1 Caracterização *in silico* das ORF's e proteínas

O genoma da linhagem *Corynebacterium pseudotuberculosis* FRC41 (CP002097.1) foi utilizada como organismo modelo neste trabalho. As proteínas alvo do estudo (PLD, PknG, SpaC, SodC e NanH) foram escolhidas com base em busca na literatura e em análises de genômica comparativa. As sequências nucleotídicas de cada uma das *Open Reading Frames* (ORF's) e suas respectivas sequência protéicas foram obtidas através de consulta ao banco de dados do NCBI – *National Center for Biotechnology Information* - (<http://www.ncbi.nlm.nih.gov/>). No caso da proteína PLD, trabalhamos com a versão atóxica caracterizada pela substituição do aminoácido His<sup>20</sup> por Ser (Hodgson *et al.*, 1999). Na Tabela 1 estão listados os identificadores das sequências.

**Tabela 1.** Nomenclatura e identificadores dos genes e proteínas de *C. pseudotuberculosis* FRC41 empregados no estudo.

Gene	Locus tag	Produto gênico	Número de acesso
<i>pld</i>	cpfrc_00029	Fosfolipase D (PLD)	YP_003782429.1
<i>pknG</i>	cpfrc_01829	Serina/treonina proteína quinase (PknG)	YP_003784229.1
<i>spaC</i>	cpfrc_01870	Proteína ancorada à superfície (SpaC)	YP_003784270.1
<i>sodC</i>	cpfrc_00610	Cu/Zn superóxido dismutase (SodC)	YP_003783011.1
<i>nanH</i>	cpfrc_00386	Neuraminidase (sialidase) (NanH)	YP_003782786.1

#### 4.1.1 Predição de parâmetros físico-químicos

Para a predição dos parâmetros físico-químicos das proteínas, foi utilizado o *Web-software* ProtParam, o qual permite computar vários parâmetros físicos e químicos que podem ser deduzidos a partir de uma sequência protéica – (<http://web.expasy.org/protparam/>) (Gasteiger *et al.*, 2005).

#### 4.1.2 Predição de peptídeo sinal

Para a predição de presença de peptídeo sinal nas sequências e seu sítio de clivagem, foi utilizada o *Web-software* SignalP 4.1 – (<http://www.cbs.dtu.dk/services/SignalP/>). O SignalP 4.1 utiliza uma combinação de redes neurais artificiais treinadas com sequências contendo segmentos transmembrânicos e não contendo estes segmentos. O resultado é apresentado a partir de três diferentes *scores*: C, S e Y. O C-score distingue o sítio de clivagem de peptídeo sinal. Já o S-score distingue posições que estão contidas no peptídeo sinal das que estão na porção madura da proteína. Enquanto que o Y-score é uma média geométrica dos dois *scores* anteriores. Além destes três *scores*, é gerado o D-score, uma média ponderada entre o S-score e Y-score que é utilizada para discriminar peptídeos sinal e peptídeos não sinal (Petersen *et al.*, 2011).

#### 4.1.3 Predição de localização subcelular

O *Web-software* PSORTb 3.0.2 – (<http://www.psort.org/psortb/index.html>) foi utilizado para predição de localização subcelular das proteínas. PSORTb consiste em múltiplos módulos analíticos que analisam características biológicas que influenciam ou caracterizam determinada localização subcelular. Estes módulos atuam como um preditor binário. Em um organismo Gram-negativo, as localizações subcelulares possíveis para as proteínas são: citoplasmática, membrana citoplasmática, membrana externa e espaço extracelular. Já os organismos Gram-positivos apresentam as seguintes localizações subcelulares: citoplasmática, membrana citoplasmática, parede celular e espaço extracelular. Os resultados de cada módulo de predição são combinados e validados estatisticamente. O resultado da predição é disponibilizado como uma lista com os possíveis sítios de localização e o valor de probabilidade associado a eles (Yu *et al.*, 2010).

#### 4.1.4 Busca por domínios conservados

Uma busca por domínios conservados nas sequências protéicas foi conduzida através do *Web-software* InterProScan5 – (<http://www.ebi.ac.uk/Tools/pfa/iprscan5/>), uma ferramenta que combina diferentes métodos de reconhecimento de padrões de proteínas e permite a identificação de domínios protéicos conservados (Quevillon *et al.*, 2005).



#### 4.1.5 Alinhamento de sequências e busca por similaridade

Cada uma das sequências protéicas foi submetida a uma busca por similaridade utilizando o algoritmo de alinhamento local de sequências BLAST (*Basic Local Alignment Search Tool*) contra a base de dados do *Web service* Uniprot – (<http://www.uniprot.org/>) (The UniProt Consortium, 2014).

Primeiramente, a busca foi realizada sobre o banco de dados geral da plataforma, o UniprotKB, adotando-se um *threshold* de 0.01 e o limite de alinhamento de 500 *hits*. A partir das sequências que apresentaram identidade com as proteínas alvo, foram selecionadas para um alinhamento múltiplo de sequências aquelas pertencentes a microrganismos representantes do grupo CMNR (gêneros *Corynebacterium*, *Mycobacterium*, *Nocardia* e *Rhodococcus*) com o objetivo de avaliar a conservação das proteínas PLD, PknG, SpaC, SodC e NanH neste grupo de microrganismos. As linhagens selecionadas como representantes do grupo CMNR foram as seguintes: *C. pseudotuberculosis* 1002 (biovar ovis – isolada de caprino); *C. pseudotuberculosis* 258 (biovar equi); *C. pseudotuberculosis* C231 (biovar ovis – isolada de ovino); *Corynebacterium diphtheriae* HC02; *C. glutamicum* ATCC 14067; *Mycobacterium tuberculosis* SUMu007; *Nocardia farcinica* (strain IFM 10152); *Rhodococcus pyridinivorans* SB3094; *Rhodococcus* sp. (strain RHA1). Para cada alinhamento múltiplo de sequência é gerado um valor de *E-value*, o qual mensura a probabilidade de um alinhamento de sequências ter ocorrido ao acaso. Quanto menor o *E-value*, mais significativa a medida. As sequências pertencentes a estes microrganismos e que apresentaram identidade com as sequências de PLD, PknG, SpaC, SodC e NanH de *C. pseudotuberculosis* FRC41 passaram pelo alinhamento múltiplo de sequências através do *software* ClustalW2 – (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>) (Thompson *et al.*, 1994). Os alinhamentos foram visualizados através do *software* Jalview – (<http://www.jalview.org/>) (Waterhouse *et al.*, 2009), o qual possibilita visualização, edição e análise de alinhamentos múltiplos de sequências.

Com o objetivo de avaliar a conservação das proteínas (PLD, PknG, SpaC, SodC e NanH de *C. pseudotuberculosis* FRC41) em organismos eucarióticos, foi realizada uma busca através do BLAST por sequências com maior porcentagem de identidade com as proteínas citadas. Para isso, foi utilizado o banco de dados de sequências eucarióticas do Uniprot (UniprotKB – Eukaryota). Também foi adotado um *threshold* de 0,01 e o limite de alinhamento de 500 *hits*. A partir do resultado gerado pelo BLAST, buscamos pelas sequências de espécies do gênero *Ovis*, *Bos* e *Equus* além das espécies *Mus musculus* e *Homo sapiens*. As sequências selecionadas também foram alinhadas através de ClustalW e o visualizou-se o alinhamento através do *software* Jalview. Nos casos em que as proteínas de *C. pseudotuberculosis* apresentaram maior conservação em organismos eucariotos,

buscou-se identificar quais seriam as regiões de maior conservação dessas sequências. Estas observações são importantes para abordagens futuras de desenvolvimento de vacinas e/ou drogas.

#### 4.1.6 Predição de epítomos

O *Web-software* MED (*Mature Epitope Density*) – (<http://med.mmci.uni-saarland.de/>) (Santos *et al.*, 2013) foi utilizado com o objetivo de prever a densidade de epítomos de MHC de classe I nas proteínas alvo deste estudo. O MED faz a predição de epítomos integrando as ferramentas SurfG+ (Barinov *et al.*, 2009), TMHMM (Krogh *et al.*, 2001) e NetMHC. A sequência correspondente ao peptídeo sinal foi removida quando presente, restando apenas a sequência predita madura. Essa sequência é submetida ao NetMHC 3.0 para predição dos alelos de MHC depositados no *software* e são filtrados os peptídeos preditos como peptídeos de ligação forte. Então é calculado o MED score para cada sequência proteica por uma equação que divide o número de epítomos lineares preditos para cada sequência pelo número de peptídeos de 9-*mer* possíveis refletindo a densidade de epítomos na sequência. Para assegurar uma diferenciação qualitativa, é realizada uma normalização baseada na afinidade de ligação ao MHC. Então é gerado o MED score em nanomolar por *mer* (nM/*mer*) (Lundegaard *et al.*, 2008).

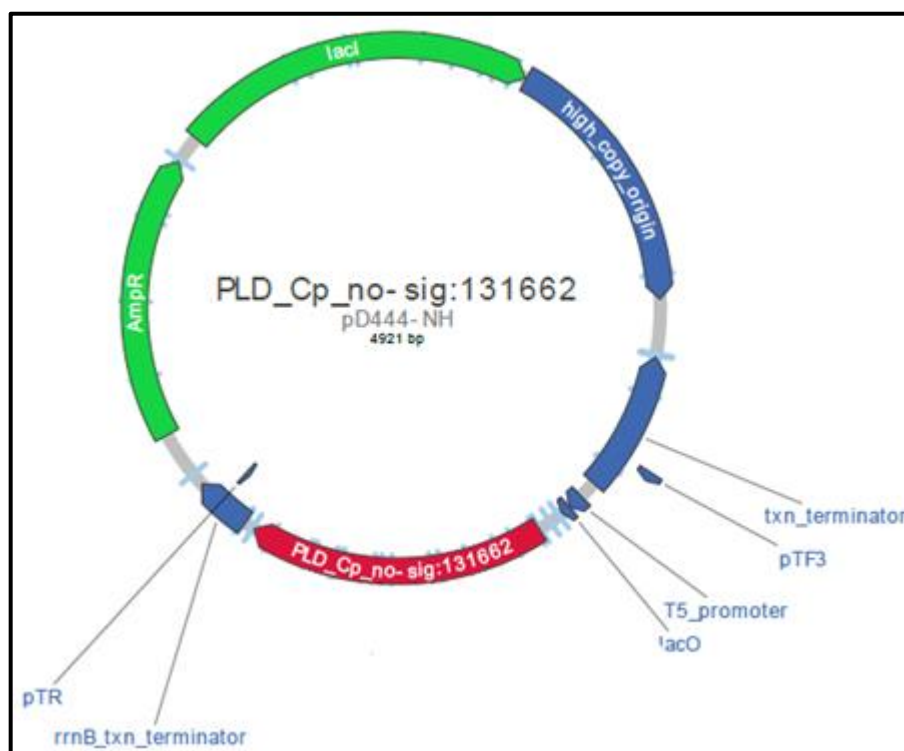
Para a predição de percentual de epítomos de MHC de classe II com alta afinidade de ligação, adotou-se o *Web-software* NetMHCII 2.2 – (<http://www.cbs.dtu.dk/services/NetMHCII/>) que prediz a ligação de peptídeos a alelos de HLA-DR, HLA-DQ, HLA-DP e alelos de MHC de classe II de camundongo. Todos os possíveis peptídeos de 15-*mer* dentro de uma sequência de aminoácidos recebem uma classificação baseada na afinidade de ligação do peptídeo. O cálculo é baseado no *core* de 9 aminoácidos dentro do peptídeo de 15-*mer*. Os peptídeos podem ser classificados em peptídeos de forte ligação, baixa ligação ou não receberem classificação por não terem atingido o *score* de ligação (Nielsen & Lund, 2009).

A fim de identificar porções antigênicas das proteínas com base na busca por epítomos de células B, utilizou-se o *Web-software* Bcepred – (<http://www.imtech.res.in/raghava/bcepred/index.html>). A predição foi realizada levando-se em conta a acessibilidade dos aminoácidos. (Saha & Raghava, 2004).

## 4.2 Clonagem molecular e expressão das proteínas recombinantes

### 4.2.1 Vetor de expressão pD444-NH e construção das ORF's sintéticas das proteínas selecionadas

Para a expressão das proteínas recombinantes, foi utilizado um vetor de expressão comercial da linha “pD444-NH Bacterial Expression Vectors” comercializado pela empresa *DNA2.0 Inc.* EUA (<https://www.dna20.com/>). O plasmídeo pD444-NH apresenta algumas características relevantes, dentre as quais destacam-se o promotor T5 induzível por IPTG - análogo de lactose não metabolizável por *E.coli*, marcador de resistência ao antibiótico ampicilina, adição de cauda de histidina (6xHis) à porção N-terminal do inserto e alto número de cópias (Figura 5).



**Figura 5.** Representação esquemática do vetor pD444-NH contendo o inserto *pld* (pD444-NH:*pld*).

As ORF's de cada uma das proteínas (*pld*, *pkng*, *spac*, *sodc*, *nanh*) foram sintetizadas e clonadas individualmente ao vetor pD444-NH pela empresa *DNA2.0*. Primeiramente, as sequências nucleotídicas codificadoras de cada uma das proteínas foi obtida via consulta ao banco de dados *GenBank* (<http://www.ncbi.nlm.nih.gov>), através do seu número de acesso (Tabela 1). Após predição de peptídeo sinal através do programa

SignalP 4.0 (Item 4.1), aquelas proteínas que apresentaram predição positiva de peptídeo sinal, tiveram a sequência codificadora do mesmo removida. As sequências codificadoras foram enviadas à empresa DNA 2.0 para síntese e convertidas a uma sequência modificada pela substituição dos códons originais por códons preferenciais de *E.coli*, espécie hospedeira escolhida para a expressão das proteínas recombinantes neste estudo. As sequências das *ORF*'s otimizadas com códons preferenciais de *E. coli* podem ser consultadas no Anexo 1.

As construções geradas pela clonagem de cada uma das *ORF*'s em plasmídeo pD444-NH foram denominadas: pD444-NH:*pld*; pD444-NH:*pknG*; pD444-NH:*spac*; pD444-NH:*sodC*; pD444-NH:*nanH* (Tabela 2).

**Tabela 2.** Plasmídeos gerados a partir da clonagem das *ORF*'s em vetor de expressão pD444-NH.

Identificação do Plasmídeo	Inserto ( <i>ORF</i> )	Proteína Recombinante
pD444-NH: <i>pld</i>	<i>pld</i>	rPLD
pD444-NH: <i>pknG</i>	<i>pknG</i>	rPknG
pD444-NH: <i>spac</i>	<i>spac</i>	rSpaC
pD444-NH: <i>sodC</i>	<i>sodC</i>	rSodC
pD444-NH: <i>nanH</i>	<i>nanH</i>	rNanH

#### 4.2.2 Preparação do DNA plasmidiano

Os plasmídeos foram recebidos da empresa DNA2.0 em papéis filtro circulares. Para a eluição do material, cada papel filtro foi colocado em um microtubo de 0,6mL contendo um orifício em sua porção inferior e o papel filtro permaneceu em contato com a parede do microtubo. Sobre cada papel contendo DNA, adicionou-se, com o auxílio de uma pipeta, 100µL de Tris-HCl 100mM, pH 7,5. O sistema permaneceu em repouso por cerca de 30 min a temperatura ambiente. Em seguida, cada microtubo de 0.6mL contendo o papel filtro com DNA foi colocado em um microtubo de polipropileno coletor de 2mL e centrifugou-se por 2 min. A fração recolhida foi armazenada e utilizada para transformação em *E.coli*.

## 4.2.3 Transformação bacteriana

### 4.2.3.1 Preparo de células eletrocompetentes

Neste estudo foram adotadas para clonagem molecular e expressão das proteínas recombinantes as seguintes linhagens: OverExpress™ C41(DE3), OverExpress™ C41(DE3) pLysS, OverExpress™ C43(DE3), OverExpress™ C43(DE3) pLysS e BL21 Star™ (DE3). Os genótipos dessas linhagens são listados na Tabela 3.

**Tabela 3.** Linhagens bacterianas utilizadas para clonagem molecular e expressão das proteínas recombinantes.

Linhagem	Genótipo
OverExpress™ C41(DE3)	F <sup>-</sup> ompT gal dcm hsdS <sub>B</sub> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> )(DE3)
OverExpress™ C41(DE3) pLysS	F <sup>-</sup> ompT gal dcm hsdS <sub>B</sub> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> )(DE3)pLysS (Cm <sup>r</sup> )
OverExpress™ C43(DE3)	F <sup>-</sup> ompT gal dcm hsdS <sub>B</sub> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> )(DE3)
OverExpress™ C43(DE3) pLysS	F <sup>-</sup> ompT gal dcm hsdS <sub>B</sub> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> )(DE3)pLysS (Cm <sup>r</sup> )
BL21 Star™ (DE3)	F <sup>-</sup> ompT hsdSB(rB <sup>-</sup> , mB <sup>-</sup> ) gal dcm rne131 (DE3)

Para a preparação das células eletrocompetentes, inoculou-se 100µL de cada linhagem de interesse separadamente em 5mL de meio LB Difco e incubou-se a 37°C por 18h sob agitação a 150rpm. Então, 5mL da cultura foram adicionados a 500mL de meio LB e incubou-se a 37°C sob agitação a 150rpm até que fosse atingida a densidade ótica lida a um comprimento de onda de 600nm (DO<sub>600nm</sub>) de 0,5-0,6, correspondente à fase exponencial de crescimento. Tendo alcançado a DO<sub>600</sub> desejada, a cultura permaneceu em repouso por 30 min em gelo. Decorrido este tempo, a cultura foi dividida em 6 tubos tipo *falcon* de 50mL e centrifugou-se a 4000rpm por 15min a 4°C, sendo o sobrenadante descartado. O precipitado foi ressuscitado em 40mL de glicerol 10% estéril, centrifugou-se novamente e o sobrenadante foi descartado. Os precipitados dos tubos 5 e 6 foram ressuscitados em 20 mL de glicerol 10% estéril e o seu conteúdo dividido nos outros 4 tubos restantes que tiveram os precipitados ressuscitados e aos quais foi adicionado glicerol 10% até completar-se o volume de 40mL. Procedeu-se nova centrifugação e o sobrenadante foi descartado. Os precipitados dos tubos 3 e 4 foram ressuscitados com 10mL de glicerol 10% e o seu conteúdo dividido nos outros dois tubos restantes que tiveram os precipitados ressuscitados e aos quais foi adicionado glicerol 10% até completar-se o volume de 40mL. Então, centrifugou-se novamente e o sobrenadante foi descartado. O precipitado do tubo 2 foi ressuscitado em 10mL de glicerol 10% e o conteúdo transferido

para o tubo 1 que teve o precipitado ressuspensionado e ao qual foi adicionado glicerol 10% até completar-se o volume de 20mL. O tubo foi centrifugado, o sobrenadante foi descartado e o precipitado ressuspensionado em 1 mL de glicerol 10%. Esse conteúdo foi dividido em alíquotas de 100µL que foram armazenadas a -80°C.

#### **4.2.3.2 Transformação de linhagens de *E. coli* com os plasmídeos pD444-NH:*pId*, pD444-NH:*pkng*, pD444-NH:*spac*, pD444-NH:*sodc* e pD444-NH:*nanh***

Cada um dos plasmídeos (pD444-NH:*pId*, pD444-NH:*pkng*, pD444-NH:*spac*, pD444-NH:*sodc* e pD444-NH:*nanh*) foram transformadas via eletrotransformação em cada uma das linhagens de *E. coli* escolhidas para este trabalho (OverExpress™ C41(DE3), OverExpress™ C41(DE3) pLysS, OverExpress™ C43(DE3), OverExpress™ C41(DE3), pLysS e BL21 Star™ (DE3)).

Para tanto, cerca de 100ng a partir da alíquota original do plasmídeo foram adicionados à alíquota de 100µL de bactéria eletrocompetente permanecendo em repouso no gelo por 5min. Decorrido este tempo, a mistura foi transferida a uma cubeta de eletroporação de 2mm BioRad® e as eletrotransformações foram realizadas utilizando o eletroporador *Gene Pulser Xcell™ Total System* (BioRad®) com um pulso de 2,4KV de voltagem, 25µF de capacitância e resistência de 200Ω. Após este procedimento, foi adicionado 1mL de meio LB sem antibiótico às células transformadas. As bactérias foram incubadas em banho a 37°C por 2h. Em seguida, 100µL da cultura foram plaqueados em meio LB-ágar suplementado com ampicilina (100µg/mL) e incubados a 37°C por 18h.

Foram selecionados clones a partir das placas resultantes da transformação para estoque e extração de DNA plasmidiano.

#### **4.2.4 Extração em pequena escala de DNA plasmidiano**

As colônias selecionadas para extração de DNA plasmidiano foram isoladamente transferidas para tubos tipo *falcon* de 50mL contendo 5mL de meio LB suplementado com ampicilina 100µg/mL. As culturas foram incubadas a 37°C sob agitação a 150rpm por 18h. Então, para cultura estoque, adicionou-se 1mL de glicerol 80% à 1mL da cultura e armazenou-se à -80°C. À partir da cultura incubada por 18h, foi retirada também uma alíquota de 1,5mL que foi transferida à microtubo polipropileno de 2mL e centrifugada a 14000rpm. O sobrenadante foi descartado e as células coletadas. A extração de DNA plasmidiano foi realizada através do método de lise alcalina descrito em Sambrook & Russel (2000) com algumas modificações.

As células coletadas após a centrifugação foram ressuspensas com 200µL de Solução I (Glicose 50mM; Tris-HCl 25mM pH 8,0; EDTA 10mM pH 8,0) e adicionou-se 1,5µL de RNase Invitrogen. A mistura foi incubada por 5min a temperatura ambiente. Em seguida, foram adicionados 200µL de Solução II (SDS 1%; NaOH 0,2N) e homogeneizou-se invertendo o tubo por 10 a 15 vezes. Então, foram adicionados 200µL de Solução III (KAc 5M, ácido acético glacial 11,5%), homogeneizou-se por inversão do tubo e a mistura foi centrifugada a 14000rpm por 7min. Em seguida, o sobrenadante foi transferido para novo microtubo de 2mL e adicionou-se 420µL de isopropanol. A mistura foi homogeneizada por inversão do tubo, mantida em repouso por 30min a -20°C e, em seguida, centrifugada por mais 15 minutos. Descartou-se o sobrenadante e o precipitado foi lavado com 500µL de etanol 70% invertendo o tubo de 3 a 4 vezes e centrifugou-se a 14000rpm por 5min. O sobrenadante foi descartado e o precipitado seco em estufa e ressuspendido em 30µL de água ultrapura estéril.

A integridade e qualidade do DNA plasmidiano extraído foram avaliadas por eletroforese utilizando “1 Kb *Plus DNA Ladder*” (Invitrogen) como padrão. Para isso, 1µL da amostra, acrescida de tampão de amostra 1:10 (azul de bromofenol 0,5%, glicerol 25%) foi resolvido em gel de agarose 1% preparado em tampão TBE 0,5X (Tris-base 89mM pH 7,2; ácido bórico 89mM, EDTA 2mM pH 8,3). A resolução foi realizada a 100V por aproximadamente 1h. O gel foi corado com brometo de etídio (0,5µg/mL), visualizado e fotodocumentado sobre transiluminador de luz ultravioleta através do sistema de documentação fotográfica “MultDoc-it/UVP”. As fotos foram processadas através do programa *PhotoRead*.

#### **4.2.5 Indução da expressão das proteínas recombinantes em *E. coli***

##### **4.2.5.1 Indução da expressão das proteínas recombinantes em diferentes linhagens de expressão**

As proteínas recombinantes receberam as seguintes denominações: rPLD, rPknG, rSpaC, rSodC e rNanH.

Foi avaliada a expressão em pequena escala de cada uma das proteínas pelas cinco linhagens de *E. coli* adotadas no estudo: OverExpress™ C41(DE3), OverExpress™ C41(DE3) pLysS, OverExpress™ C43(DE3), OverExpress™ C43(DE3) pLysS, BL21 Star™ (DE3) a fim de selecionar a linhagem que oferecesse o melhor nível de expressão de cada proteína. Para isso, 5µL da cultura estoque de cada uma das linhagens de *E. coli* transformadas com cada um dos plasmídeos (Item 4.2.3.2) foram inoculados em 5mL de meio LB suplementado com ampicilina (100µg/mL) e incubados a 37°C sob agitação a

150rpm por 18h. Então, 1mL de cada cultura crescida por 18h foi inoculado em 10mL de meio LB acrescido de ampicilina (100µg/mL) e incubou-se novamente a 37°C sob agitação a 150rpm até que fosse atingida a  $DO_{600nm}$  de 0,5-0,7. Atingida a densidade óptica desejada, 1mL da cultura foi coletado e transferido para um microtubo de 1,5mL que foi centrifugado a 14.000rpm por 2min, sendo o sobrenadante descartado. Esta fração foi denominada tempo zero (T0) não induzido. Em seguida, as culturas foram induzidas pela adição de Isopropiltio-β-D-galactopiranosídeo (IPTG) a uma concentração final de 1mM. A cultura permaneceu incubada a 37°C sob agitação a 150rpm por um tempo de 5h. Ao final das 5h de indução, foram coletadas alíquotas de 1mL que foram transferidas para microtubos de 1,5mL e centrifugadas a 14.000rpm por 2min. Estas frações foram denominadas tempo cinco de indução (T5) induzido. As alíquotas T0 e T5 foram utilizadas para comparar o nível de expressão das proteínas recombinantes oferecido por cada uma das linhagens de *E.coli*. As frações coletadas foram resolvidas por eletroforese em gel de poliacrilamida sob condições desnaturantes (SDS-PAGE). A partir disso, foi selecionada a linhagem com melhor nível de expressão de cada proteína.

#### **4.2.5.2 Resolução eletroforética de proteínas em gel de poliacrilamida sob condições desnaturantes – SDS-PAGE**

As amostras foram centrifugadas a 14.000rpm por 2 minutos e os precipitados foram ressuspendidos em 20µL de Tampão de Amostra (0,00625M Tris-HCl pH 6,8; SDS 2%, glicerol 10%, azul de bromofenol 0,0005%, β-mercaptoetanol 2,5%) e aquecidas a 100°C por 10 min. A resolução eletroforética foi realizada em tampão Tris-glicina (25mM Tris-base, pH8,3; glicina 250mM e SDS 0,1%) a 75V até que fosse atingido o gel de resolução e posteriormente a 150V utilizando sistema de eletroforese vertical da Biorad. As concentrações do gel de resolução para cada uma das proteínas estão listadas na Tabela 4. O gel foi incubado em solução corante (Metanol 50%, ácido acético 10%, Coomassie Brilliant Blue R-250 0,25%) por 3h sob agitação a temperatura ambiente. Em seguida, o gel foi descorado por uma solução descorante (Metanol 10%, ácido acético 10%) por tempo suficiente para que o padrão de bandas pudesse ser visualizado. A fotodocumentação dos géis foi realizada utilizando ImageScanner Amersham Biosciences e visualização pelo LabScann 5.00.



**Tabela 4.** Peso molecular e concentração do gel de resolução (SDS-PAGE) das proteínas recombinantes.

Proteína	Peso molecular	Concentração do gel de resolução
rPLD	31kDa	12%
rPknG	83kDa	10%
rSpaC	86kDa	10%
rSodC	18kDa	15%
rNanH	71,5kDa	10%

#### 4.2.5.3 Avaliação da cinética de expressão das proteínas recombinantes

Após a escolha de uma linhagem de *E.coli* para cada proteína recombinante, procedeu-se à indução da expressão das proteínas recombinantes em um litro de cultura. O procedimento descrito a seguir foi empregado para a expressão de cada proteína recombinante na linhagem de *E. coli* escolhida, separadamente.

Para isso, 20 $\mu$ L do clone congelado foram adicionados a 10mL de meio LB suplementado com ampicilina (100 $\mu$ g/mL) e incubados a 37°C sob agitação a 150rpm por 18h. Então, 10mL desta cultura foram inoculados em 1L de meio LB acrescido de ampicilina (100 $\mu$ g/mL). O volume foi dividido em quatro frascos erlenmeyer de 500mL e a cultura foi incubada a 37°C sob agitação a 150rpm até ser atingida a DO<sub>600nm</sub> de 0,5-0,8. Então, 1mL da cultura foi coletado, transferido a um microtubo de 1,5mL, centrifugado a 14000rpm por 2min e o sobrenadante foi descartado. Esta fração foi denominada tempo zero (T0) não induzido. Então, a cultura foi induzida pela adição de IPTG a uma concentração final de 1mM. A cultura permaneceu incubada a 37°C sob agitação a 180rpm por 5h sendo coletadas alíquotas de 1mL a cada hora, denominadas T1, T2, T3, T4 e T5. Essas alíquotas foram centrifugadas a 14000rpm por 2min sendo o sobrenadante descartado. As alíquotas T0-T5 foram resolvidas por eletroforese em gel de poliacrilamida em condições desnaturantes (SDS-PAGE) (Item 4.2.5.2) para avaliação da cinética de expressão de cada proteína recombinante na linhagem de *E.coli* escolhida.

#### 4.2.5.4 Teste de solubilidade das proteínas recombinantes

O teste de solubilidade foi realizado a partir da linhagem de *E. coli* que apresentou melhor desempenho na expressão proteica de cada alvo. O teste teve como finalidade verificar se a proteína encontra-se na fração solúvel ou insolúvel após a lise celular por sonicação para liberação das proteínas.

Para isso, alíquotas do tempo 5 de indução (T5) foram centrifugadas a 14.000rpm por 2min e tiveram o sobrenadante descartado (Item 4.2.5.2). Então, as células foram lisadas em tampão PBS 1X ( $\text{Na}_2\text{HPO}_4$  100 mM,  $\text{KH}_2\text{PO}_4$  17 mM, NaCl 1,4 mM, KCl 27 mM) por sonicação em 3 ciclos de 30s à 30KHz com intervalos de 10s entre os ciclos. Em seguida, as amostras foram centrifugadas a 14.000rpm por 2min e o sobrenadante foi transferido para outro microtubo. Foram adicionados 40 $\mu\text{L}$  de tampão de amostra no precipitado e 20 $\mu\text{L}$  do Tampão e Amostra em 20 $\mu\text{L}$  do sobrenadante. As amostras foram aquecidas a 100°C por 10min e resolvidas em gel de poliacrilamida sob condições desnaturantes (SDS-PAGE) (Item 4.2.5.2).

#### **4.2.6 Purificação da proteína recombinante**

Inicialmente 5 mL do pré-inóculo foram transferidos para um novo tubo contendo 1L de meio líquido seletivo (LB + Antibiótico específico). O tubo foi mantido sob agitação e aeração até atingir a densidade ótica ideal para indução da expressão com IPTG,  $\text{OD}_{600}$  entre 0,5 e 0,7. A leitura da amostra foi realizada em espectrofotômetro - Spectronic Genesys 2. A indução da expressão foi realizada pela adição de IPTG 0,2 mM a temperatura de 25°C a fim de obter proteínas na fração solúvel após a sonicação. Após 10 h de indução, foram coletadas amostras que foram analisadas através de eletroforese em gel desnaturante de poliacrilamida (SDS-PAGE).

A purificação foi conduzida em coluna de afinidade utilizando resina de níquel (ProBond) conforme instruções do fabricante (Invitrogen).

O meio foi centrifugado por 30 min a 9000rpm e 4°C e o precipitado foi ressuspenso e estocado por 1h em 20 mL de tampão de lise (50 mM tampão fosfato, 200 mM NaCl, 1 mM PMSF - fluoreto de fenilmetilsulfonil, pH 8) contendo 0,5 mg/mL de lisozima. Em seguida, a solução foi submetida a pulsos de 15 s a 300 W em sonificador para lise celular. Esta amostra foi centrifugada por 30 min a 17000rpm a 10 °C. A fase aquosa foi coletada e utilizada para a purificação da proteína de interesse.

Primeiramente, a resina (Ni-NTA) foi equilibrada com o mesmo tampão utilizado para lisar as células. Logo em seguida, o sobrenadante do lisado celular foi passado através da coluna de purificação (coluna de gravidade, Bio-Rad) para que a proteína (His-tag) fosse adsorvida. Após a adsorção, a resina foi lavada utilizando tampão de lavagem (20mM  $\text{NaH}_2\text{PO}_4$ , 500mM NaCl, 60 mM imidazol, 2% glicerol e 1mM PMSF, pH 7,4) para remover contaminantes adsorvidos inespecificamente na resina de níquel. Por fim, as proteínas foram eluídas utilizando tampão de eluição (20 mM tris, 300 mM NaCl, 300 mM imidazol , pH 7.4).

Utilizando coluna mono-Q, a proteína passou por uma nova purificação por troca iônica e, em seguida, utilizando sistema AKTA purifier foi realizada uma gel filtração, onde a proteína foi separada de acordo com sua massa molecular. A proteína foi concentrada utilizando concentradores Millipore e analisada em SDS-PAGE 12%.

#### **4.2.7 Western Blot**

A fração protéica purificada e resolvida por eletroforese SDS-PAGE foi transferida para membrana de PVDF (Amesham) em equipamento Owl utilizando-se diferença de potencial elétrico (ddp) de 5V por 1h e 30min. A membrana foi incubada em 10mL de solução de bloqueio do *Western Breeze<sup>®</sup> Chromogenic Western Blot Immunodetection Kit* (Invitrogen) por 30min sob agitação e em seguida, lavada com 20mL de água destilada por 5min. Então, a membrana foi incubada com anticorpo monoclonal anti-histidina (Sigma) na diluição 1:5000 por 1h. Posteriormente, foi adicionado sobre a membrana o anticorpo anti-IgG de camundongo conjugado com fosfatase alcalina permanecendo a membrana incubada com este anticorpo por 30min. A membrana foi lavada três vezes após cada incubação com anticorpo. A revelação foi realizada com a solução cromógena de BCIP/NBT do *Western Breeze Kit*.

## 5. RESULTADOS

### 5.1. Caracterização *in silico* das ORFs e proteínas

#### 5.1.1 Predição de parâmetros físico-químicos

Através da ferramenta ProtParam foram preditos os seguintes parâmetros para as proteínas PLD, PknG, SpaC, SodC e NanH de *C. pseudotuberculosis* FRC41: número de aminoácidos, peso molecular, ponto isoelétrico (pI) teórico, fórmula molecular e número de átomos (Gasteiger *et al.*, 2005). Na Tabela 5, estão apresentados os parâmetros obtidos para as proteínas. Além disso, PLD, PknG, SpaC, SodC e NanH foram classificadas como proteínas estáveis através desta ferramenta.

As sequências submetidas a essa análise correspondem à proteína madura. Portanto, a porção correspondente ao peptídeo sinal, naquelas proteínas que o apresentaram, foi removida para a análise.

**Tabela 5.** Parâmetros físico-químicos preditos através do ProtParam.

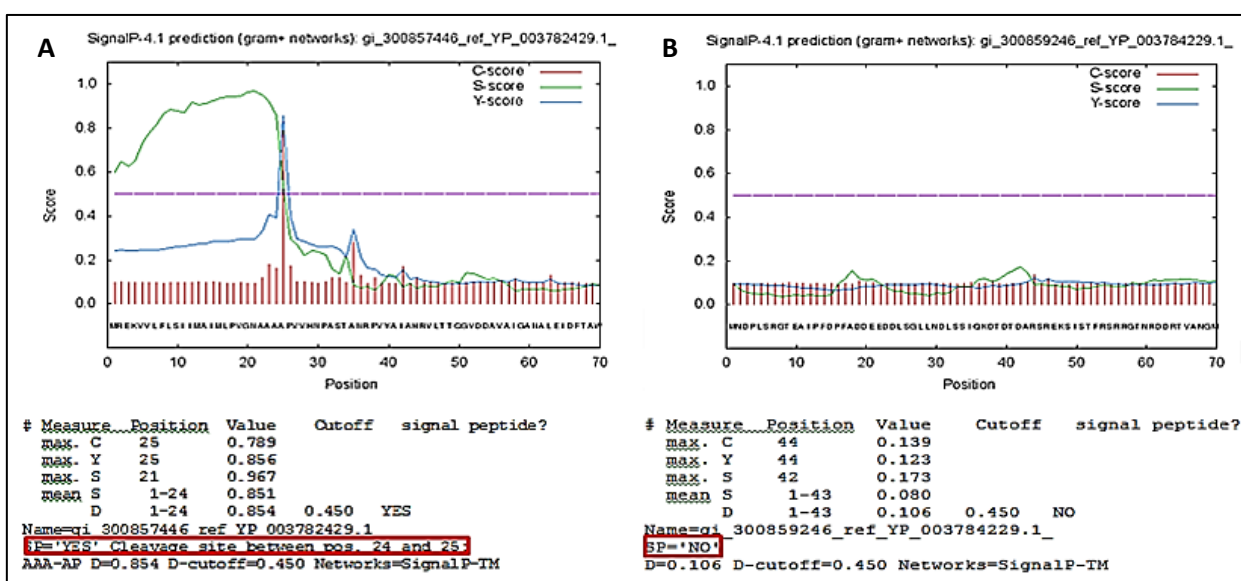
Proteína	Nº de aminoácidos	Peso molecular (kDa)	pI teórico	Fórmula molecular	Número de átomos
PLD	283	31,264	8.77	C <sub>1396</sub> H <sub>2143</sub> N <sub>401</sub> O <sub>412</sub> S <sub>4</sub>	4356
PknG	749	83,3494	5.13	C <sub>3704</sub> H <sub>5833</sub> N <sub>1025</sub> O <sub>1124</sub> S <sub>20</sub>	11706
SpaC	796	85,9649	5.13	C <sub>3777</sub> H <sub>5971</sub> N <sub>1033</sub> O <sub>1231</sub> S <sub>13</sub>	12025
SodC	171	17,6914	5.70	C <sub>766</sub> H <sub>1184</sub> N <sub>224</sub> O <sub>252</sub> S <sub>4</sub>	2430
NanH	663	71,4775	4.94	C <sub>3123</sub> H <sub>4906</sub> N <sub>884</sub> O <sub>1008</sub> S <sub>16</sub>	9937

#### 5.1.2 Predição de peptídeo sinal

Para predição de presença de peptídeo sinal e seu sítio de clivagem foi adotada o *Web-software* Signal 4.0 (Petersen *et al.*, 2011).

De acordo com esta ferramenta, as proteínas PLD, SodC e NanH possuem peptídeo sinal em sua sequência. O gráfico de predição de peptídeo sinal da PLD pode ser visualizado na Figura 6.A, enquanto que os gráficos referentes às proteínas SodC e NanH podem ser consultados no Anexo 2. Nas figuras, a indicação da posição do sítio de clivagem predito do peptídeo sinal está destacada em vermelho. Já as proteínas PknG e SpaC não

tiveram peptídeo sinal predito em suas sequências, conforme pode ser visualizado na Figura 6.B para a PknG e no Anexo 2 para a SpaC.



**Figura 6.** A) Predição de peptídeo sinal da proteína PLD através do SignalP 4.0. B) Predição de peptídeo sinal da proteína PknG através do SignalP 4.0.

### 5.1.3 Predição de localização subcelular

Através da ferramenta PSORTb foi realizada a predição de localização subcelular de cada proteína (Yu *et al.*, 2010). Os resultados estão apresentados na Tabela 6. As proteínas recebem um score que pode variar de 0 a 10 para cada um das localizações subcelulares (citoplasmática, membrana citoplasmática, parede celular, extracelular) sendo que 7,5 é o ponto de corte. Ou seja, a proteína é predita como tendo a localização subcelular para a qual recebeu score acima de 7,5.

**Tabela 6.** Predição de localização subcelular pela ferramenta PSORTb.

Proteína	Número de acesso	Score				Predição final
		Cit.* <sup>1</sup>	Membr. Cit.* <sup>2</sup>	Par. Cel.* <sup>3</sup>	Extracel.* <sup>4</sup>	
PLD	YP_003782429.1	0.00	3.33	3.33	3.33	Desconhecida
PknG	YP_003784229.1	9.89	0.09	0.01	0.02	Citoplasmática
SpaC	YP_003784270.1	0.01	0.01	9.97	0.01	Parede celular
SodC	YP_003783011.1	0.24	0.05	0.80	8.91	Extracelular
NanH	YP_003782786.1	0.00	0.14	0.16	9.70	Extracelular

\*<sup>1</sup>Cit = Citoplasmática;

\*<sup>2</sup>Membr. Cit = Membrana Citoplasmática;

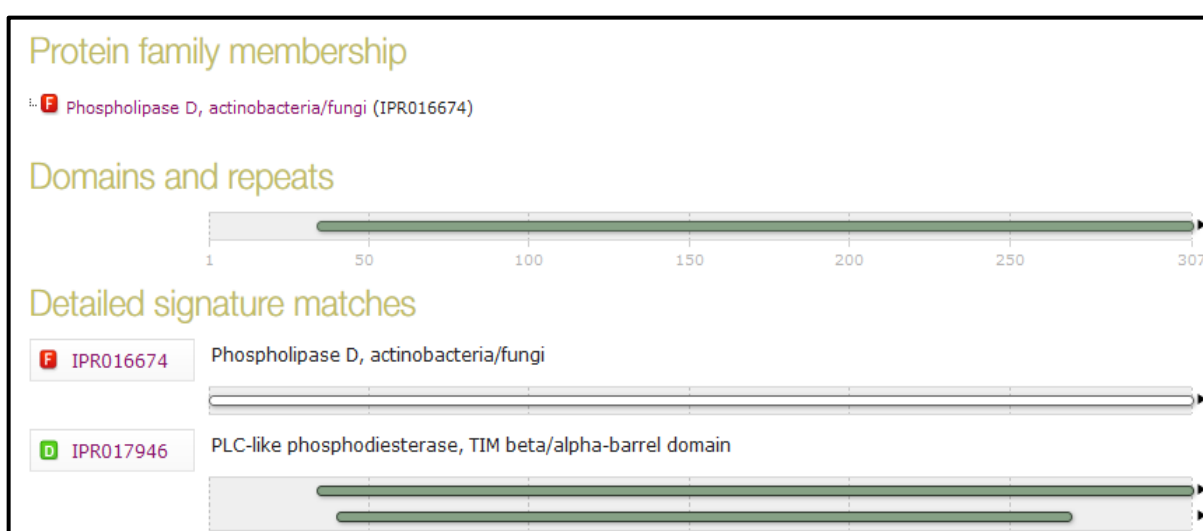
\*<sup>3</sup>Par. Cel.= Parede Celular;

\*<sup>4</sup>Extracel. = Extracelular.

### 5.1.4 Predição de domínios conservados

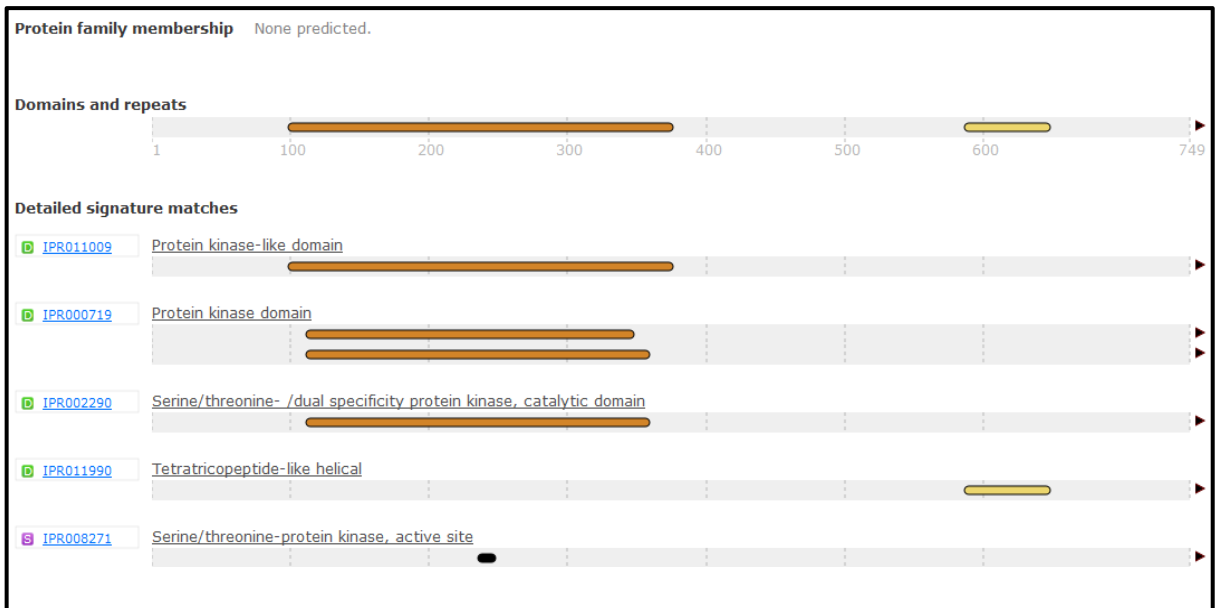
Uma busca por domínios conservados nas sequências protéicas foi realizada através da ferramenta InterProScan 5 (Quevillon *et al.*, 2005).

Na sequência protéica de PLD de *C. pseudotuberculosis* FRC41 foi identificado o domínio “fosfodiesterase PLC-*like*, TIM-Barril”, que se estende do aminoácido 35 ao 307. Este é um domínio estrutural que consiste na estrutura “TIM beta/alfa-Barril” formada por uma fita beta seguida de uma alfa hélice repetidas oito vezes e encontrado nas fosfodiesterases PLD-*like*. A proteína foi também classificada como pertencente à família “Fosfolipase D, actinobacteria/fungi” (Figura 7).



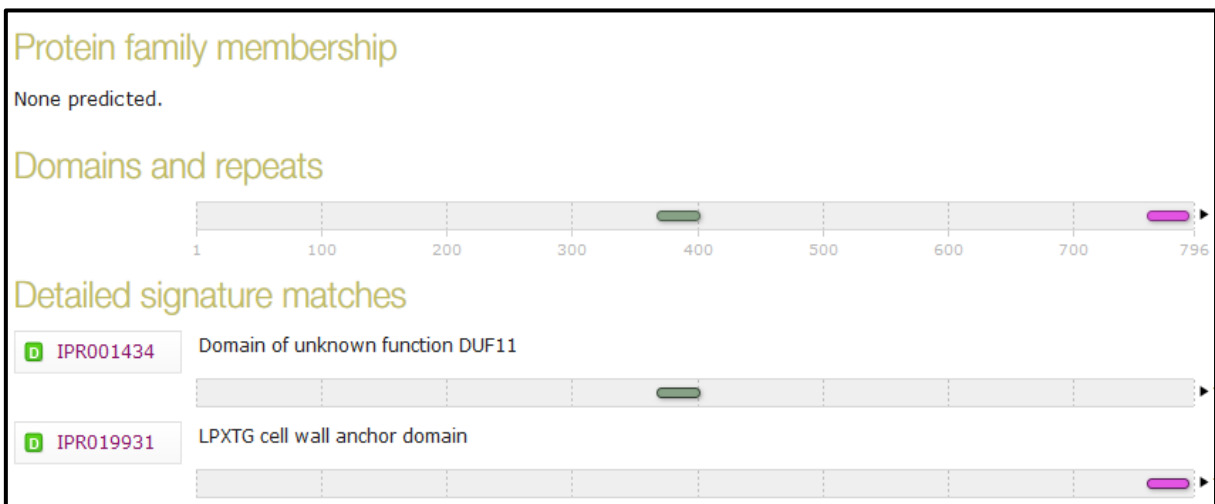
**Figura 7.** Representação esquemática de domínio conservado da proteína PLD de *C. pseudotuberculosis* FRC41. Identifica-se o domínio “PLC-*like* fosfodiesterase, TIM beta/alfa-barril” (em cinza). Análise realizada através do IntreProScan5.

Na sequência protéica de PknG de *C. pseudotuberculosis* FRC41, foi identificado o domínio “proteína quinase” na região compreendida entre os aminoácidos 111 e 347. Identificou-se ainda uma sequência característica de sítio ativo de “serina/treonina proteína quinase” entre os aminoácidos 235 a 247, além do domínio “tetratricopeptídeo-*like* helicoidal” (aminoácidos 586 a 647) (Figura 8).



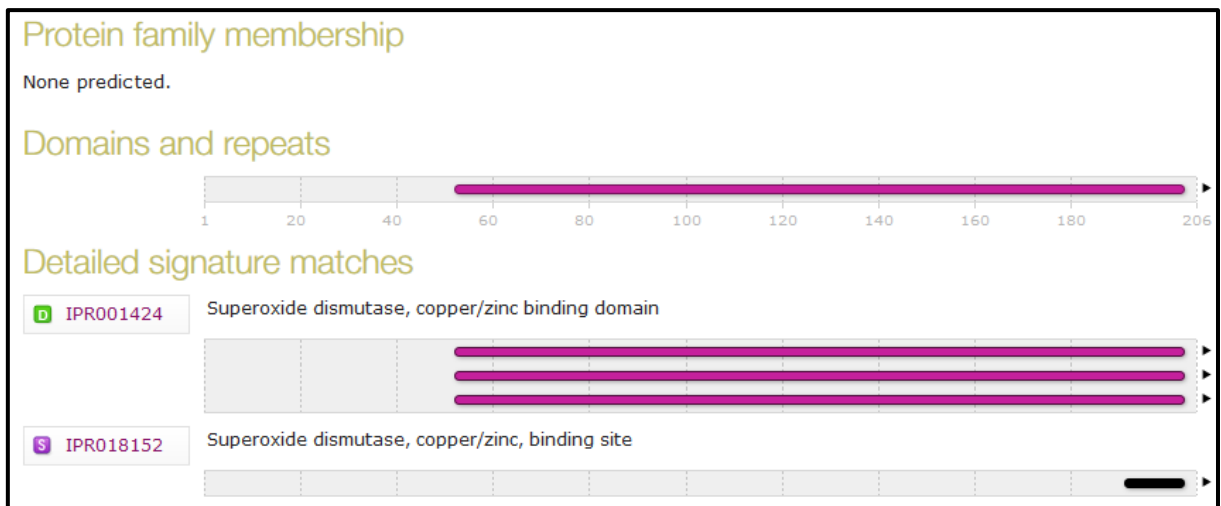
**Figura 8.** Representação esquemática de domínio conservado da proteína PknG de *C. pseudotuberculosis* FRC41. Identifica-se o domínio “proteína quinase” (marrom), domínio “Tetratricopeptídeo-*like* helicoidal” (amarelo) e região sítio ativo “Serina/Treonina proteína quinase (preto). Análise realizada através do InterProScan5.

Na proteína Spac de *C. pseudotuberculosis* FRC41 foram identificados dois domínios. Um domínio de função desconhecida (DUF11) (aminoácidos 369 a 400) e o “domínio de ancoragem a parede celular – LPXTG” (aminoácidos 760 a 790) (Figura 9).



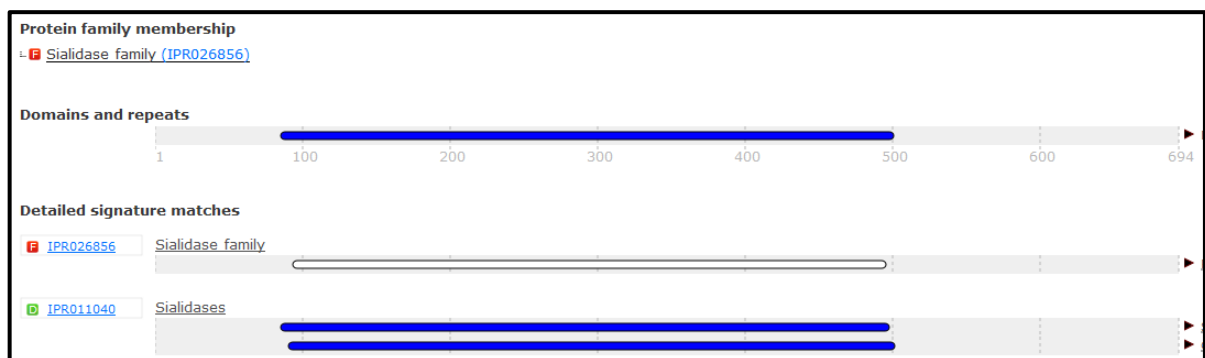
**Figura 9.** Representação esquemática de domínio conservado da proteína SpaC de *C. pseudotuberculosis* FRC41. Identifica-se o domínio de função desconhecida DUF11 (cinza) e o domínio LPXTG (lilás). Análise realizada através do InterProScan5.

A busca por domínios conservados em SodC de *C. pseudotuberculosis* FRC41 revelou a presença de um “domínio de ligação Cu/Zn de superóxido dismutase” (aminoácidos 53-203). Foi identificado também o “sítio de ligação a Cu/Zn de superóxido dismutase” (aminoácidos 192-203) (Figura 10).



**Figura 10.** Representação esquemática de domínio conservado da proteína SodC de *C. pseudotuberculosis* FRC41. Identifica-se o domínio de ligação a Cu/Zn de superóxido dismutase (lilás) e também o sítio de ligação a Cu/Zn (preto). Análise realizada através do IProScan5

Em NanH de *C. pseudotuberculosis* FRC41, foi identificado o domínio das “sialidases” (Aminoácidos 90-500) e a proteína é classificada como pertencente à família “sialidase”. (Figura 11).



**Figura 11.** Representação esquemática de domínio conservado da proteína NanH de *C. pseudotuberculosis* FRC41. Identifica-se o domínio “sialidases” (azul). Análise realizada através da ferramenta IProScan5.

### 5.1.5 Alinhamento de sequências e busca por similaridade

As sequências de PLD, PknG, SpaC, SodC e NanH foram submetidas a uma busca por similaridade através do BLAST no Uniprot a partir do banco de dados “UniprotKB” e “UniprotKB Eukaryota” conforme descrito no Item 4.1.5 (Metodologia). Foram escolhidos microrganismos representantes do grupo CMNR e organismos eucariotos pertencentes aos gêneros *Ovis*, *Bos* e *Equus* e as espécies *Mus musculus* e *Homo sapiens*.



Os alinhamentos tiveram como objetivo avaliar a conservação das sequências nos organismos selecionados. A seguir, são apresentados os resultados obtidos para cada uma das proteínas. Os alinhamentos completos obtidos encontram-se no Anexo 3.

### 5.1.5.1 PLD

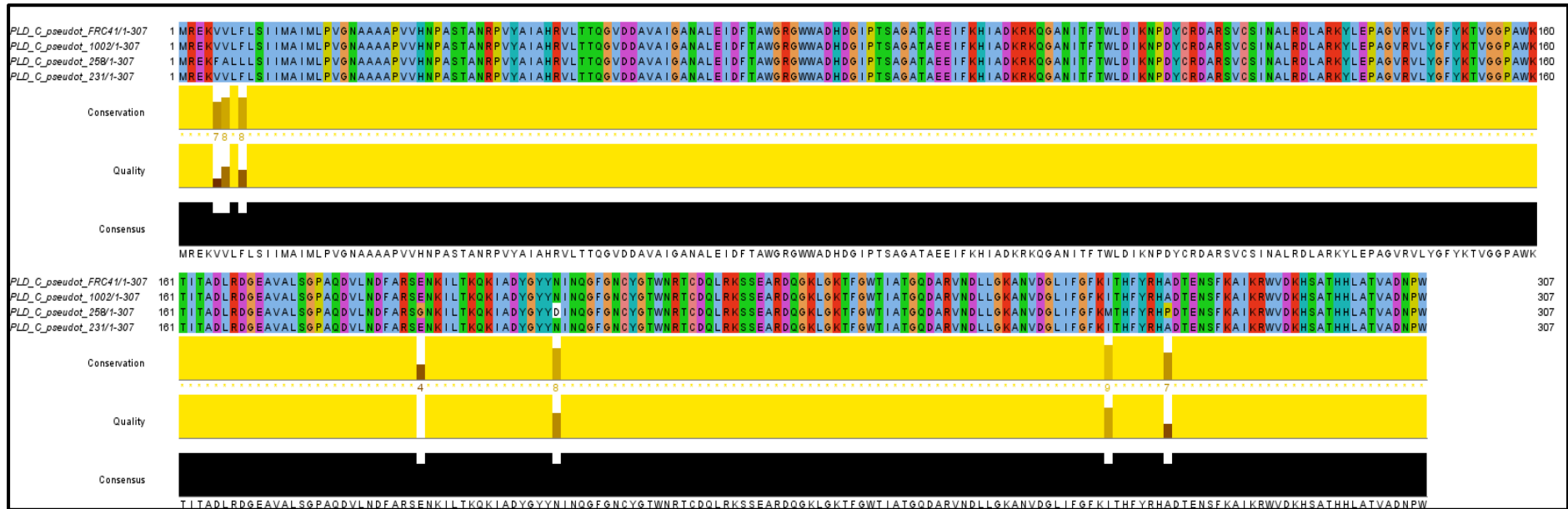
Na busca por sequências de proteínas de microrganismos representantes do grupo CMNR que apresentassem identidade com a proteína PLD de *C. pseudotuberculosis* FRC41, através de BLAST com os parâmetros acima mencionados (*threshold* 0.01 e 500 *hits*), alcançou-se um limite de 40% de identidade. As sequências obtidas a partir desta busca estão listadas na Tabela 7, bem como a porcentagem de identidade apresentada por elas. Observa-se alta identidade entre as sequências identificadas e PknG de *C. pseudotuberculosis* FRC41. O valor de E-value é fornecido tanto para o alinhamento a partir de PLD como para os alinhamentos referentes às demais proteínas que serão apresentados nos tópicos subsequente.

**Tabela 7.** Proteínas de microrganismos do grupo CMNR que apresentaram identidade com a proteína PLD de *C. pseudotuberculosis* FRC41.

Organismo	Proteína	Número de acesso	Identidade com PLD de <i>C. pseudotuberculosis</i> FRC41	E-value
<i>C. pseudotuberculosis</i> 1002	Fosfolipase D	YP_005680306.1	100%	0.0
<i>C. pseudotuberculosis</i> 258	Fosfolipase D	YP_006351625.1	98%	0.0
<i>C. pseudotuberculosis</i> C231	Fosfolipase D	YP_005682396.1	100%	0.0

As sequências de fosfolipase D de *C. pseudotuberculosis* linhagens 1002, 258 e C231 foram alinhadas através do ClustalW e o alinhamento foi visualizado através da ferramenta Jalview (Figura 12). Observa-se que a sequência de PLD é bastante conservada nestas três linhagens. A conservação é representada pelas barras no gráfico de conservação abaixo da representação do alinhamento, onde marrom indica menor conservação e amarelo representa maior conservação. Além disso, os aminoácidos são coloridos de acordo com a sua similaridade. Estas observações são válidas para os demais gráficos de alinhamento gerados pelo ClustalW e visualizados pelo Jalview.

As sequências de PLD das linhagens FRC41, 1002, 258 e C231 de *C. pseudotuberculosis* possuem um tamanho de 307 aminoácidos, considerando toda a proteína, inclusive a região do peptídeo sinal. Destes 307 aminoácidos, apenas 12 não foram alinhados nessas quatro sequências.



**Figura 12.** Alinhamento de seqüências de proteínas de microrganismos do grupo CMNR que apresentaram identidade com a proteína PLD de *C. pseudotuberculosis* FRC41 realizado através do ClustaW e visualizado pelo Jalview. A conservação é representada pelas barras no gráfico de conservação abaixo da representação do alinhamento, onde marrom indica menor conservação e amarelo representa maior conservação. Além disso, os aminoácidos são coloridos de acordo com a sua similaridade.

A busca por sequências que apresentassem identidade com a PLD de *C. pseudotuberculosis* FRC41, quando a busca através do BLAST foi restrita ao banco de dados de organismos eucariotos, não revelou nenhuma sequência homóloga.

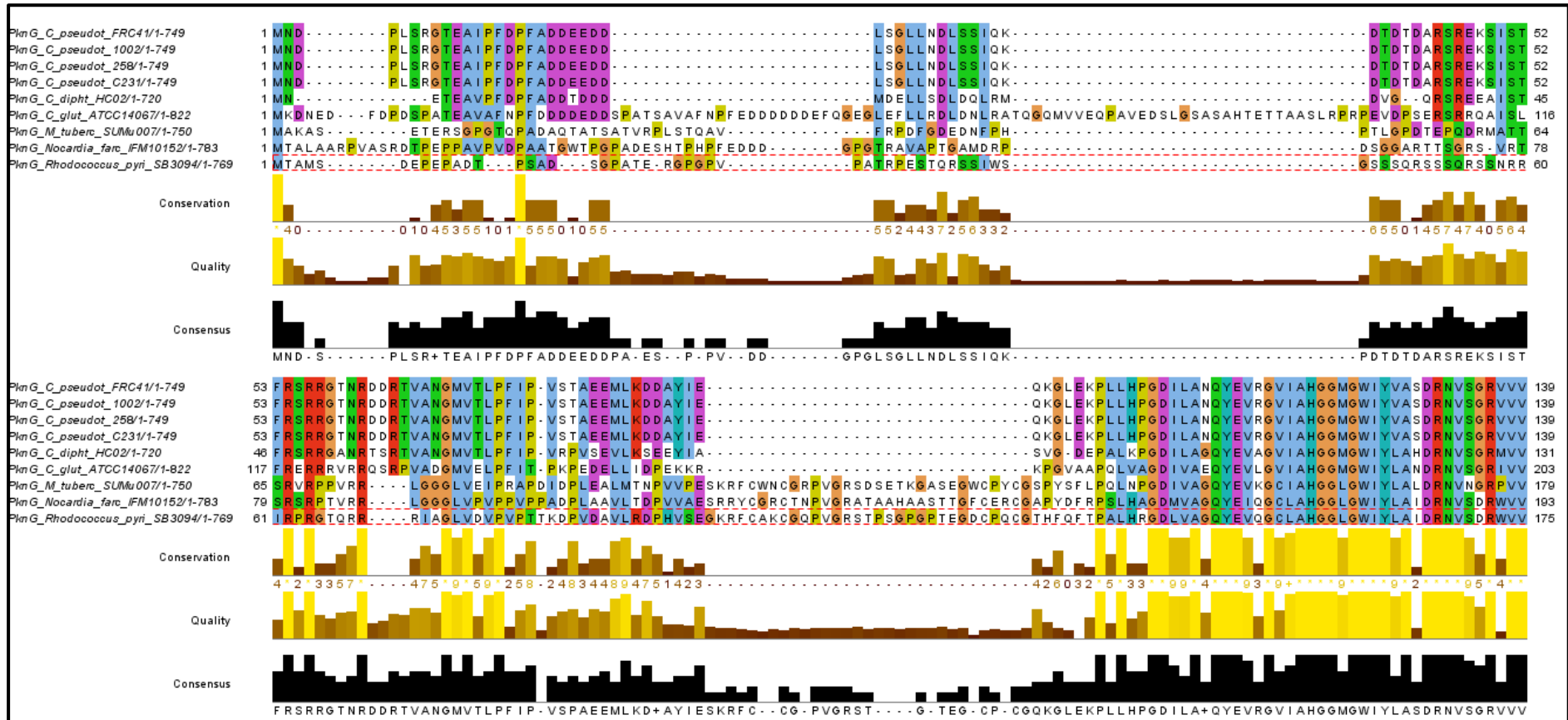
#### 5.1.5.2 PknG

A busca através do BLAST por proteínas de microrganismos representantes do grupo CMNR que apresentassem maior identidade com a proteína PknG de *C. pseudotuberculosis* FRC41 levou à identificação das sequências listadas na Tabela 8. A partir dos parâmetros adotados, alcançou-se um limite de 41% de identidade. Observa-se alta porcentagem de identidade (99 a 100%) entre as sequências de PknG das linhagens 1002, 258 e C231 de *C. pseudotuberculosis*. Além das linhagens de *C. pseudotuberculosis*, foi observada identidade com serina/treonina proteínas quinases de outras espécies do gênero: linhagem HC02 de *C. diptheriae* (74%) e linhagem ATCC14067 de *C. glutamicum* (61%). Serina/treonina proteína quinase de *M. tuberculosis* SuMu007 e *R. pyridinivorans* SB3094 apresentaram 46% de identidade com PknG de *C. pseudotuberculosis* FRC41. Já *N. farcinica* IFM10152 apresentou uma serina-treonina proteína quinase com 41% de identidade com PknG de *C. pseudotuberculosis* FRC41.

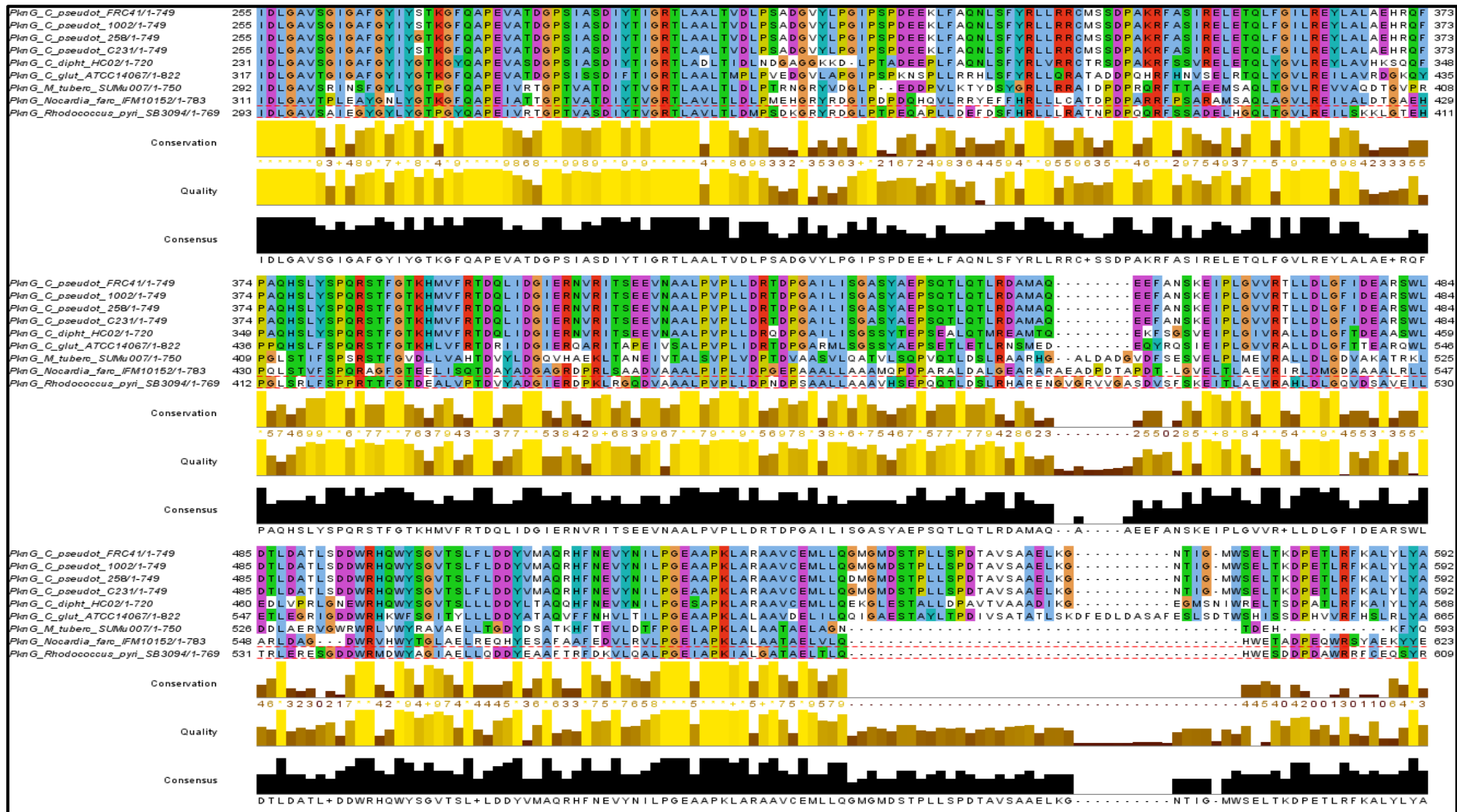
**Tabela 8.** Proteínas de microrganismos do grupo CMNR que apresentaram identidade com a proteína PknG de *C. pseudotuberculosis* FRC41.

Organismo	Proteína	Número de acesso	Identidade com PknG de <i>C. pseudotuberculosis</i> FRC41	E-value
<i>C. pseudotuberculosis</i> 1002	PknG	YP_005682090.1	100%	0.0
<i>C. pseudotuberculosis</i> 258	PknG	YP_006353407.1	99%	0.0
<i>C. pseudotuberculosis</i> C231	PknG	YP_005684181.1	100%	0.0
<i>C. diptheriae</i> HC02	PknG	YP_005165735.1	74%	0.0
<i>C. glutamicum</i> ATCC14067	Serina/treonina proteína quinase	KEI24302.1	61%	0.0
<i>M. tuberculosis</i> SUMu007	PknG	EFP36082.1	46%	1x10 <sup>-169</sup>
<i>N. farcinica</i> IFM10152	Putative serina/treonina proteína quinase	YP_121560.1	41%	4.0x10 <sup>-173</sup>
<i>R. pyridinivorans</i> SB3094	Serina/treonina proteína quinase	YP_008987279.1	46%	1.0x10 <sup>-179</sup>

A partir do alinhamento múltiplo das sequências pelo ClustalW, observa-se que a porção N-terminal dessas proteínas apresenta uma baixa conservação compreendendo a região que se estende aproximadamente até o aminoácido 93 de PknG de *C. pseudotuberculosis* FRC41 (Figura 13). Já as regiões centrais e C-terminais das proteínas são bem conservadas. Na figura 14 é possível visualizar o alinhamento das regiões localizadas próximo às porções C-terminais das proteínas. Na região posterior ao aminoácido de posição 568 de PknG da linhagem *C. pseudotuberculosis* FRC41 observa-se também uma menor conservação da sequência com formação de espaço no alinhamento.



**Figura 13.** Alinhamento de porção N-terminal de sequências de serina/treonina proteínas quinases de microrganismos representantes do grupo CMNR que apresentaram identidade com a proteína PknG de *C.pseudotuberculosis* FRC41. Alinhamento múltiplo de sequências realizado através do ClustaW e visualizado pelo Jalview. A conservação é representada pelas barras no gráfico de conservação abaixo da representação do alinhamento, onde marrom indica menor conservação e amarelo representa maior conservação. Além disso, os aminoácidos são coloridos de acordo com a sua similaridade. Observa-se uma baixa conservação e formação de espaço no alinhamento.



**Figura 14.** Alinhamento de porção central de seqüências de serina/treonina proteínas quinases de microrganismos representantes do grupo CMNR que apresentaram identidade com a proteína PknG de *C. pseudotuberculosis* FRC41. Alinhamento múltiplo de seqüências realizado através ClustaW e visualizado pelo Jalview. A conservação é representada pelas barras no gráfico de conservação abaixo da representação do alinhamento, onde marrom indica menor conservação e amarelo representa maior conservação. Além disso, os aminoácidos são coloridos de acordo com a sua similaridade. Observa-se regiões de alta conservação.

Buscando pelas sequências de proteínas dos organismos eucariotos listados no item 5.1.5, os organismos eucariotos que apresentaram sequências de proteínas com identidade com PknG de *C. pseudotuberculosis* FRC41 foram: *Ovis aries*, *Bos taurus*, *Equus caballus*, *Mus musculus* e *Homo sapiens* e atingiu-se um limite de 29% de identidade. As proteínas eucarióticas em questão apresentaram por volta de 30% de identidade com PknG de *C. pseudotuberculosis* FRC41 (Tabela 9).

**Tabela 9.** Proteínas de organismos eucariotos que apresentaram identidade com a proteína PknG de *C. pseudotuberculosis* FRC41.

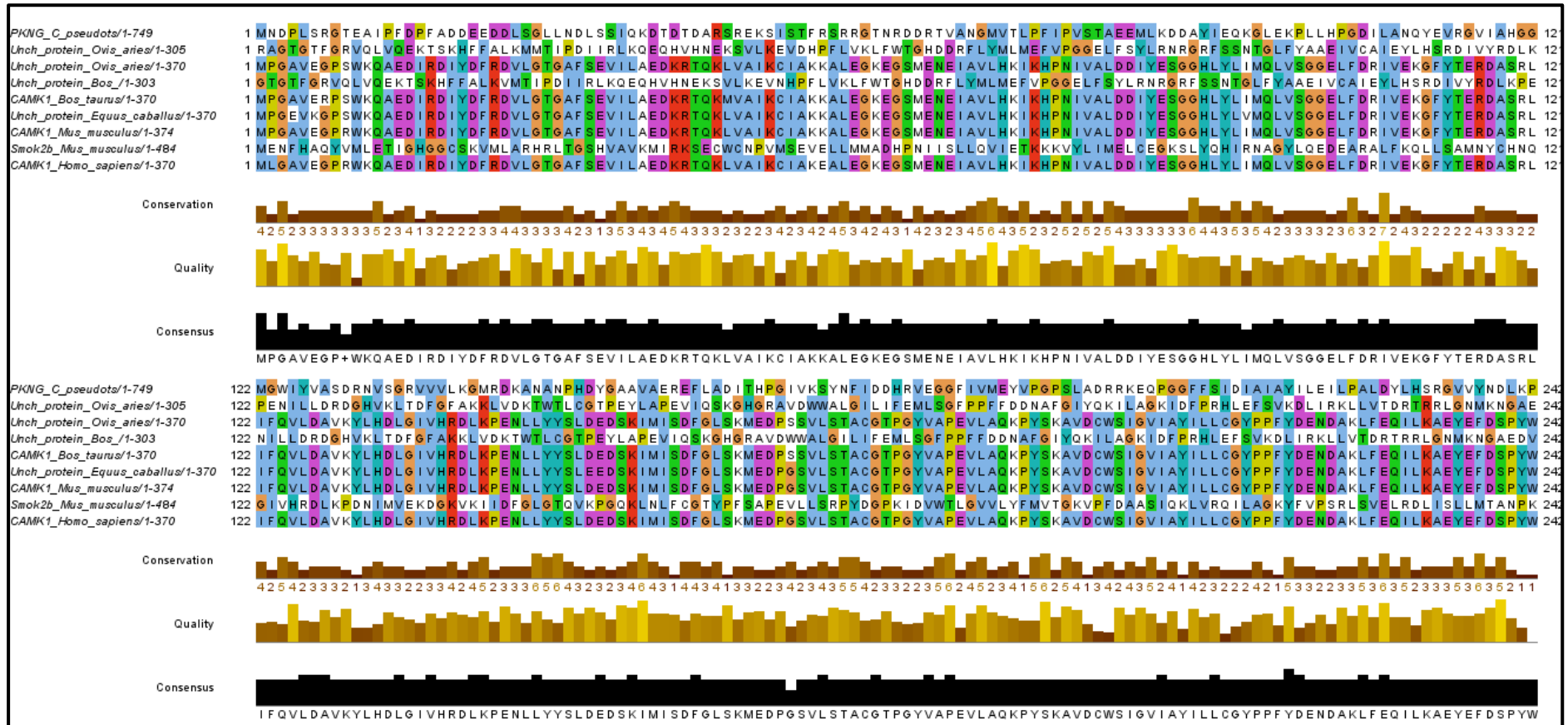
Organismo	Proteína	Número de acesso	Identidade com PknG de <i>C. pseudotuberculosis</i> FRC41	E-value
<i>Ovis aries</i>	Unch. Prot. <sup>*1</sup> PRKX <sup>*2</sup>	XP_004022575	30%	4x10 <sup>-12</sup>
<i>Ovis aries</i>	Unch. Prot. <sup>*1</sup> CAMK1 <sup>*3</sup>	XP_004018333	29%	4x10 <sup>-12</sup>
<i>Bos Taurus</i>	Unch. Prot. <sup>*1</sup> PRKX <sup>*2</sup>	XP_002700469	30%	3x10 <sup>-12</sup>
<i>Bos Taurus</i>	CAMK1 <sup>*3</sup>	NP_001070336	29%	1x10 <sup>-12</sup>
<i>Equus caballus</i>	Unch. Prot. <sup>*1</sup> CAMK1 <sup>*3</sup>	XP_005600446.1	31%	9x10 <sup>-13</sup>
<i>Mus musculus</i>	CAMK1 <sup>*3</sup>	NP_598687	29%	3x10 <sup>-13</sup>
<i>Mus musculus</i>	Smok2b <sup>*4</sup>	NP_001161385	31%	1x10 <sup>-13</sup>
<i>Homo sapiens</i>	CAMK1 <sup>*3</sup>	NP_003647	29%	2x10 <sup>-12</sup>

\*<sup>1</sup> Unch. Prot.- Uncharacterized protein – Proteína não caracterizada

\*<sup>2</sup> PRKX- Serina/treonina proteína quinase X

\*<sup>3</sup> CAMK1- Proteína quinase I Cálcio/calmodulina dependente

A partir do alinhamento múltiplo de sequências, observa-se uma baixa conservação entre as sequências. Além disso, a proteína PknG possui uma sequência composta por 784 aminoácidos, enquanto as sequências das proteínas eucarióticas alinhadas possuem tamanhos que variam de 303 a 484 aminoácidos. A porção N-terminal de PknG de *C. pseudotuberculosis* se alinha com as demais sequências, porém, apresenta baixa conservação (Figura 15). A menor conservação é observada na região C-terminal, a partir do aminoácido de posição 370 (Anexo 3).



**Figura 15.** Alinhamento de porção N-terminal de seqüências proteicas de organismos eucariotos que apresentaram identidade com a proteína PknG de *C. pseudotuberculosis* FRC41. Alinhamento múltiplo de seqüências realizado através do ClustaW e visualizado pelo Jalview. A conservação é representada pelas barras no gráfico de conservação abaixo da representação do alinhamento, onde marrom indica menor conservação e amarelo representa maior conservação. Além disso, os aminoácidos são coloridos de acordo com a sua similaridade. Observa-se uma baixa conservação das seqüências.



### 5.1.5.3 SpaC

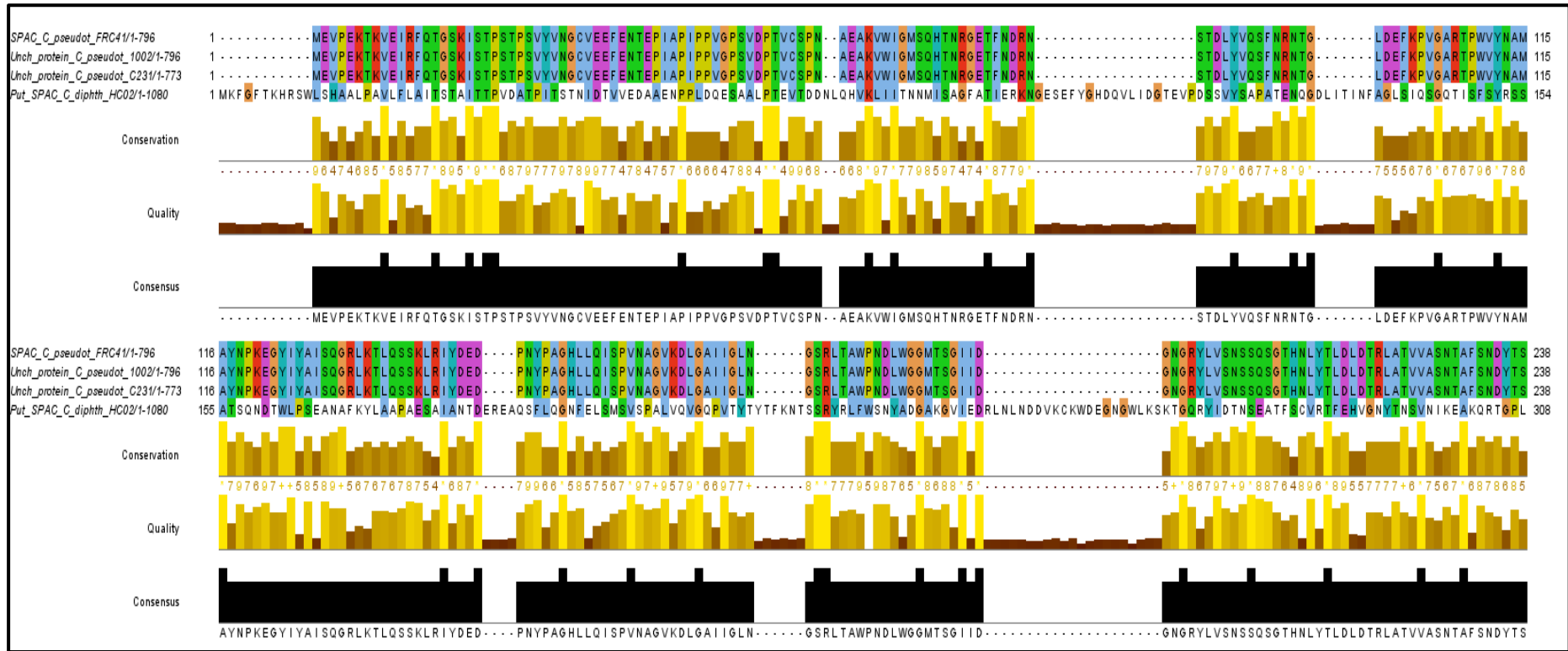
A busca através do BLAST por proteínas de microrganismos representantes do grupo CMNR que apresentassem identidade com SpaC de *C. pseudotuberculosis* FRC41 alcançou um limite de 26% de identidade. Foram identificadas a partir desta busca, sequências protéicas pertencentes aos seguintes microrganismos: *C. pseudotuberculosis* 1002, *C. pseudotuberculosis* C231 e *C. diptheriae* HC02. Observou-se alta identidade entre PknG de *C. pseudotuberculosis* FRC41 e as proteínas de *C. pseudotuberculosis* linhagens 1002 e C231. Já a proteína de *C. diptheriae* identificada, apresentou identidade de 27% com SpaC de *C. pseudotuberculosis* FRC41 (Tabela 10).

**Tabela 10.** Proteínas de microrganismos do grupo CMNR que apresentaram identidade com a proteína SpaC de *C. pseudotuberculosis* FRC41.

Organismo	Proteína	Número de acesso	Identidade com SpaC de <i>C. pseudotuberculosis</i> FRC41	E-value
<i>C. pseudotuberculosis</i> 1002	Uncharacterized protein	YP_005682125	99%	0.0
<i>C. pseudotuberculosis</i> C231	Uncharacterized protein	YP_005684218	100%	0.0
<i>C. diptheriae</i> HC02	Subunidade fimbrial ancorada a superfície	YP_005165747	27%	5x10 <sup>-15</sup>

A partir do alinhamento múltiplo das sequências listadas na Tabela 10, revelou-se que essas proteínas são bem conservadas nestes microrganismos, principalmente quando se compara as espécies de *C. pseudotuberculosis*, que possuem também tamanho semelhante (Figura 16). Já a proteína *C. diptheriae* possui 1080 aminoácidos e apresenta menor identidade com SpaC de *C. pseudotuberculosis* FRC41.

Identidade entre SpaC de *C. pseudotuberculosis* FRC41 e eucariotos não foi identificada.



**Figura 16.** Alinhamento de porção N-terminal de seqüências proteicas de microrganismos representantes do grupo CMNR que apresentaram identidade com a proteína SpaC de *C. pseudotuberculosis* FRC41. Alinhamento múltiplo de seqüências realizado através do ClustaW e visualizado pelo Jalview. A conservação é representada pelas barras no gráfico de conservação abaixo da representação do alinhamento, onde marrom indica menor conservação e amarelo representa maior conservação. Além disso, os aminoácidos são coloridos de acordo com a sua similaridade. Observa-se alta conservação entre as seqüências.

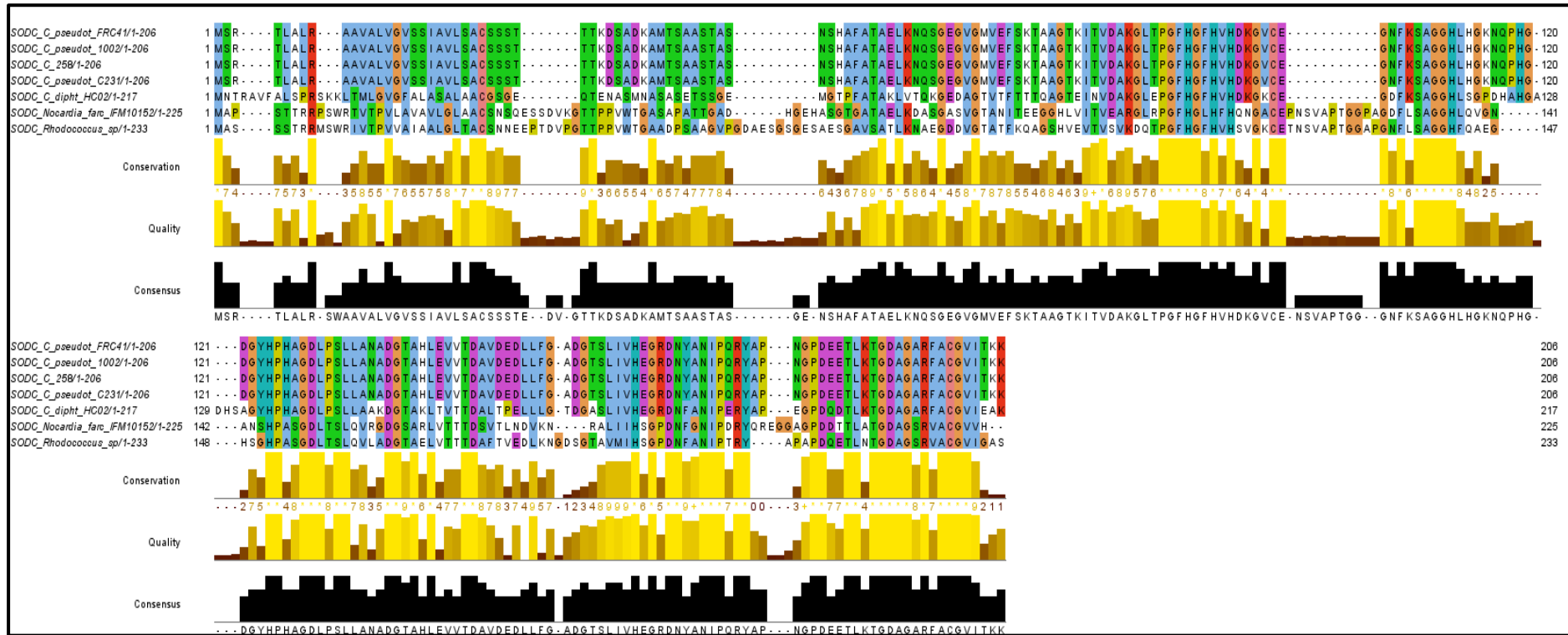
#### 5.1.5.4 SodC

A busca por proteínas de microrganismos do grupo CMNR, que apresentassem identidade com SodC de *C. pseudotuberculosis* FRC41, levou à identificação de sequências pertencentes aos seguintes organismos: *C. pseudotuberculosis* 1002; *C. pseudotuberculosis* 258; *C. pseudotuberculosis* C231; *C. diphtheriae* HC02; *M. tuberculosis* SUMuu007; *N. farcinica* IFM10152; *Rhodococcus* sp. RHA1 (Tabela 11). Observou-se alta identidade de SodC de *C. pseudotuberculosis* FRC41 com as linhagens 1002, 258 e C231 de *C. pseudotuberculosis*, enquanto que os demais membros apresentaram menor identidade. Foi atingido um limite de 40% de identidade.

**Tabela 11.** Proteínas de microrganismos do grupo CMNR que apresentaram identidade com a proteína SodC de *C. pseudotuberculosis* FRC41.

Organismo	Proteína	Número de acesso	Identidade com SodC de <i>C. pseudotuberculosis</i> FRC41	E-value
<i>C. pseudotuberculosis</i> 1002	SodC	ADL20500	100%	$3 \times 10^{-122}$
<i>C. pseudotuberculosis</i> 258	SodC	YP_006352192	100%	$3 \times 10^{-122}$
<i>C. pseudotuberculosis</i> C231	SodC	YP_005682971	100%	$3 \times 10^{-122}$
<i>C. diphtheriae</i> HC02	SodC	YP_005165965	68%	$2 \times 10^{-75}$
<i>N. farcinica</i> IFM 10152	SodC	YP_121514	44%	$4 \times 10^{-35}$
<i>Rhodococcus</i> sp. RHA1	SodC	YP_702138	49%	$2 \times 10^{-42}$

Através do alinhamento das sequências listadas na Tabela 11 com a proteína SodC de *C. pseudotuberculosis* FRC41, foi possível observar que essas sequências são bem conservadas (Figura 17).



**Figura 17.** Alinhamento de porção N-terminal de seqüências proteicas de microrganismos representantes do grupo CMNR que apresentaram identidade com a proteína SodC de *C. pseudotuberculosis* FRC41. Alinhamento múltiplo de seqüências realizado através do ClustaW e visualizado pelo Jalview. A conservação é representada pelas barras no gráfico de conservação abaixo da representação do alinhamento, onde marrom indica menor conservação e amarelo representa maior conservação. Além disso, os aminoácidos são coloridos de acordo com a sua similaridade. Observa-se alta conservação entre as seqüências.

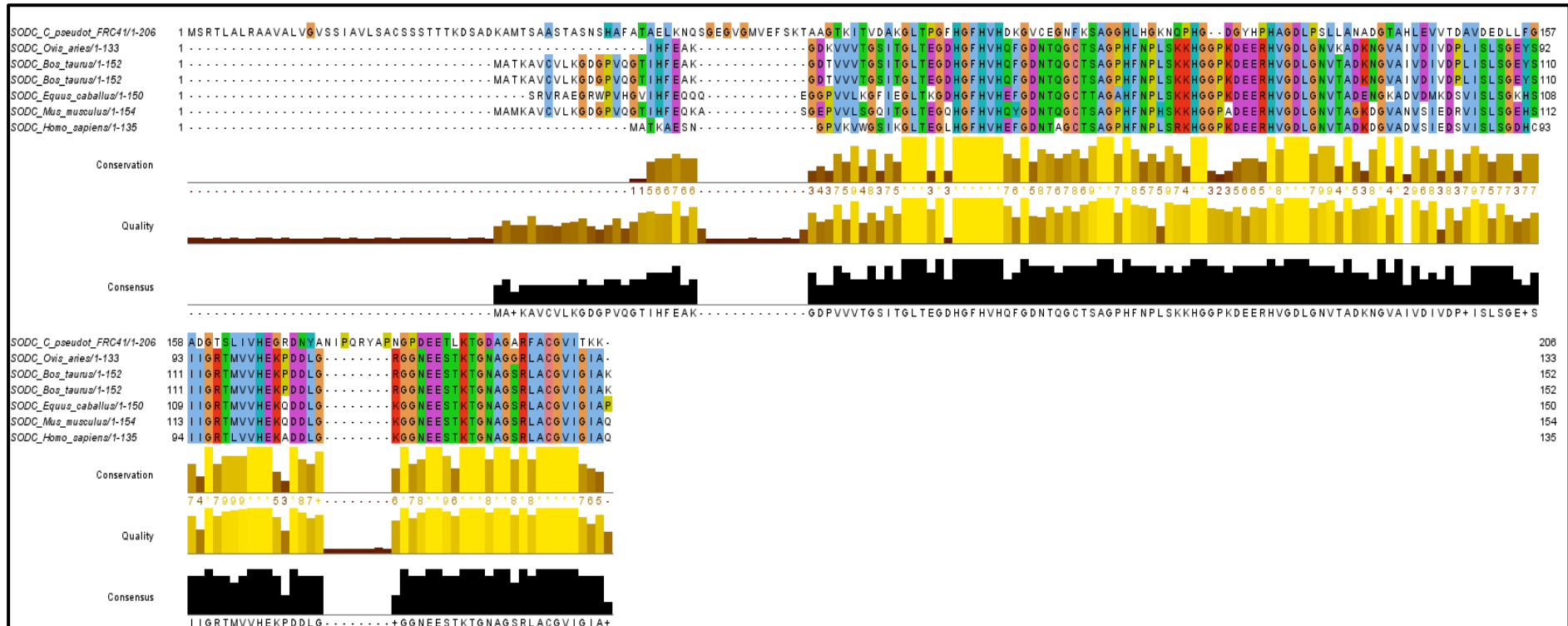
Na busca por sequências proteicas de organismos eucariotos que também apresentassem similaridade com SodC de *C. pseudotuberculosis* FRC41 foram identificadas sequências pertencentes aos seguintes organismos: *Ovis aries*, *Bos taurus*, *Equus caballus*, *Mus musculus* e *Homo sapiens*. Foi atingido um limite de 34% de identidade.

Essas sequências estão listadas na Tabela 12.

**Tabela 12.** Proteínas de organismos eucariotos que apresentaram identidade com a proteína SodC de *C. pseudotuberculosis* FRC41.

Organismo	Proteína	Número de acesso	Identidade com SodC de <i>C. pseudotuberculosis</i> FRC41	E-value
Ovis aries	Sod	AER27859.1	36%	$3 \times 10^{-15}$
Bos taurus	Sod1	P00442.2	34%	$2 \times 10^{-15}$
Equus caballus	Sod1	P00443.2	32%	$8 \times 10^{-15}$
Mus musculus	Sod1	P08228.2	34%	$2 \times 10^{-16}$
Homo sapiens	Sod1	NP_000445.1	34%	$3 \times 10^{-16}$

A partir do alinhamento dessas sequências foi possível observar conservação entre elas (Figura 18). Porém, a proteína SodC de *C. pseudotuberculosis* FRC41 é maior que as outras proteínas que são pertencentes a eucariotos, possuindo 206 aminoácidos em sua sequência enquanto que as outras proteínas possuem um tamanho variando entre 133 e 154 aminoácidos. A região N-terminal de SodC de *C. pseudotuberculosis* FRC41, até o aminoácido de posição 70, é uma região pouco conservada.



**Figura 18.** Alinhamento de seqüências proteicas de organismos eucariotos que apresentaram identidade com a proteína SodC de *C. pseudotuberculosis* FRC41. Alinhamento múltiplo de seqüências realizado através do ClustaW e visualizado pelo Jalview. A conservação é representada pelas barras no gráfico de conservação abaixo da representação do alinhamento, onde marrom indica menor conservação e amarelo representa maior conservação. Além disso, os aminoácidos são coloridos de acordo com a sua similaridade. Observa-se alta conservação entre as seqüências.

### 5.1.5.5 NanH

Ao buscar por sequências protéicas dos microrganismos representantes do grupo CMNR que apresentassem identidade com NanH de *C. pseudotuberculosis* FRC41, encontrou-se sequências pertencentes aos seguintes microrganismos: *C. pseudotuberculosis* 1002, *C. pseudotuberculosis* C231, *C. diptheriae* HC02, *C. glutamicum* ATCC 14067 (Tabela 13). A identidade entre as sequências de NanH das linhagens 1002, 258 e C231 de *C. pseudotuberculosis* e a da linhagem FRC41 é de 100%. A identidade de NanH de *C. diptheriae* HC02 com NanH de *C. pseudotuberculosis* FRC41 é de 50%. *C. glutamicum* apresentou uma sequência de uma proteína não caracterizada com 50% de identidade com NanH de *C. pseudotuberculosis* FRC41. Alcançou-se um limite de 26% de identidade.

**Tabela 13.** Proteínas de microrganismos do grupo CMNR que apresentaram identidade com a proteína NanH de *C. pseudotuberculosis* FRC41.

Organismo	Proteína	Número de acesso	Identidade com NanH de <i>C. pseudotuberculosis</i> FRC41	E-value
<i>C. pseudotuberculosis</i> 1002	NanH	YP_005680659.1	100%	0.0
<i>C. pseudotuberculosis</i> C231	NanH	YP_005682747.1	100%	0.0
<i>C. diptheriae</i> HC02	NanH	YP_005164275.1	50%	0.0
<i>C. glutamicum</i> ATCC 14067	Uncharacterized protein	KEI23275.1	36%	1x10 <sup>-55</sup>

A partir do alinhamento múltiplo de sequências através do ClustalW, observa-se que as sequências são bem conservadas entre estes organismos. Porém, a proteína de *C. glutamicum* possui 384 aminoácidos, enquanto que as outras quatro sequências possuem por volta de 700 aminoácidos (Figura 19).





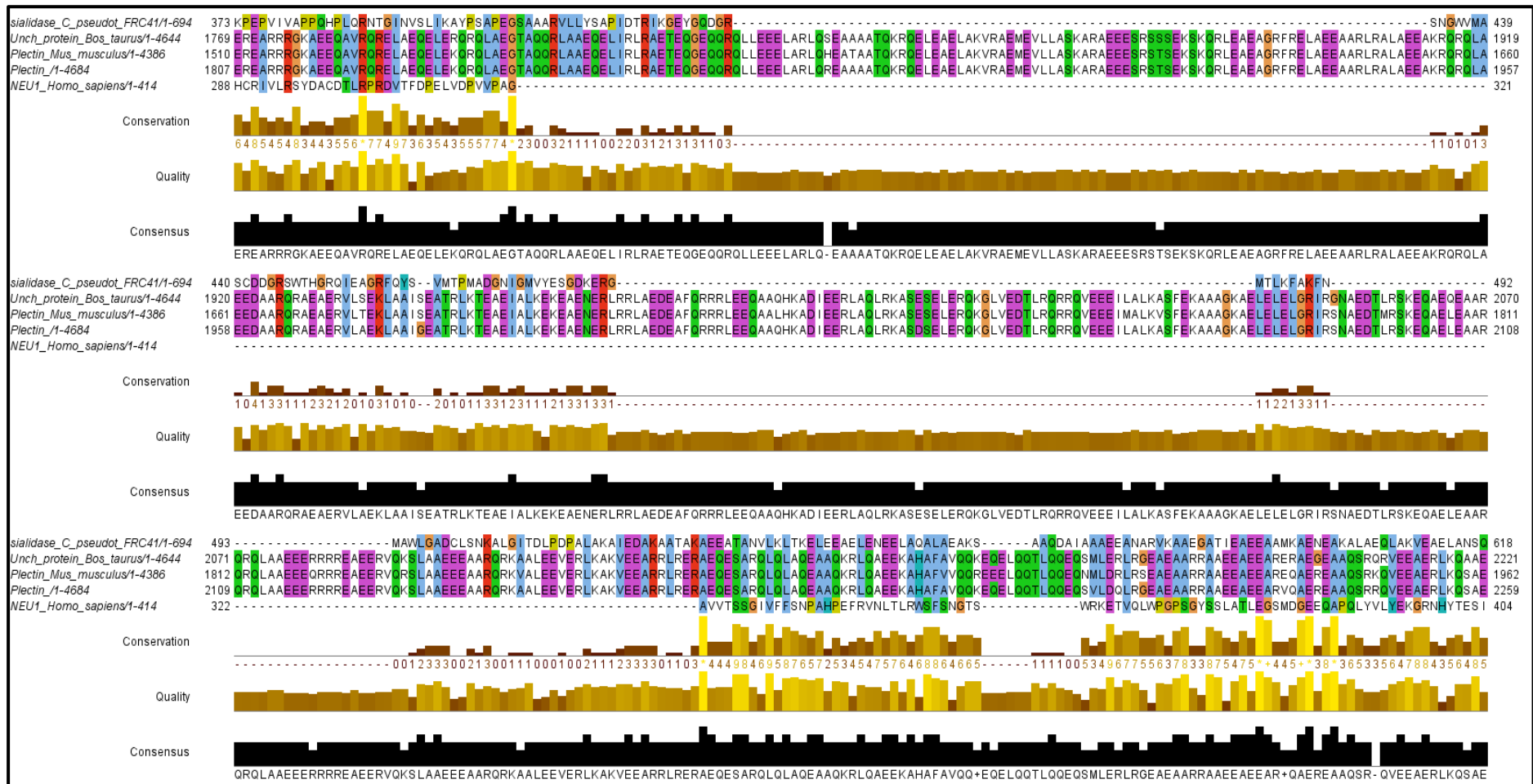
Ao buscar proteínas de organismos eucariotos que apresentassem identidade com NanH de *C. pseudotuberculosis* FRC41, identificou-se sequências pertencentes aos seguintes organismos: *Bos taurus*, *Mus musculus* e *Homo sapiens*. As proteínas identificadas apresentaram identidade de cerca de 30% e o limite de identidade alcançado foi de 25% (Tabela 14).

**Tabela 14.** Proteínas de organismos eucariotos que apresentaram identidade com a proteína NanH de *C. pseudotuberculosis* FRC41.

Organismo	Proteína	Número de acesso	Identidade com NanH de <i>C. pseudotuberculosis</i> FRC41	E-value
<i>Bos taurus</i>	Uncharacterized protein Plectina		30%	$1 \times 10^{-8}$
<i>Mus musculus</i>	Plectina	NP_001157675.1	30%	$2 \times 10^{-8}$
<i>Homo sapiens</i>	Plectina	NP_958782.1	30%	$2 \times 10^{-9}$
<i>Homo sapiens</i>	NEU1*	CAG32998.1	27%	$2 \times 10^{-7}$

\*Sialidase lisossomal

A partir do alinhamento das sequências de proteínas eucarióticas que apresentaram identidade com NanH de *C. pseudotuberculosis* FRC41, observa-se que as mesmas são pouco conservadas. Além disso, existe uma grande diferença de número de aminoácidos entre as sequências. As sequências de Plectina de *Bos taurus*, *Mus musculus* e *Homo sapiens* apresentam sequências com mais de 4300 aminoácidos, enquanto que a NanH de *C. pseudotuberculosis* FRC41 possui 694 aminoácidos e NEU1 de *Homo sapiens* possui 414 aminoácidos. A porção de NanH que exibe menor conservação é a que vai do aminoácido 432 a 519, aproximadamente (Figura 20).



**Figura 20.** Alinhamento de seqüências proteicas de organismos eucariotos que apresentaram identidade com a proteína NanH de *C. pseudotuberculosis* FRC41. Alinhamento múltiplo de seqüências realizado através do ClustaW e visualizado pelo Jalview. A conservação é representada pelas barras no gráfico de conservação abaixo da representação do alinhamento, onde marrom indica menor conservação e amarelo representa maior conservação. Além disso, os aminoácidos são coloridos de acordo com a sua similaridade. Observa-se região onde há baixa conservação e menor alinhamento da seqüência de NanH de *C. pseudotuberculosis* FRC41 e as outras seqüências.

### 5.1.6 Predição de epítomos

A ferramenta *Mature Epitope Density* (MED)1.0 foi utilizada para prever a densidade de epítomos lineares de 9-mer do Complexo de Histocompatibilidade Principal (MHC) de Classe I nas estruturas maduras das proteínas alvos deste estudo. Os resultados obtidos através desta ferramenta para as porções maduras completas das proteínas PknG, Pld, SpaC, SodC e NanH de *C. pseudotuberculosis* FRC41 estão dispostos na Tabela 15. A proteína PknG foi a que apresentou maior densidade de epítomos (12,67), enquanto as outras proteínas apresentaram densidades medianas.

**Tabela 15.** Predição de densidade de epítomos de MHCI nas proteínas PLD, PknG, SpaC, SodC e NanH de *C. pseudotuberculosis* FRC41 através da ferramenta MED.

Proteína	MED- score
PLD	6,78
PknG	12,67
SpaC	7,14
SodC	6,38
NanH	7,84

Nos casos onde a proteína apresentou identidade com proteínas de organismos eucariotos (Item 5.1.2), foi avaliada também a porção de menor conservação quanto a densidade de epítomos, situação observada nas proteínas PknG, SodC e NanH. Estas observações são importantes para uma futura abordagem de desenvolvimento de vacina,

Para a proteína PknG, selecionamos a porção correspondente ao aminoácido 370 a 749. Em SodC, foram selecionados os aminoácidos 1 a 70 e na proteína NanH, selecionou-se a porção que vai do aminoácido 432 a 519. As porções selecionadas correspondem às regiões de menor conservação das proteínas. A porção de SodC compreende o peptídeo sinal, que por sua vez foi removido durante a execução do MED. As porções selecionadas de PknG e de SodC apresentaram uma densidade de epítomos superior à apresentada pela proteína completa. Já a porção selecionada de NanH manteve o MED-score muito próximo ao calculado para a proteína completa (Tabela 16).

**Tabela 16.** Predição de densidade de epítomos de MHC I nas regiões de menor conservação das proteínas PknG, SodC e NanH através da ferramenta MED.

Proteína (região)	MED- score
PknG (370-749)	15,05
SodC (1-70)	10,05
NanH (432-519)	7,72

Para a predição de epítomos de MHC de classe II foi adotada a ferramenta NetMHCII 2.2 (Nielsen & Lund, 2009). Realizou-se uma avaliação quantitativa de epítomos de ligação a MHC de classe II na qual foi somado o número de peptídeos classificados como peptídeos de forte ligação através da ferramenta NetMHCII. Na Tabela 17, são apresentados: o número ligações possíveis de peptídeos de 15-mer aos diferentes alelos de MHCII (possibilidades de ligação - p), o número de peptídeos que foram classificados como peptídeos de forte ligação (n) e o percentual de peptídeos de forte ligação.

**Tabela 17.** Predição quantitativa de epítomos de MHCII nas proteínas PLD, PknG, SpaC, SodC e NanH de *C. pseudotuberculosis* FRC41 através da ferramenta NetMHCII.

Proteína	Possibilidades de ligação (p)	Peptídeos de forte ligação (n)	Percentual de peptídeos de forte ligação
PLD	7532	227	3%
PknG	20580	1032	5%
SpaC	21896	557	2%
SodC	4396	73	2%
NanH	18172	485	3%

Também foi realizada a predição de epítomos de MHCII para as porções das proteínas PknG, SodC e NanH de menor conservação nos eucariotos investigados (Tabela 18).

**Tabela 18.** Predição quantitativa de epítomos de MHCII nas regiões de menor conservação das proteínas PknG, SodC e NanH através da ferramenta NetMHCII.

<b>Proteína (região)</b>	<b>Possibilidades de ligação (p)</b>	<b>Peptídeos de forte ligação (n)</b>	<b>Percentual de peptídeos de forte ligação</b>
PknG (370-740)	10220	458	4%
SodC (1-70)	588	14	2%
NanH (432-519)	2072	98	4%

Nota-se que o percentual de peptídeos de forte ligação manteve a mesma ordem de grandeza para as sequências completas das proteínas e para as porções de menor conservação em eucariotos.

Para a predição de epítomos lineares de células B, foi utilizada a ferramenta BcePred (Saha & Raghava, 2006). Os epítomos preditos foram mapeados nas sequências das proteínas PLD, PknG, SpaC, SodC e NanH bem como nas porções de menor conservação de PknG, SodC e NanH. Os resultados podem ser visualizados no Quadro 1. As sequências destacadas em negrito correspondem ao epítopo predito.

**Quadro 1.** Predição de epítomos lineares de células B através da ferramenta BcePred.

**A)**<sup>1</sup>APVVHN**PASTANRPVY**AIASRVLTQGVDDAVAIGANALEIDFTAWGRGWADHDGIPTSAGATAEEIF**KHIADKRKQGANITFTWLDIKNP  
DYCRDARS**VCSINAL**RDLARKYLEPAG**VRVLYGFYKTVGGPAWKTTADLRDGEAVALSGPAQDVL**NDFARSENKILTKQKIADYGYYNINQG  
FGNCYGTWNRTCDQLRKSSEARDQGKLGKTF**GWTIATG**QDARVND**LLGKANVDGLIFGF**KITHFYRHADTENSFKAIKRWVDKHSATHHLA  
TVADNPW**<sup>283</sup>

**B)**<sup>1</sup>**MNDPLSRG**TEAIPF**DPFADDEEDDLS**GLLNDL**SSIQKDTDTDARSREKSISTFRSRRGTNRDDRTVAN**GMVTLFPFIPVST**AEEMLKDDAYIE  
KQGLEKPLLHPGDILANQYEVRG**VIAHGGMGWI**VASDRNVSGRVVV**LKGM**RDKANANPHDY**GAAVAEREFLADITHPGIVKSYNFID**DHRV  
EGGFIVMEYVPGPSLADRRKEQPGGFF**SIDIAIAYILEILPALDYLHSRGV**VYNDLKPDNI**IVAEDQVKLIDLGAVSGIGAFGYI**YSTKGFQAPEVA  
TDGPSIASDIYTIGRTLAALTVDLPSADGVYLPGIPSPDEEKLF**AQNLSFY**RLRRCMSSDPAKR**FASIRELETQLFGILREYLAL**AEHRQFPAQ  
HSLYSPQRSTFGTK**HMV**FRTDQLIDGIERNVRIT**SEEVNAALPV**LLDRTDPGAILISGAS**YA**EPSQTLQTLRDAMAQEEFANSKEIP**LGVV**RTL  
LDLGFIDEARSWLDTLDATL****SDDWRHQWYSG**VTSLFLDDYVM**AQRHFNEVY**NILPGEAAPKLARA**AVCEMLLQGMGMDSTPLLSPDTAVSAA  
ELKGNTIGM****WSELTKDPETLRFK**ALYLYAL  
**VWRTNP**TTVSSAFGLARQLVAENQVDLAVST**LDQVPQNSTHRRMAELT**AILHLISGD**LSESRI**RAARR**LESIPTNEPRFLQ**IKIAIMNAAL**TWL  
RQSRLEAAAADNDLFDYRFTQV**GLRTGLYDSLRL**LARSAPNVHHR**YTLVDMAN**QVRPMT**WF<sup>749</sup>

**C)**<sup>1</sup>**MEVPEKTKVEIRFQTGSKISTPSTPS**VYVNGC**VEEFENTEPI**APIPPVGPSVDPTVC**SPNAEAK**VWIG**MSQHTNRGETFNDRNSTDLYVQS  
FNRNTGLDEFKPVGAR**TPWVYN**AMAYNPKEGYI**AI**SQGRLKTLQSSKLRIYDEDPNYPAG**HLLQISPVNAGVKDLGAIIGLNGSRLTAWPNDL  
WGGMTSGIIDGNRYLVS**NSSQSGTHN**LYTLDDLTRLATVVASNTAF**SNDYTSTGKTDSNYVWG**IKNSS**PAVLERIDVRDGSRKEFSLDGV  
KDPLGQNVEKGI**YGTAWTYGNGSLGFGN**NATGSVYQIDIGNESGTTIDDLHL**KIVAK**RKGPTSQNNDATSN**GILSPVITDLKV**TKKLEKIEGNQ  
VSWTITVENVGPCPSSGFTLQDVVPEGYA****EVKGDPQSNGWYQD**ISVNGNVINASHGPIA**VNEKATYKVTANQ**SIS**NEKCLQNT**ASIY**ANEKD  
LIEDN**NVASDGACIPAIT**KTVVDQNGDEKIDGQD**GSVAAGN**LRKVTYKIEVKNPKGFPETKYS**LTDPQFADSV**KLERLKVISDYGKKNQEV  
QAADISVNGLVLAEDGTPI****EGNASKDTVHTY**MVEAF**YSGPESSAH**STDECKD**STPKF**GLFNSAKLKV**VSEKTSEGC**APIVRN**VPIKIQLKKV  
DAENKETELQAT**FELYRVSDGGKV**ATLQPRPNQQEQTEVEPGKYRLIETQ**SPVGY**QLLPRPVEFEVTRTADGKADVTIDAAS**FPLSK**SADQG  
KDPNL**VILTVADVRVGTLPKTGGHGV**AIYLVMGVLLVLVGVSWSLYRNQLISS**<sup>796</sup>

**D)**<sup>1</sup>DKAMTSAASTASNSHAFATA**EELKNQSGEG**VGMVEFS**KTAAGTK**ITVDAKGLTPGFHGFHVHDKG**VCEGNFKSAGGHLHGKNQPHGDGY  
HPHAGDLPSLLANADGTAHLEVVTDAVDEDLLFGADGTS**LI**VHEGRDNYANIPQRYAPNGPDEETLKTGD**AGARFACGVITKK<sup>171</sup>

**E)**<sup>1</sup>TLAEEGSTIVPAPIDNNGLFDAAPPAPVARGAVG**EGKLPDPVTSEFFDSKVIRD**VDPAGQRCFRIPAIATAVDGTLLVA**FDNRYGPDAQSR**  
**TWCRDAPYENTKRNSREKQTDIEIYRSHDSGKSFEEDG**FI**AKGTEDPRDLSYTDPSIVVDHETGKIFAVFVRGYDYRFFD**ATAGV**NRGSVED**  
**PIQKRSVQDTVII**ESNDNGQ**TWGNMKLLSELTEKIKVRHKGQEL**AGRGRFATSGNGI**QLKYGENAGRLLFPMSVHYSPNGAGAVANVALY**  
SDDHGATWNVGQGVGG**EGETNGDENKLVELSDGRIMMNSKSTGSNPSNPRGAERWRSYSEDQGETWSKPEPVIVA**PPQHPLQRNTGIN  
VSLI**KAYPSAP**EGSAAARVLLYSAPID**TRIKGEYGGDGRSNGWVMASCDDGRSWTHGRQIEAGRFQYSVMTPMADGNIGMVYESGDKERG**  
**MTLK**FAKFNMAWLGADCLSNKALGITDLPDPALAK**KAIEDAKAATAKAEATANVLKLTKELEEAELENEELAQ**ALAEAKSAAQDAIAAA**EEA**  
**NARVKA**AEGATIEAEEA**AMKAENEAKALAEQLAKVEAELANSQDQAKALAEAKEAAEIARKAAEEALKLEKEKSGKAGGTDNTENKGF**  
**QELLRIFPGFAPIFSFLASIWGGMQKLLAF**<sup>663</sup>

**F)**<sup>1</sup>**RQFPAQHSLYSPQRSTFGTKH**MFRTD**QLIDGIERNVRIT**SEEVNAALPVPL**LDRTDPGAILISGAS**YA**EPSQTLQTLRDAMAQEEFANS**  
**KEIPL**GVVRTLLDLGFIDEARSWLDTLDATL**SDDWRHQWYSG**VTSLFLDDYVMA**QRHFNEVY**NILPGEAAPKLARA**AVCEMLLQGMGMDST**  
PLLSPDTAVSAAELKGNTIGM**WSELT**KDP**ETLRFK**ALYLYAL**VWRTNPTTV**SSAFGLARQLVAENQVDLAVST**LDQVPQNSTHRRMAELTAI**  
LHLISGD**LES**RRRAARR**LES**IPTNEPR**LQIK**IAIMNAAL**TWLRQSRLEAAAADNDLFDYRFTQV**GLRTGLYDSLRL**LARSAPNVHHR**YTLV  
DMAN**QVRPMT**WF<sup>379</sup>

**G)**<sup>1</sup>DKAMTSAASTASNSHAFATA**EELKNQSGEG**VGMVEF<sup>35</sup>

**H)**<sup>1</sup>RSNGWVMASC**DDGRSWTHGRQIEAGRFQYSVMTPMADGNIGMVYESGDKERGMTLK**FAKFNMAWLGADCLSNKALGITDLPDPALAK  
A<sup>88</sup>

Predição de epítomos lineares de células B para as porções completas das proteínas PLD (A), PknG (B), SpaC (C), SodC (D), NanH (E) e para a porção de baixa conservação de PknG (F), SodC (G) e NanH (H). Os epítomos preditos são destacados em vermelho.

## 5.2 Expressão heteróloga em *E. coli* das proteínas recombinantes rPLD, rPknG, rSpaC, rSodC e rNanH de *C. pseudotuberculosis* FRC41

### 5.2.1 Construção dos vetores de expressão contendo as ORF's sintéticas

As sequências das ORF's de PLD, PknG, SpaC, SodC e NanH de *C. pseudotuberculosis* FRC41 foram enviadas à empresa DNA2.0 e sintetizadas com substituição dos códons por códons preferenciais de *E. coli*, a linhagem hospedeira escolhida para os experimentos de expressão neste trabalho. As sequências foram clonadas no vetor de expressão pD444-NH gerando as construções que foram apresentadas na Tabela 2. O material plasmidiano foi recebido em papel filtro e eluído em Tris-HCl 100mM pH7,5.

Os mapas dos plasmídeos pD444-NH:*pld* pode ser visualizado na Figura 5 e os mapas de pD444-NH:*pknG*, pD444-NH:*spac*, pD444-NH:*sodc* e pD444-NH:*nanh* podem ser visualizados na Anexo 4. Nestes esquemas, o inserto encontra-se destacado em vermelho. Em verde, estão marcados o marcador de resistência à ampicilina (AmpR) e repressor Lac (*lacI*). A origem de replicação pUC (*high\_copy\_origin*), promotor T5 e operador lac (*lacO*) estão marcados em azul.

### 5.2.2 Avaliação da produção das proteínas rPLD, rPknG, rSpaC, rSodC e rNanH em diferentes linhagens de expressão

A partir das linhagens *E. coli* OverExpress™ C41(DE3), OverExpress™ C41(DE3)pLysS, OverExpress™ C43(DE3), OverExpress™ C41(DE3) pLysS e BL21 Star™(DE3) transformadas com os plasmídeos contendo as ORF's de cada uma das proteínas recombinantes, foram selecionadas colônias para indução da expressão gênica. A indução da expressão foi realizada pela adição de IPTG 1mM à cultura bacteriana e foram coletadas alíquotas no tempo de 0h (não induzido) e 5h (induzido). As amostras foram analisadas via eletroforese em gel de poliacrilamida sob condições desnaturantes. Dessa forma, foi avaliada a expressão de cada uma das proteínas pelas 5 linhagens de expressão de *E. coli* citadas acima. As linhagens que apresentaram o maior nível de expressão de cada proteína foram selecionadas para os próximos ensaios.

Nos géis onde foram analisadas as amostras retiradas das culturas nas quais foi induzida a expressão da proteína rPLD, de 31kDa, pode ser visualizada banda com o peso molecular esperado. O nível de expressão foi ligeiramente maior na linhagem *E. coli* BL21 Star™(DE3), sendo esta linhagem selecionada para os próximos ensaios (Figura 21).

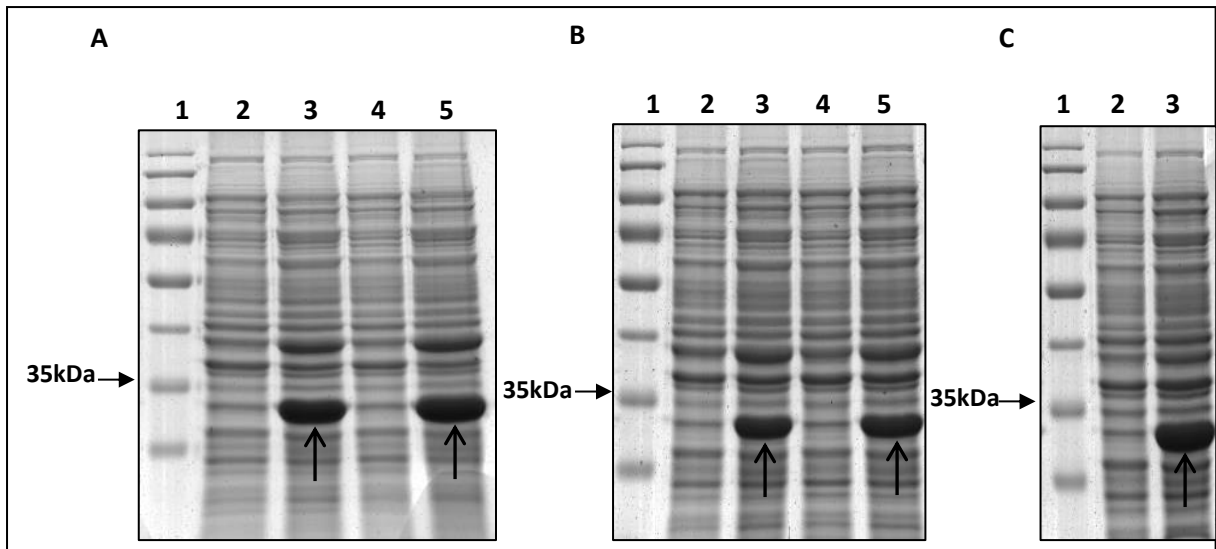


Na análise da expressão da proteína rPknG, de 83kDa, são visualizadas no gel bandas com o peso molecular esperado nas amostras das culturas induzidas de *E.coli* OverExpress™ C41(DE3), OverExpress™ C41(DE3)pLysS e *E. coli* BL21 Star™(DE3). O maior nível de expressão foi identificado na linhagem *E. coli* BL21 Star™(DE3), a qual foi selecionada para os próximos ensaios (Figura 22).

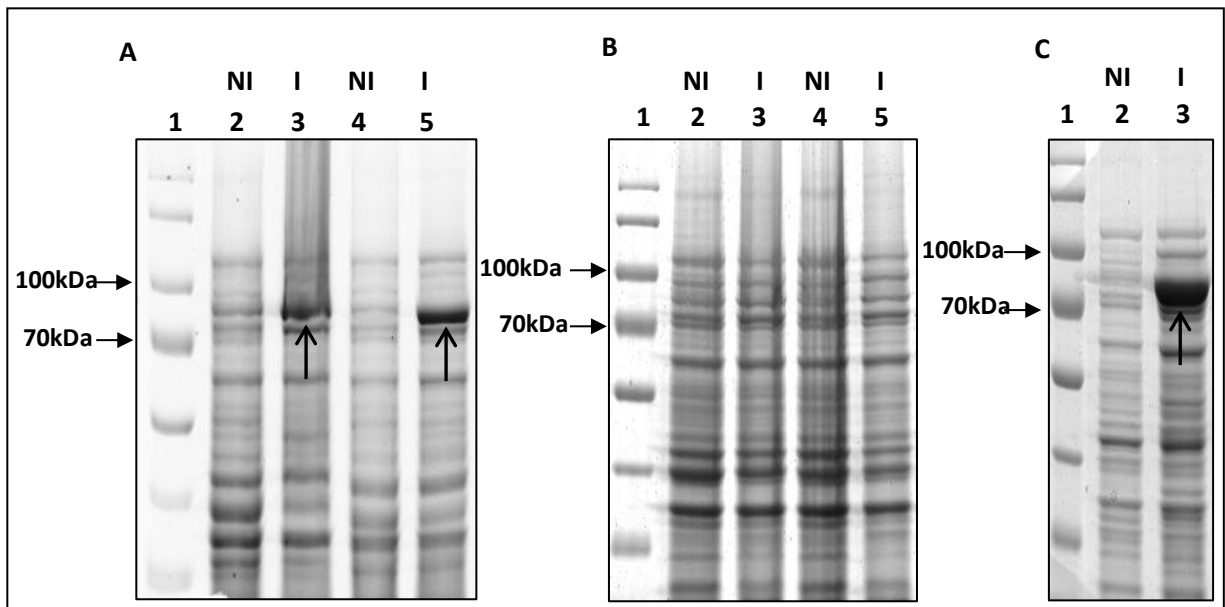
A partir da indução da expressão da proteína rSpaC (86kDa) nas cinco linhagens de *E. coli*, observa-se o aparecimento de bandas nas culturas induzidas das linhagens *E.coli* OverExpress™ C41(DE3), OverExpress™ C41(DE3)pLysS, *E.coli* OverExpress™ C43(DE3) e OverExpress™ C43(DE3)pLysS que são ausentes nas amostras não induzidas. Porém, essas bandas não apresentam o peso molecular esperado, sendo que, de acordo com a comparação com o marcador de peso molecular, essas proteínas aparecem na altura do marcador de 100kDa. A linhagem C41(DE3) apresentou nível de expressão ligeiramente maior que as demais linhagens e foi selecionada para os próximos ensaios de expressão de rSpaC (Figura 23).

Analisando as culturas nas quais foi induzida a expressão da proteína rSodC (18kDa), apenas a linhagem *E.coli* OverExpress™ C41(DE3) não apresentou banda com o peso molecular esperado na visualização em gel de poliacrilamida. As outras quatro linhagens apresentaram bandas com intensidades semelhantes e com o peso molecular esperado. A linhagem escolhida para os próximos passos foi a *E. coli* OverExpress™ C41(DE3)pLysS (Figura 24).

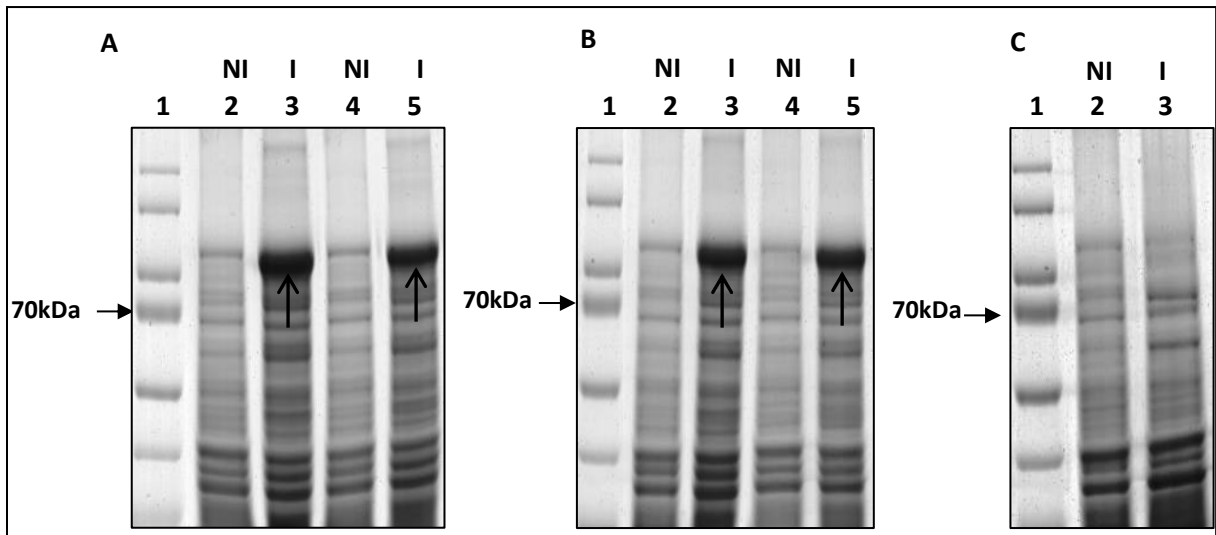
A partir da análise das culturas nas quais foi induzida a expressão da proteína NanH (71,5kDa), observa-se banda com o peso molecular próximo ao esperado nas amostras das culturas das linhagens a *E.coli* OverExpress™ C41(DE3), C43(DE3) e BL21 Star™(DE3). A linhagem *E.coli* OverExpress™ C41(DE3) apresentou nível de expressão maior que as outras linhagens, sendo selecionada para as próximas etapas. (Figura 25).



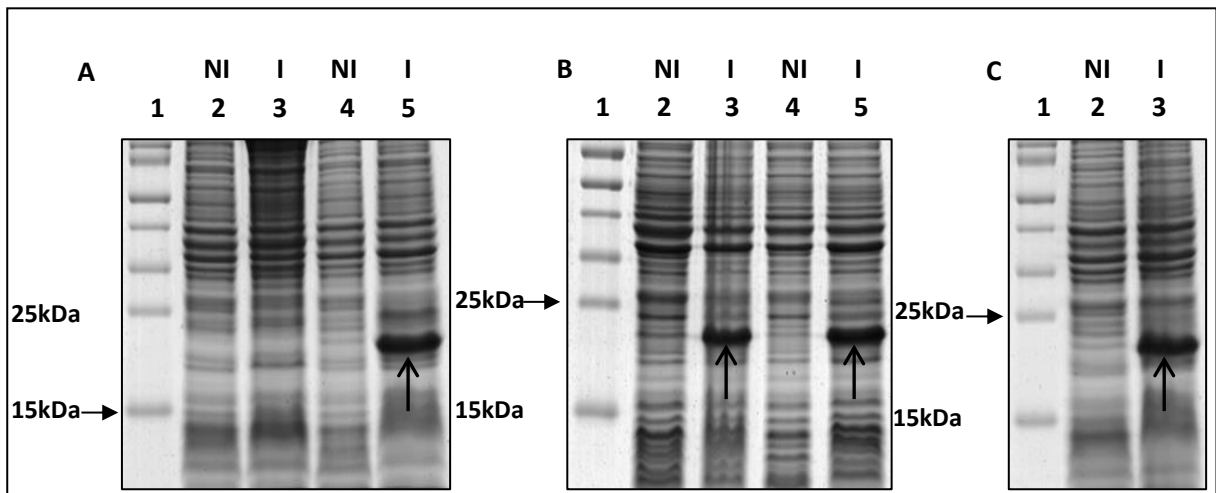
**Figura 21.** Avaliação da expressão da proteína recombinante rPLD (31kDa). Eletroforese em gel de poliacrilamida 12%. Coloração por Coomassie Blue. NI – Não induzido, tempo de 0h. I – Induzido, tempo de 5h. Canaletas 1: PageRuler Prestained Protein Ladder (Thermo Scientific). **A)** Canaletas 2 (NI) e 3 (I) - Linhagem C41(DE3). Canaletas 4 (NI) e 5 (I) - Linhagem C41(DE3)pLysS. **B)** Canaletas 2 (NI) e 3 (I) - Linhagem C43(DE3). Canaletas 4 (NI) e 5 (I) - Linhagem C43(DE3)pLysS. **C)** Canaletas 2 (NI) e 3 (I) - Linhagem BL21 Star (DE3). As setas indicam banda correspondente à expressão da proteína recombinante.



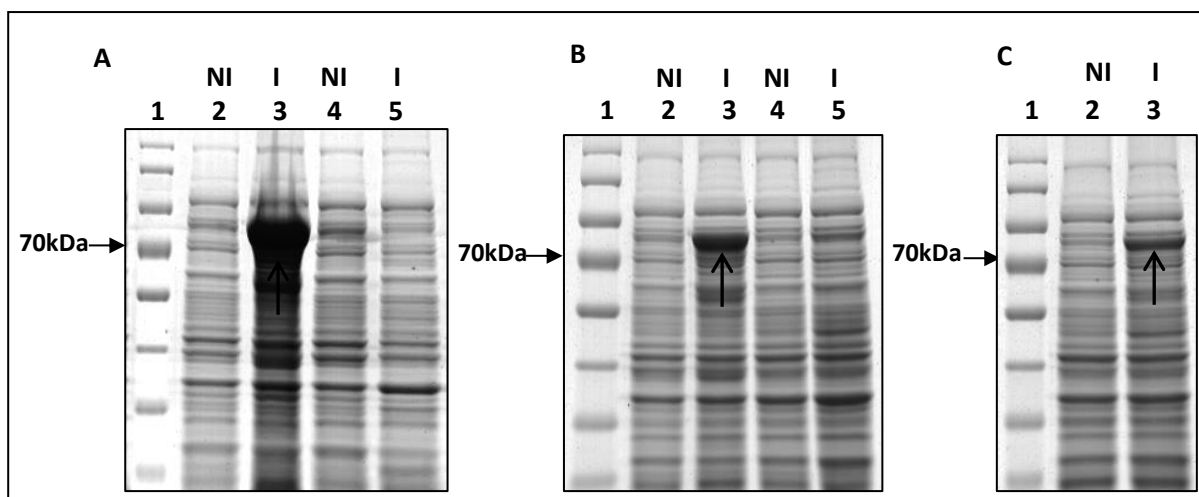
**Figura 22.** Avaliação da expressão da proteína recombinante rPknG (83kDa). Eletroforese em gel de poliacrilamida 10%. Coloração por Coomassie Blue. NI – Não induzido, tempo de 0h. I – Induzido, tempo de 5h. Canaletas 1: PageRuler Prestained Protein Ladder (Thermo Scientific). **A)** Canaletas 2 (NI) e 3 (I) - Linhagem C41(DE3). Canaletas 4 (NI) e 5 (I) - Linhagem C41(DE3)pLysS. **B)** Canaletas 2 (NI) e 3 (I) - Linhagem C43(DE3). Canaletas 4 (NI) e 5 (I) - Linhagem C43(DE3)pLysS. **C)** Canaletas 2 (NI) e 3 (I) - Linhagem BL21 Star (DE3). As setas indicam banda correspondente à expressão da proteína recombinante.



**Figura 23:** Avaliação da expressão da proteína recombinante rSpaC (86kDa). Eletroforese em gel de poliacrilamida 10%. Coloração por Coomassie Blue. NI – Não induzido, tempo de 0h. I – Induzido, tempo de 5h. Canaletas 1: PageRuler Prestained Protein Ladder (Thermo Scientific). **A)** Canaletas 2 (NI) e 3 (I) - Linhagem C41(DE3). Canaletas 4 (NI) e 5 (I) - Linhagem C41(DE3)pLysS. **B)** Canaletas 2 (NI) e 3 (I) - Linhagem C43(DE3). Canaletas 4 (NI) e 5 (I) - Linhagem C43(DE3)pLysS. **C)** Canaletas 2 (NI) e 3 (I) - Linhagem BL21 Star (DE3). As setas indicam banda correspondente à expressão da proteína recombinante.



**Figura 24:** Avaliação da expressão da proteína recombinante rSodC (18kDa). Eletroforese em gel de poliacrilamida 15%. Coloração por Coomassie Blue. NI – Não induzido, tempo de 0h. I – Induzido, tempo de 5h. Canaletas 1: PageRuler Prestained Protein Ladder (Thermo Scientific). **A)** Canaletas 2 (NI) e 3 (I) - Linhagem C41(DE3). Canaletas 4 (NI) e 5 (I) - Linhagem C41(DE3)pLysS. **B)** Canaletas 2 (NI) e 3 (I) - Linhagem C43(DE3). Canaletas 4 (NI) e 5 (I) - Linhagem C43(DE3)pLysS. **C)** Canaletas 2 (NI) e 3 (I) - Linhagem BL21 Star (DE3). As setas indicam banda correspondente à expressão da proteína recombinante.



**Figura 25:** Avaliação da expressão da proteína recombinante rNanH (71,5kDa). Eletroforese em gel de poliacrilamida 10%. Coloração por Coomassie Blue. NI – Não induzido, tempo de 0h. I – Induzido, tempo de 5h. Canaletas 1: PageRuler Prestained Protein Ladder (Thermo Scientific). **A)** Canaletas 2 (NI) e 3 (I) - Linhagem C41(DE3). Canaletas 4 (NI) e 5 (I) - Linhagem C41(DE3)pLysS. **B)** Canaletas 2 (NI) e 3 (I) - Linhagem C43(DE3). Canaletas 4 (NI) e 5 (I) - Linhagem C43(DE3)pLysS. **C)** Canaletas 2 (NI) e 3 (I) - Linhagem BL21 Star (DE3). As setas indicam banda correspondente à expressão da proteína recombinante.

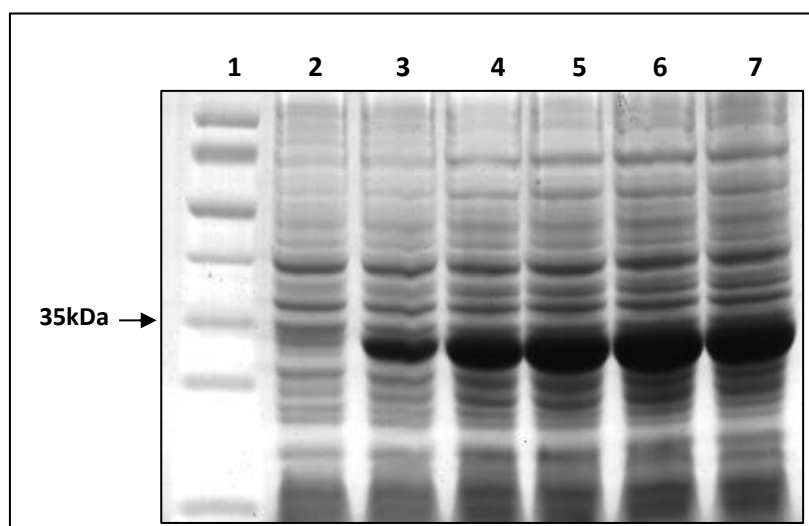
### 5.2.3 Avaliação da cinética de expressão das proteínas rPLD, rPknG, rSpaC, rSodC e rNanH

A partir da avaliação da expressão das proteínas recombinantes pelas cinco linhagens de *E.coli*, foi selecionada uma linhagem para cada proteína, conforme descrito no tópico anterior e exposto na Tabela 19. Então, foi conduzida a avaliação da cinética de expressão das proteínas em cinco horas de indução de expressão, sendo coletadas amostras a cada hora. A expressão foi induzida pela adição de IPTG 1mM.

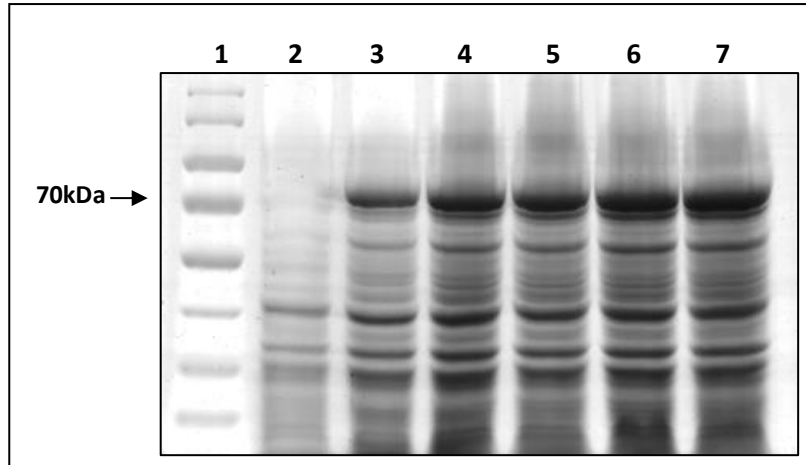
As amostras foram analisadas em gel de poliacrilamida e todas as induções apresentaram bandas com intensidade crescente ao longo do tempo. Não houve grande diferença no nível de expressão entre os tempos de 4h e 5h (Figuras 26 a 30). As bandas visualizadas apresentaram o tamanho molecular esperado, com exceção das amostras referente à proteína rSpaC, nas quais as bandas ficaram acima da altura esperada.

**Tabela 19.** Linhagens de *E.coli* selecionadas para a expressão das proteínas recombinantes.

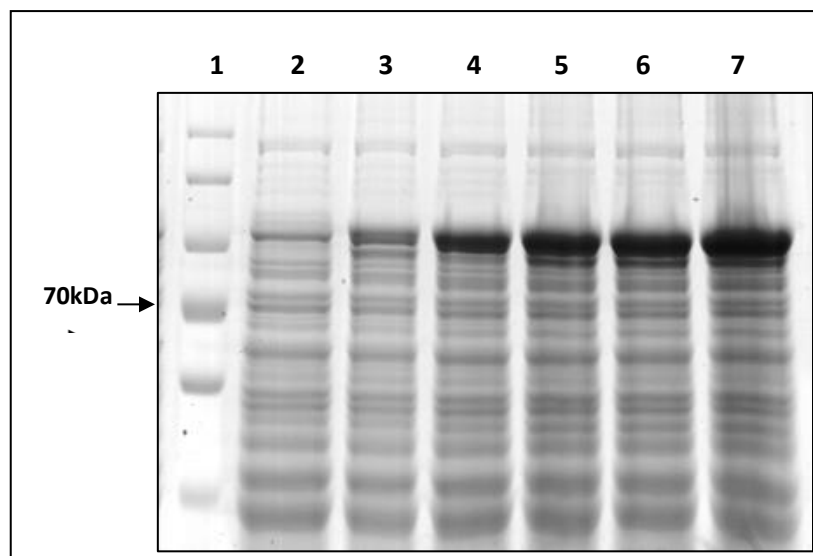
Linhagem de <i>E.coli</i> para expressão	Proteína	Peso molecular
BL21 Star™(DE3)	rPLD	31kDa
BL21 Star™(DE3)	rPknG	83kDa
OverExpress™ C41(DE3)	rSpaC	86kDa
OverExpress™ C43(DE3) pLysS	rSodC	18kDa
OverExpress™ C41(DE3)	rNanH	71,5kDa



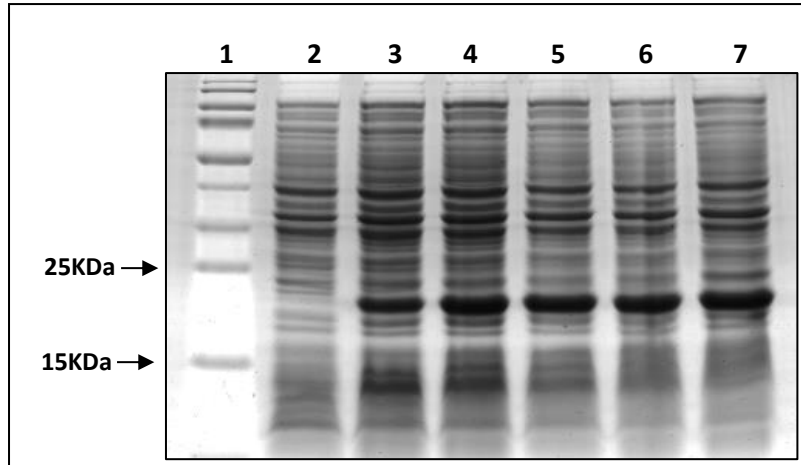
**Figura 26.** Avaliação da cinética de expressão da proteína recombinante rPLD (31kDa) na linhagem BL21 Star™(DE3). Eletroforese em gel de poliacrilamida 12%. Coloração por Coomassie Blue. Coluna 1: PageRuler Prestained Protein Ladder (Thermo Scientific). Colunas 2 a 7 correspondem aos tempos 0h a 5h de expressão, respectivamente.



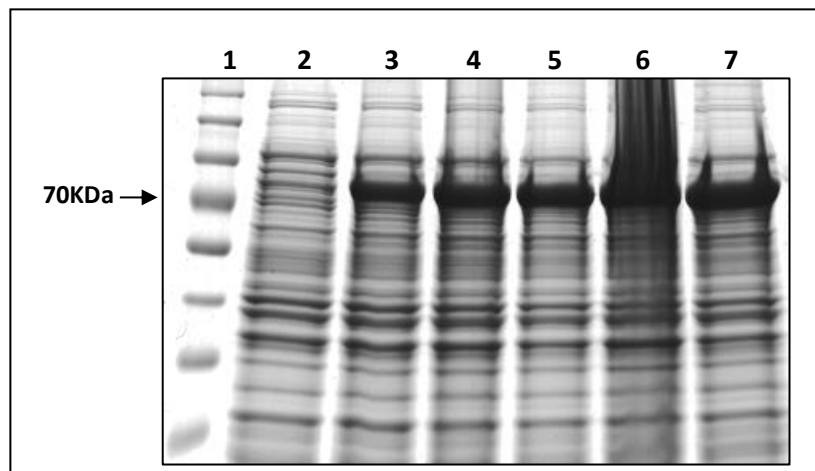
**Figura 27.** Avaliação da cinética de expressão da proteína recombinante rPknG (83kDa) na linhagem BL21 Star™(DE3). Eletroforese em gel de poliacrilamida 10%. Coloração por Coomassie Blue. Coluna 1: PageRuler Prestained Protein Ladder (Thermo Scientific). Colunas 2 a 7 correspondem aos tempos 0h a 5h de expressão, respectivamente.



**Figura 28.** Avaliação da cinética de expressão da proteína recombinante rSpaC (86kDa) na linhagem OverExpress™ C41(DE3). Eletroforese em gel de poliacrilamida 10%. Coloração por Coomassie Blue. Coluna 1: PageRuler Prestained Protein Ladder (Thermo Scientific). Colunas 2 a 7 correspondem aos tempos 0h a 5h de expressão,



**Figura 29.** Avaliação da cinética de expressão da proteína recombinante rSodC (18kDa) na linhagem OverExpress™ C43(DE3)pLysS. Eletroforese em gel de poliacrilamida 15%. Coloração por Coomassie Blue. Coluna 1: PageRuler Prestained Protein Ladder (Thermo Scientific). Colunas 2 a 7 correspondem aos tempos 0h a 5h de expressão,

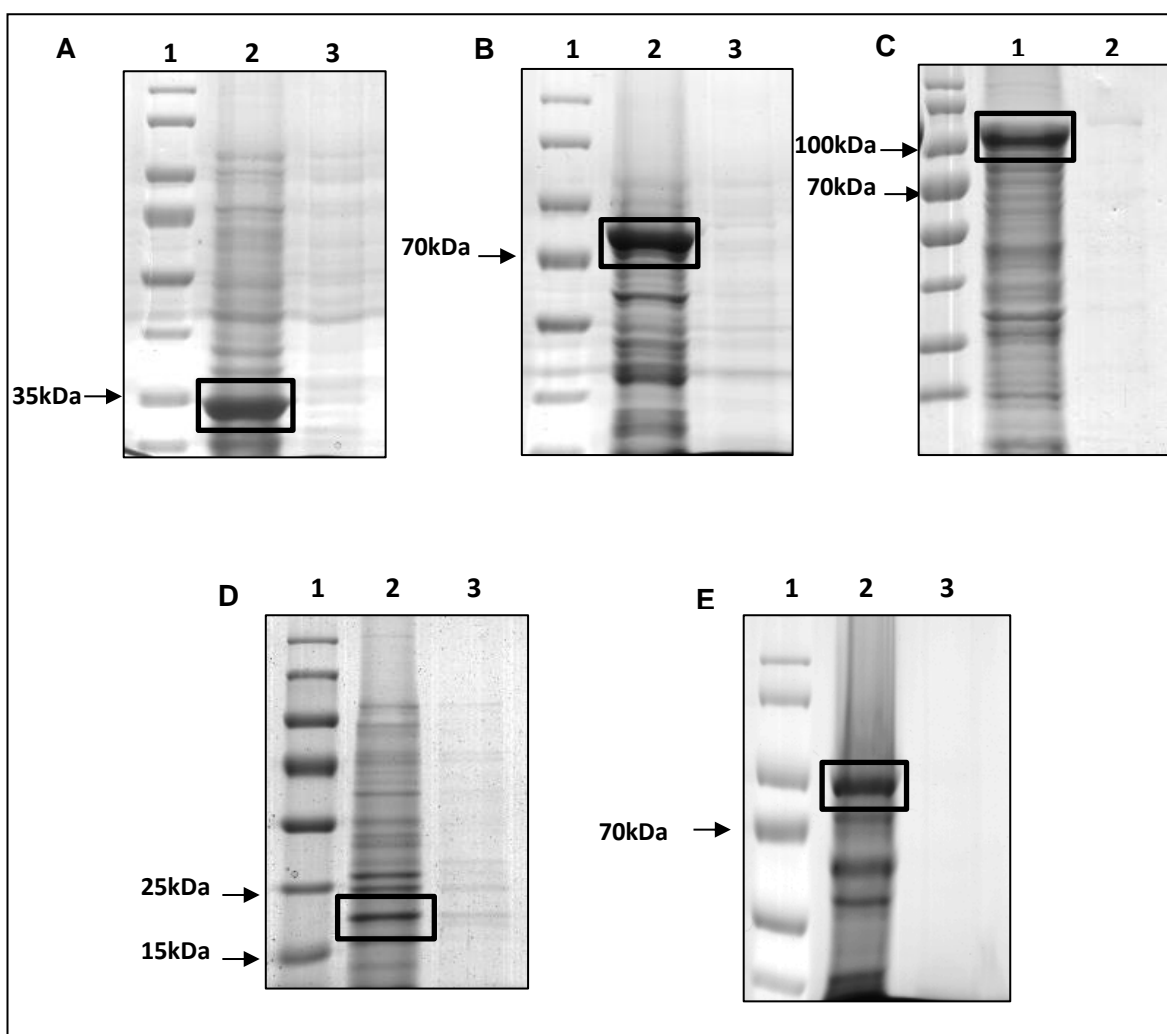


**Figura 30.** Avaliação da cinética de expressão da proteína recombinante rNanH (71,5kDa) na linhagem OverExpress™ C43(DE3)pLysS. Eletroforese em gel de poliacrilamida 10%. Coloração por Coomassie Blue. Coluna 1: PageRuler Prestained Protein Ladder (Thermo Scientific). Colunas 2 a 7 correspondem aos tempos 0h a 5h de expressão, respectivamente.

## 5.2.4 Teste de solubilidade das proteínas recombinantes

Após a avaliação da expressão das proteínas recombinantes em *E. coli*, as células foram lisadas por sonicação e centrifugadas obtendo-se o precipitado e o sobrenadante. Foram utilizadas as mesmas linhagens nas quais foi avaliada a cinética de expressão (Tabela 19).

As bandas referentes a cada uma das proteínas puderam ser visualizadas em gel de poliacrilamida na amostra correspondente ao precipitado, indicando a insolubilidade das proteínas. As bandas apresentaram o peso molecular esperado, com exceção da amostra onde foi induzida a expressão de rSpaC, a qual foi visualizada na altura próxima ao marcador de 100kDa (Figura 32).



**Figura 31.** Teste de solubilidade das proteínas recombinantes. Eletroforese em gel de poliacrilamida 12%. Coloração por Coomassie Blue. Canaletas 1: PageRuler Prestained Protein Ladder (Thermo Scientific). Canaletas 2: Fração insolúvel. Canaletas 3: Fração solúvel. **A)** rPLD. **B)** rPknG. **C)** rSpaC. **D)** rSodC. **E)** rNanH.



### 5.2.5 Purificação da proteína PknG

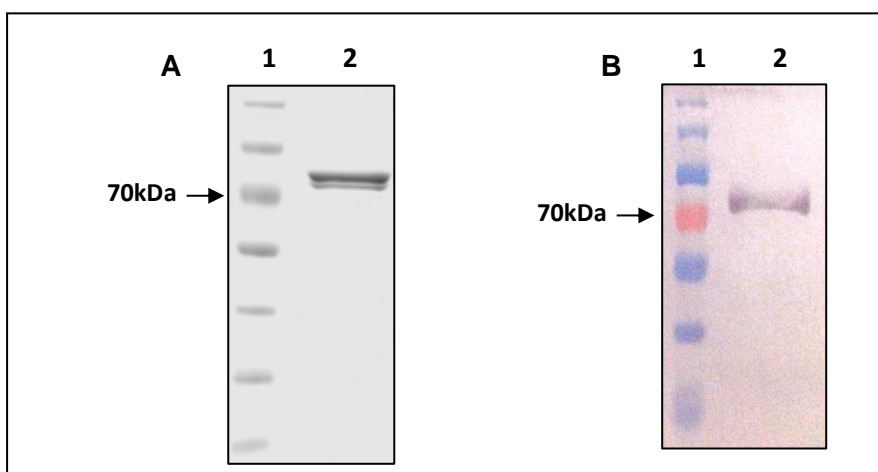
Para purificação foi escolhida primeiramente a proteína rPknG, que obteve melhor predição de epítomos além de alto nível de expressão. Temos como perspectiva a purificação das outras quatro proteínas recombinantes. Os experimentos para purificação da proteína rPknG foram desenvolvidos em colaboração com o professor Raghuvir Krishnaswamy Arni da Universidade Estadual Paulista Julio de Mesquita Filho – Unesp – São José do Rio Preto.

A proteína PknG passou por três passos de purificação após a expressão em um litro de cultura. Primeiramente, a proteína passou por coluna de afinidade utilizando resina de níquel. A fração recolhida passou por gel filtração em sistema AKTA purifier e em seguida, purificação por troca iônica utilizando coluna mono-Q. A purificação foi satisfatória, o que pode ser visualizado na Figura 43 onde observa-se banda única na altura esperada. Atingiu-se um rendimento de 10mg/L.

### 5.2.6 Confirmação da expressão da proteína rPknG por *Western blot*

A proteína rPknG purificada foi transferida para membrana de PVDF e foi realizado o *Western blot* utilizando anticorpo Anti-His.

Observa-se na Figura 43 que houve revelação da banda correspondente à proteína rPknG com o peso molecular esperado confirmando a presença da proteína recombinante pela cauda 6xHis.



**Figura 32.** Purificação da proteína rPknG a partir da expressão na linhagem BL21 Star™(DE3) e confirmação por *Western blot*. **A)** Purificação da proteína rPknG a partir da expressão na linhagem BL21 Star™(DE3). Eletroforese em gel de poliácridamida 10%. Coluna 1: PageRuler Prestained Protein Ladder (Thermo Scientific). Coluna 2: Proteína rPknG purificada. **B)** *Western blot* com a proteína rPknG a partir da proteína rPknG purificada. Coluna 1: PageRuler Prestained Protein Ladder (Thermo Scientific). Coluna 2: Proteína rPknG purificada.

## 6. DISCUSSÃO

O conhecimento e entendimento dos fatores de virulência associados à LC são de grande importância para a compreensão dos mecanismos de patogenicidade de *C. pseudotuberculosis* bem como para o desenvolvimento de novas abordagens de controle da doença. A vacinação é indicada como medida de controle mais adequada à LC. Porém, alguns aspectos ainda precisam ser melhorados em relação às vacinas disponíveis, como a melhoria da eficácia protetora e também da praticidade de utilização, uma vez que a imunização pode ter sua eficácia reduzida devido ao uso incorreto por parte dos criadores (Dorella, 2009). Além disso, até o momento, nem todas as vacinas licenciadas apresentam a mesma eficácia para caprinos e ovinos (Williamson, 2001). Neste sentido, várias estratégias vacinais vem sendo testadas contra a LC.

Novas estratégias tem sido possibilitadas pelo sequenciamento de genomas de *C. pseudotuberculosis* por permitir a identificação de novos potenciais fatores de virulência os quais, por sua vez, podem ser importantes alvos para o desenvolvimento de drogas e vacinas. Entre estes novos potenciais fatores de virulência, encontram-se as proteínas alvo deste estudo que foram apontadas como potenciais fatores de virulência a partir da análise do genoma da linhagem *C. pseudotuberculosis* FRC41. Selecionamos, baseados em dados da literatura e em análises de genômica comparativa, as proteínas PknG, SpaC, SodC e NanH. Tivemos como alvo também a proteína PLD, apontada como principal fator de virulência de *C. pseudotuberculosis* e responsável pelo desencadeamento de resposta imunológica contra a LC em estudos onde foi avaliada uma forma atóxica desta proteína. Não se obteve sucesso na utilização da PLD sozinha como vacina em dose única (Burrell, 1983; Eggleton *et al.*, 1991a; Eggleton *et al.*, 1991b). Porém, acreditamos que possa ser alcançado um bom resultado pela combinação da PLD a outras proteínas com potencial vacinal. Além disso, existe a necessidade de sua caracterização estrutural que, por sua vez, possibilitaria o estudo de drogas contra a LC.

Buscando a caracterização *in silico* das proteínas alvo deste estudo, foram realizadas primeiramente as predições de peptídeo sinal na sequência, localização subcelular e domínios conservados.

A proteína PknG, uma serina-treonina proteína quinase que pode estar envolvida com o metabolismo de glutamina e inibição da fusão fagossomo-lisossomo, foi predita como citoplasmática e não apresentou peptídeo sinal, o que é condizente com os dados na literatura relacionados a proteínas homólogas à PknG de *C. pseudotuberculosis* (Tabela 6 e Figura 6B) (Niebisch *et al.*, 2006; Scherr & colaboradores 2009; Burnside & Rajagopal, 2011).

Em micobactéria, acredita-se que PknG seja secretada intra-macrófago contribuindo para a inibição da fusão fagossomo-lisossomo, por vias alternativas, uma vez que não possui sinal de exportação (Walburger *et al.*, 2004). Possivelmente, este mecanismo ocorra também em *C. pseudotuberculosis* por apresentar um processo infeccioso similar a micobactéria. PknG apresentou em sua sequência predição do domínio “proteína quinase”, sítio ativo “serina/treonina proteína quinase” e domínio “tetratricopeptídeo-*like* helicoidal” (Figura 8), como era esperado. O domínio proteína quinase contém a função catalítica de proteínas quinases e está presente nas três classes: serina/treonina proteína quinase, tirosina proteína quinase, proteína quinase de especificidade dual (Hanks *et al.*, 1988; Hanks & Quinn, 1991). A fosforilação de um ou mais resíduos de aminoácidos é responsável por mudanças funcionais de uma ampla gama de proteínas através de mudanças conformacionais (Hanks *et al.*, 1988). As subunidades catalíticas de proteínas quinases são estruturas altamente conservadas (Stout *et al.*, 2004). O sítio ativo “serina/treonina proteína quinase” é caracterizado pela presença de um resíduo de ácido aspártico conservado na porção central do domínio catalítico (Knighton *et al.*, 1991). O domínio “tetratricopeptídeo-*like* helicoidal” (TPR) é caracterizado pelo motivo de repetições de tetratricopeptídeo e está presente em diferentes proteínas sendo conservado em diversos organismos (Lamb *et al.*, 1995; Das *et al.*, 1998; Goebel & Yanagida, 1991). Este motivo consiste de 3 a 16 repetições em *tandem* de uma sequência degenerada de 34 resíduos de aminoácidos formando um suporte que proporciona a interação proteína-proteína e a montagem de complexos protéicos. O TPR está envolvido em diversos processos biológicos como o funcionamento de chaperonas, ciclo celular, transcrição e complexos de transporte de proteínas (Blatch & Lässle, 1999). De acordo com suas características, PknG integra vias de sinalização celular e a compreensão de suas funções e estrutura pode contribuir para o desenvolvimento de drogas. Além disso, esta proteína pode ter potencial vacinal como já apresentado por outras serina-treonina proteínas quinases como, por exemplo, a STPK de *Streptococcus pneumoniae* (Burnside & Rajagopal, 2011; Giefing *et al.*, 2008; Giefing *et al.*, 2010).

Diferente da proteína PknG, as proteínas PLD, SodC e NanH apresentaram peptídeo sinal na predição através da ferramenta SignalP (Figuras 5 a Anexo 2), resultado que é condizente com a predição de localização subcelular dessas proteínas. PLD foi classificada pelo programa PSORTb como tendo localização desconhecida (Tabela 6) e, de acordo com as instruções disponíveis no PSORTb, proteínas que são classificadas como tendo localização desconhecida, podem ter na verdade múltiplos sítios de localização. Segundo dados da literatura, sabe-se que PLD é uma proteína secretada (Songer *et al.*, 1990).

Além disso, a proteína PLD apresentou o domínio “fosfodiesterase PLC-*like*, TIM beta/alfa-Barril” (Figura 7), que representa um domínio estrutural formado por uma folha beta e uma alfa hélice, repetidas oito vezes formando uma estrutura cilíndrica estável (barril

( $\alpha/\beta$ )<sub>8</sub> (TIM)). O dobramento (barril ( $\alpha/\beta$ )<sub>8</sub> (TIM) é amplamente distribuído em proteínas de diferentes classes e funções, porém, exibe baixa similaridade de sequência (Gromiha et al., 2004). PLD foi classificada como pertencendo à família “Fosfolipases D” e representa o fator de virulência mais estudado de *C. pseudotuberculosis*.

Já as proteínas SodC e NanH foram preditas como proteínas extracelulares pelo PSORTb (Tabela 6), condizente com os dados na literatura para proteínas homólogas a elas (Trost et al., 2010; Dussurget et al., 2001; Kim et al., 2009). A proteína SodC apresentou, como esperado, o “domínio de ligação a Cu/Zn de superóxido dismutase” e o “sítio de ligação Cu/Zn de superóxido dismutase” (Figura 10), o qual é caracterizado por dois padrões: dois resíduos de histidina que ligam o átomo de cobre (cofator metálico) e um resíduo de cisteína localizado na porção C-terminal envolvido em uma ligação dissulfeto. SodC contribui para a resistência de microrganismos contra o estresse oxidativo gerado, por exemplo no ambiente intra-macrofágico, como observado em *M. tuberculosis* e *S. choleraesuis* (Dussurget et al., 2001; Piddington et al., 2001; Sansone et al., 2002), entre outros microrganismos.

NanH, por sua vez, foi classificada como pertencente à família das “sialidases” que inclui diversas classes de sialidases e apresentou o domínio “sialidase” (Figura 11). As sialidases eucarióticas bacterianas e virais compartilham regiões conservadas de motivos de folhas beta e são responsáveis pela hidrólise de ligação terminal de ácido siálico em vários substratos naturais, como glicoproteínas, glicolipídeos, gangliosídeos e polissacarídeos (Monti et al., 2002). As sialidases bacterianas agem sobre moléculas do hospedeiro e são apontadas como fator de virulência. Elas sequestram ácidos siálicos de glicoproteínas ou outros gliconjugados e podem transferir esses resíduos para a superfície celular do microrganismo contribuindo, dessa forma, para evasão do sistema imune e interação com células do hospedeiro (Vimr & Lichtensteiger, 2002; Vimr et al., 2004). Além disso, essas enzimas são estudadas como alvo vacinal em diversas doenças patogênicas (Taylor, 1996; Mattos-Guaraldi et al., 1998; Tai, 2006).

A proteína SpaC, por sua vez, foi predita como uma proteína de parede celular (Tabela 6), porém, não apresentou peptídeo sinal (Anexo 1). Foi identificado em sua sequência, o domínio LPXTG de ancoramento a parede celular (Figura 9). O domínio LPXTG encontra-se no sinal de endereçamento celular C-terminal de proteínas de superfície de microrganismos Gram-positivos (CWSS – do inglês *cel wall sorting signal*) (Schneewind et al., 1993). Os precursores de proteínas de superfície são direcionados à via secretória de microrganismos Gram-positivos através do peptídeo sinal N-terminal (Bae & Schneewind, 2003). Então, o CWSS contendo o motivo LPXTG, ou outra sequência de reconhecimento, é reconhecido por sortase que catalisa a clivagem entre os resíduos de Thr e Gly levando ao ancoramento do novo resíduo Thr C-terminal à parede celular (Ton-That & Schneewind,

2004). Regiões hidrofóbicas e resíduos positivamente carregados de CWSS retem a proteína translocada na membrana plasmática até que o substrato seja reconhecido por sortase (Schneewind *et al.*, 1992). A presença do domínio LPXTG bem como a localização na parede celular é condizente com os dados na literatura quanto às proteínas homólogas à SpaC, uma adesina presente na extremidade de *pili* e envolvida com a aderência à células do hospedeiro (Mandlik *et al.*, 2008). De acordo com o exposto, seria esperada a presença do peptídeo sinal N-terminal, o que não ocorreu de acordo com a predição realizada pelo SignalP. Uma possibilidade é que a sequência depositada no GenBank apresente um *start codon upstream* ao sugerido e que este *start codon upstream* seja precedido de peptídeo sinal.

Na busca por alvos vacinais, as proteínas secretadas ou proteínas de membrana são interessantes uma vez que fazem parte da interface parasito-hospedeiro. Estas proteínas podem interagir com o hospedeiro através da adesão às células do hospedeiro, invasão, multiplicação, evasão da resposta imune, danos ao hospedeiro e resistência aos estresses gerados pela célula do hospedeiro e (Bhavisar *et al.*, 2007; Stavrinides *et al.*, 2008; Sibbal *et al.*, 2009; Simeone *et al.*, 2009; Wooldridge *et al.*, 2009). Porém, as proteínas intracelulares também podem ser expostas ao sistema imune do hospedeiro após a lise do microrganismo. Além disso, foi demonstrado por Lee e colaboradores (2009) que bactérias gram-positivas possuem um sistema de secreção por vesículas de membrana que medeia a comunicação intracelular e que pode ser capaz de expor proteínas intracelulares.

As proteínas alvo deste estudo foram avaliadas também quanto à sua conservação em organismos representantes do grupo CMNR - *C. pseudotuberculosis* linhagens 1002 (biovar ovis – isolada de caprino); 258 (biovar equi) e C231 (biovar ovis – isolada de ovino); *Corynebacterium diphtheriae* HC02; *C. glutamicum* ATCC 14067; *Mycobacterium tuberculosis* SUMu007; *Nocardia farcinica* (strain IFM 10152); *Rhodococcus pyridinivorans* SB3094; *Rhodococcus* sp. (strain RHA1). Foi avaliada também a conservação das sequências em organismos eucariotos - gênero *Ovis*, *Bos* e *Equus*, que representam hospedeiros de *C. pseudotuberculosis*; *Mus musculus*, o principal modelo animal para estudo *in vivo*; e *Homo sapiens*.

Observou-se que as sequências são conservadas em alguns organismos do grupo CMNR avaliados e constatou-se alta identidade ao comparar-se as proteínas (PLD, PknG, SpaC, SOdC e NanH) da linhagem *C. pseudotuberculosis* FRC41 com as sequências das linhagens 1002, C231 e 258 de *Corynebacterium pseudotuberculosis* (Tabelas 7, 8, 10, 11, 13; Figuras 12, 13, 14, 16, 17, 19). Além dos representantes de *C. pseudotuberculosis* aqui mostrados, essas proteínas também são conservadas nos outros genomas de *C. pseudotuberculosis* (dados não mostrados). Esta característica é interessante no que diz respeito ao desenvolvimento de vacinas ou drogas que visem a inibição da função protéica,

uma vez que estas proteínas podem representar potenciais alvos em diferentes hospedeiros e até mesmo entre biovares diferentes de *C. pseudotuberculosis*.

Já a avaliação da conservação das sequências em organismos eucariotos torna-se importante devido à possibilidade de reação cruzada com antígenos do hospedeiro devido à similaridade do antígeno vacinal com antígenos codificados pelo genoma do hospedeiro (Sette & Rappuoli, 2010). Observou-se que as proteínas PknG, SodC e NanH apresentaram conservação em sequências dos organismos eucariotos investigados com identidade de sequência por volta de 30% (Tabelas 9, 12, 14; Figuras 15, 18, 20). Avaliando as regiões de maior conservação, percebe-se que, em alguns casos, os domínios conservados encontram-se nessa região. É o caso do domínio “proteína quinase” de PknG que se estende do aminoácido 111 a 347, estando inserido na região de maior conservação da proteínas (posição 1 a 369). Já no caso da proteína SodC, o “domínio de ligação Cu/Zn de superóxido dismutase” estende-se sobre as porções conservadas e não conservadas da proteína. A região de menor conservação de SodC é a porção N-terminal até o aminoácido de posição 70 (considerando a porção de peptídeo sinal). Em relação à proteína NanH, o domínio sialidase estende-se da posição 90 a 500, sendo esta, em sua maior parte, uma região de maior conservação da sequência. Em NanH, a região de menor conservação compreende os aminoácidos 432 a 519.

Uma possível abordagem futura seria a avaliação das porções não conservadas das proteínas PknG, SodC e NanH como alvo vacinal comparadas às proteínas inteiras, uma vez que até mesmo pequenos peptídeos podem ter potencial vacinal. Além disso, a resposta imunológica pode ser aumentada pela combinação de diversos peptídeos, adjuvantes ou outros imunógenos ou ainda pela fusão de diferentes epítopos (Hansson *et al.*, 2000; Rogan & Babiuk, 2005).

Para o desenvolvimento de vacinas de subunidades, é importante que os alvos escolhidos apresentem epítopos envolvidos no desenvolvimento da resposta imune, o que é essencial para o desenvolvimento de uma vacina capaz de inibir infecção e doença. Avanços na bioinformática tem permitido a identificação de possíveis epítopos protetores (Rogan & Babiuk, 2005).

Linfócitos B e linfócitos T desempenham papel principal no desenvolvimento de resposta imune. As células B reconhecem epítopos de antígenos através da identificação por paratopos do anticorpo. Os epítopos de células B são classificados em epítopos lineares e conformacionais (Saha & Raghava, 2006). O mapeamento de epítopos lineares de células B pode ser feito por diferentes técnicas e tem sido o foco das pesquisas que buscam determinar epítopos de células B (Flower, 2007). Já as células T estão envolvidas na imunidade adquirida mediada por célula. O receptor de célula T (TCR) reconhece antígenos peptídicos que formam complexos com proteínas do MHC (do inglês *Major*

*Histocompatibility Complex*), que por sua vez, são expressas na superfície de várias células e no homem é chamado HLA (do inglês *Human Leucocyte Antigen*) (Klein & Sato, 2000; Ladner, 2007). O complexo HLA está localizado no cromossomo 6 e codifica mais de 40 antígenos leucocitários (Klein, 1986; Klein & Horejsi, 1998; Forbes & Trowsdale, 1999). Os genes MHC são divididos em duas classes, I e II. Genes de classe I são expressos pela maioria das células, enquanto que genes de classe II são expressos por células B, células T ativadas, macrófagos, células dendríticas e células epiteliais do timo (Klein, 1986). Os peptídeos antigênicos podem ser conjugados ao MHC por duas vias, citosólica ou endocítica. A via citosólica geralmente ocorre em decorrência de vírus e bactérias intracelulares que tem suas proteínas degradadas no citoplasma e posteriormente conjugadas a moléculas MHCI. Esses peptídeos levam a ativação de linfócitos T-CD8. Já a via endocítica, ocorre em células apresentadoras ou células estimuladas por IFN- $\gamma$ . Após fagocitose ou endocitose, os antígenos são degradados e conjugados a MHCII. A resposta é desenvolvida a partir de interação com células T-CD4 (Klein & Sato, 2000). As moléculas de MHCI, geralmente se ligam a pequenos peptídeos de até 9 aminoácidos. Já as moléculas de MHCII se ligam a peptídeos de tamanhos variados, de 11 a 30 aminoácidos, sendo que os peptídeos mais curtos e mais longos não são comuns (Rammensee *et al.*, 1995). Apesar de as moléculas de MHCII permitirem a ligação de peptídeos de comprimento variados, uma região de 9 aminoácidos (*core*) no peptídeo é essencial para a ligação às moléculas de MHCII (Madden, 1995; Rammensee *et al.*, 1995). A predição de epítomos de MHCII tem sido baseada nesse *core* de ligação de 9 aminoácidos. A predição de epítomos de MHCII é mais complicada devido à variabilidade de comprimento dos peptídeos (Klein & Sato, 2000).

Neste trabalho, buscamos a avaliação quantitativa de epítomos de MHC de classe I, MHC de classe II e epítomos de células B, através dos *softwares* Mature Epitope Density - MED 1.0 Server, NetMHCII 2.2 e Bcepred, respectivamente. Esta avaliação foi realizada tanto sobre as sequências completas das proteínas, quanto sobre as porções de menor conservação em eucariotos das proteínas PknG, SodC e NanH. É importante ressaltar que os dados obtidos a partir das ferramentas disponíveis para predição de epítomos são referentes a alelos de HLA humanos e de camundongos, conforme descrito na metodologia. Dessa forma, elas devem ser interpretadas apenas com um indicativo do potencial imunogênico das moléculas investigadas.

Na avaliação de epítomos de MHCI, a proteína PknG foi a que obteve melhor resultado através da predição pelo MED. PknG obteve um MED score de 12,67. As proteínas PLD, SpaC, SodC e NanH tiveram um MED score entre 6 e 7 (Tabela 15). Em estudo onde foi avaliada a densidade de epítomos a partir do genoma de *C. pseudotuberculosis* 1002, o MED score variou até 16.26, sendo que os melhores candidatos vacinais apresentaram densidade de epítomos a partir de 9 (Comunicação pessoal). Dessa

forma, PknG teve um resultado acima da média, enquanto que as outras proteínas tiveram um resultado mediano na predição de epítomos de MHC I.

Para os epítomos de MHC II, foi calculada a porcentagem de epítomos de alta afinidade na sequência das proteínas. As cinco proteínas apresentaram porcentagens aproximadas de epítomos de MHC II de alta afinidade, inclusive a PLD que é a proteína mais estudada como fator de virulência de *C. pseudotuberculosis* (Tabela 17).

Quando investigada a distribuição de epítomos lineares de células B, observou-se uma distribuição homogênea destes epítomos ao longo das sequências das cinco proteínas (Quadros 1-5).

Ao serem avaliadas as porções de menor conservação das proteínas PknG, SodC e NanH, observa-se que houve um aumento na densidade de epítomos de MHC I (Tabela 16). Já a porcentagem de epítomos de MHC II manteve-se próxima a encontrada para as proteínas inteiras (Tabela 18). A distribuição de epítomos de células B manteve-se homogênea (Quadros 6-8). Estes resultados indicam que as porções de menor conservação das proteínas também podem ser consideradas candidatas vacinais.

Estes dados indicam que poderíamos classificar as proteínas alvo deste estudo como potenciais alvos vacinais, em especial a PknG, de acordo com as metodologias comumente adotadas para predição de alvos vacinais. Nos casos onde foi observada conservação em eucariotos, devem ser avaliadas as respostas geradas tanto pela porção completa da proteína quanto pela região de menor conservação.

Devido à importância dos processos nos quais as proteínas PLD, PknG, SpaC, SodC e NanH de *C. pseudotuberculosis* estão envolvidas e por serem potenciais alvos vacinais, a obtenção destas proteínas em larga escala representa um passo importante no entendimento das características e potências destas moléculas.

Neste trabalho, buscamos obter estas proteínas a partir da expressão heteróloga em *E. coli*, que é um dos sistemas de expressão de proteínas recombinantes mais utilizados mesmo em escala comercial (Baneyx, 1999; Terpe 2006). Para a expressão das proteínas recombinantes em *E. coli*, primeiramente as ORF's sintéticas de cada uma das proteínas foram construídas e clonadas no vetor pD444-NH (*DNA2.0 Inc*). O vetor pD444-NH foi selecionado por apresentar algumas características que levam à otimização da expressão da proteína recombinante. Este vetor apresenta o promotor forte T5 (Bujard *et al.*, 1987), que é reconhecido pela RNA polimerase de *E. coli*, regulado pelo operador *lac* (*lacO*), induzível por IPTG, análogo de lactose não metabolizável por *E. coli*. IPTG se associa ao repressor *lac* (*lacI*), libera o operador *lac* (*lacO*) e permite a transcrição do gene de interesse. O controle da expressão por indução com IPTG é importante para evitar a expressão basal da proteína de interesse. Além disso, o vetor escolhido permite seleção por ampicilina e adiciona uma cauda de histidina (6xHis) N-terminal que possibilita a purificação



por afinidade da proteína em coluna de níquel. Além disso, visando a otimização da expressão, as sequências das *ORF*'s tiveram os códons substituídos por códons preferenciais de *E. coli* (Anexo 1). A frequência de utilização de códons por diversos organismos variam significativamente (Gouy & Gautier, 1982) e está relacionada com a concentração da população de tRNA correspondente na célula (Ikemura, 1981). Códons raros estão associados a baixos níveis de expressão devido a interrupção da tradução (Hayes *et al.*, 2002), além de mudança de fase de leitura e incorporação de aminoácidos incorretos (Kane *et al.*, 1992; McNulty *et al.*, 2003).

Foram avaliadas neste trabalho, cinco diferentes linhagens de *E. coli* para a expressão das proteínas (PLD, PknG, SpaC, SodC e NanH): OverExpress™ C41(DE3), OverExpress™ C41(DE3) pLysS, OverExpress™ C43(DE3), OverExpress™ C43(DE3) pLysS e BL21 Star™ (DE3). Essas linhagens são derivadas da linhagem *E. coli* BL21, que é a linhagem hospedeira para expressão heteróloga mais comumente utilizada. Esta linhagem é capaz de replicar em alta taxa em meio mínimo e não é patogênica (Chart *et al.*, 2000). BL21 é deficiente nas proteases *ompT* e *lon*, que poderiam degradar a proteína recombinante (Sorensen & Mortensen, 2005). As linhagens (DE3), derivadas de BL21, carregam o gene da T7 RNA polimerase do bacteriófago T7 localizado em seu cromossomo sob controle do promotor Lac (Grossman *et al.* 1998; Pan & Malcolm 2000). Isso permite que sejam utilizados sistemas de expressão baseados no promotor T7. O plasmídeo PlysS presente nas linhagens C41(DE3) pLysS e C43(DE3) pLysS, por sua vez, confere resistência ao antibiótico cloranfenicol e codifica a lisozima do fago T7, inibidor da T7 RNA polimerase, que reduz a expressão basal a partir do promotor T7 de plasmídeos não induzidos (Studier, 1991). O vetor utilizado neste trabalho (pD444-NH) é baseado no promotor derivado do fago T5, como exposto anteriormente, que é reconhecido pela RNA polimerase de *E. coli.*, e não exigiria as funcionalidades relacionadas à utilização do promotor T7. Porém, as linhagens que permitem a expressão controlada pelo promotor T7, também permitem a utilização do promotor T5. Por isso, foi possível utilizar o banco de linhagens de expressão disponível em nosso laboratório para a expressão das proteínas recombinantes a partir do vetor pD444-NH.

As linhagens C41 (DE3) e C43 (DE3) são mutantes desenvolvidos a partir da linhagem BL21 (DE3) selecionados fenotipicamente por conferirem tolerância às proteínas tóxicas. A linhagem C41 (DE3) tem pelo menos uma mutação não caracterizada que previne a morte celular decorrente da expressão de várias proteínas recombinantes tóxicas. Já a linhagem C43 (DE3), foi derivada da C41 (DE3) por ser capaz de expressar proteínas tóxicas que a linhagem C41 (DE3) não havia sido capaz de expressar (Miroux & Walker, 1996). As linhagens C41 (DE3) e C43 (DE3) podem aumentar o nível de expressão de

proteínas de membrana e proteínas globulares (Miroux & Walker, 1996; Arechaga *et al.*, 2000).

A linhagem BL21 Star (DE3) foi desenvolvida para conferir uma alta taxa de expressão de proteínas não tóxicas. Esta linhagem contém uma mutação no gene RNaseE (*rne131*), codificador da RNase E que tem a capacidade de degradar RNA. Esta mutação então confere maior estabilidade ao RNA mensageiro (mRNA) em sistemas de expressão recombinante (Lopez *et al.*, 1999).

Neste trabalho, os resultados foram variados quanto ao desempenho de cada uma das linhagens comparando-se as cinco proteínas.

A proteína rPLD foi obtida a partir de todas as linhagens, tendo nível de expressão ligeiramente maior em BL21 Star (DE3) (Figura 22). Já a proteína rPknG, foi obtida através da expressão pelas linhagens C41 (DE3), C41 (DE3) PlysS e BL21 Star (DE3), tendo a última linhagem apresentado o maior nível de expressão (Figura 23). A proteína rSpaC, por sua vez, foi obtida por todas as células, exceto BL21 Star (DE3), em níveis de expressão aproximadamente iguais. Porém, a banda obtida apresentou um tamanho de aproximadamente 100kDa, acima do esperado que seria 86kDa (Figura 24). É válido ressaltar que o SDS-PAGE não é um método acurado de determinação de peso molecular de proteínas. O peso molecular estimado através de SDS-PAGE pode se diferenciar consideravelmente do peso real. Alguns fatores podem afetar a migração em SDS-PAGE como hidrofobicidade, carga, substituição de aminoácidos, entre outros (Hames, 1990; Fasano *et al.*, 1984).

A proteína rSodC foi obtida a partir da indução em todas as linhagens, exceto C41 (DE3) (Figura 15). A linhagem C41 (DE3), no entanto, foi a que apresentou maior nível de expressão de rNanH, que também foi obtida pelas linhagens C43 (DE3) e BL21 Star (DE3), porém, em menor nível (Figura 26). Com exceção de rSpaC, as bandas observadas após SDS-PAGE correspondentes a todas as proteínas atingiram o tamanho esperado.

Em todas as linhagens nas quais foi constatada a produção das proteínas recombinantes, pode-se perceber, comparando-se o tempo 0 e o tempo 5 de expressão, que não houve expressão basal, indicando que o promotor foi finamente regulado pelo indutor (Figuras 22-26).

Avaliando-se a cinética de expressão nas linhagens selecionadas para cada uma das proteínas, percebemos mais uma vez a ausência de expressão basal e além disso, um aumento gradual do nível de expressão no decorrer do tempo. Porém, não há grande diferença no nível de expressão entre os tempos 4 e 5 de expressão, indicando que após 5 horas o nível de expressão é estabilizado e este seria o tempo suficiente para expressão com o protocolo adotado, com temperatura de 37°C e indução por 1mM de IPTG (Figuras 27-31).

No teste de solubilidade, as proteínas recombinantes foram encontradas na fração insolúvel (Figuras 32). A superprodução de proteínas no citoplasma da bactéria pode levar a formação de proteínas insolúveis formando agregados que são chamados de corpos de inclusão com perda de conformação da proteína (Baneyx, 1999). A formação de corpos de inclusão provavelmente ocorre devido ao esgotamento do sistema de chaperonas durante a produção de altas densidades de proteína heteróloga. As chaperonas são responsáveis pelo correto enovelamento e degradação proteolítica de proteínas com má formação (Kurland & Gallat, 1996; Mergulhão *et al.*, 2005). A formação de corpos de inclusão pode ser positiva no sentido de indicar uma alta produção da proteína e por proteger a proteína de degradação proteolítica. Além disso, a obtenção na forma insolúvel pode ser um método viável para a produção de proteínas tóxicas para as células (Misawa & Kumagai, 1999). Porém, a expressão de proteínas em corpos de inclusão exige uma etapa de re-enovelamento após a purificação para que adquira sua forma biologicamente ativa o que pode inviabilizar a produção de grandes quantidades da proteína biologicamente ativa (Baneyx, 1999). Para reduzir a formação de corpos de inclusão, algumas estratégias podem ser adotadas como a coexpressão de chaperonas, diminuição da taxa de expressão por redução da temperatura e alteração das condições de cultivo, como o pH (Terpe, 2006).

Visando a obtenção das proteínas recombinantes em sua forma biologicamente ativa, alterações no protocolo de expressão foram testadas em colaboração com o Laboratório de Biologia Molecular da Universidade Estadual Paulista “Júlio de Mesquita Filho” – Unesp, sob coordenação do Prof. Dr. Raghuvir Krishnaswamy Arni. Foi adotado um protocolo com temperatura reduzida (25°C) durante um período maior de tempo (10 horas) e com uma concentração menor do indutor (IPTG 0,2mM). Essas alterações foram adotadas com a finalidade de reduzir a taxa de expressão e obter a proteína recombinante na fração solúvel. A proteína recombinante que obteve um maior nível de expressão adotando-se este protocolo foi a rPknG, se comparada às outras quatro proteínas alvo deste estudo (dados não mostrados). Além disso, a PknG foi a que apresentou melhores resultados na predição de epítopos. Dessa forma, rPknG foi selecionada para ser purificada neste trabalho também em colaboração com o Laboratório de Biologia Molecular da Unesp.

A purificação foi conduzida primeiramente em coluna de níquel (Ni-NTA), o que é possibilitado pela presença da cauda de histidina na proteína recombinante. Esta etapa foi seguida de uma purificação por troca iônica em coluna mono-Q e gel filtração em sistema AKTA purifier. Desta forma, foi possível obter a proteína rPknG purificada e de forma solúvel (Figura 33). A expressão da rPknG foi confirmada com sucesso por *Western blot* da fração purificada utilizando anticorpo Anti-His (Figura 33).

Tendo obtido a proteína rPknG purificada e após a otimização da expressão das proteínas rPLD, rSpaC, rSodC e rNanH, será possível a avaliação dessas proteínas quanto à sua imunorreatividade, potencial vacinal bem como estudos estruturais.

## 7. CONCLUSÕES

A partir deste trabalho, foi possível a caracterização *in silico* das proteínas PLD, PknG, SpaC, SodC e NanH que indicou o potencial dessas proteínas como fatores de virulência e alvos vacinais. As ORF's de cada uma dessas proteínas foram clonadas em vetor de expressão pD444-NH (DNA2.0) e foi avaliada a expressão das proteínas recombinantes em cinco diferentes linhagens de expressão de *E. coli*. A partir disso, foi indicada a melhor linhagem para expressão de cada uma das proteínas. As proteínas rPLD, rPknG, rSodC e rNanH foram expressas com o tamanho esperado, porém, não foi possível obter rSpaC com o tamanho correto. O protocolo de expressão foi otimizado para obtenção da proteína rPknG em forma solúvel, o que foi seguido de sua purificação e confirmação por *Western blot*. A caracterização e obtenção dessas proteínas por expressão heteróloga é um importante passo para o estudo de diversas estratégias de controle da LC e representa um dos primeiros estudos de obtenção de proteínas de *C. pseudotuberculosis* através de sistema de expressão heteróloga.

## 8. PERSPECTIVAS

Temos como perspectiva a expressão das proteínas rPLD, rSpaC, rSodC e rNanH de forma solúvel, bem como a purificação de cada uma delas, avaliação de antigenicidade, avaliação da resposta imune gerada por elas em cultura de macrófagos e ensaio vacinal em modelo murino. Além disso, temos a perspectiva da elucidação estrutural dessas proteínas.

## 9. REFERÊNCIAS BIBLIOGRÁFICAS

- Airelo, S., Mays, A., Oliveira, P.M.A. (2001). Manual Merck de Veterinária. Editora ROCA, São Paulo, 8ed.
- Aleman, M.; Spier, S.J., Wilson, W.D., Doherr M. *Corynebacterium pseudotuberculosis* infection in horses: 538 cases (1982-1993) (1996). J Am Vet Med Assoc. 209, 804-9.
- Aleman, M.; Spier, S.J. *Corynebacterium* infection. In: Smith, P.B. (2001). Large animal internal medicine. St Louis: Mosby. 3ed,1078-84.
- Alves, F.S.F., Pinheiro, R.R., Pires, P.C. (1997). Linfadenite caseosa: patogenia-diagnóstico-controle, Artigo Técnico. Documento nº27, Embrapa Sobral CE, Ministério da Agricultura e do Abastecimento, 1977.
- Antunes, C.A. (2011). Desenvolvimento de um Teste Intradérmico de Hipersensibilidade com antígenos secretados de *Corynebacterium pseudotuberculosis* para o diagnóstico subclínico da Linfadenite Caseosa em pequenos ruminantes. 81p. Dissertação.
- Arden, S.B., Chang, W.-H., and Barksdale, L. (1972). Distribution of Neuraminidase and N-Acetylneuraminase Lyase Activities Among *Corynebacteria*, *Mycobacteria*, and *Nocardias*. J. Bacteriol. 112, 1206–1212.
- Arechaga, I., Miroux, B., Karrasch, S., Huijbregts, R., De Kruiff, B., Runswick, M.J., Walker, J.E. (2000). Characterisation of new intracellular membranes in *Escherichia coli* accompanying large scale over-production of the b subunit of F1F0 ATP synthase. FEBS Lett 482, 215–219
- Arsenault, J.O., Girard, C., Dubreuil, P., Daignault, D.O., Galarneau, J.R., Boisclair, J., Simard, C., Bélanger, D. (2003). Prevalence of and carcass condemnation from maedi-visna, paratuberculosis and caseous lymphadenitis in culled sheep from Quebec, Canada. Prev. Vet. Med. 59, 67-81.
- Arsenault, J., Dubreuil, P., Girard, C., Simard, C., Bélanger, D. (2003). Maedi-visna impact on productivity in Quebec sheep flocks (Canada). Prev. Vet. Med. 59, 125–137.
- Av-Gay, Y., and Everett, M. (2000). The eukaryotic-like Ser/Thr protein kinases of *Mycobacterium tuberculosis*. Trends Microbiol. 8, 238–244.
- Awad, F.J. (1960). Serological investigation of *Corynebacterium pseudotuberculosis* in sheep. I. Agglutination test. Am. J. Vet. Res. 81, 251-253.
- Ayers, J.L. (1977). Caseous lymphadenitis in goats and sheep: a review of diagnosis, pathogenesis, and immunity. J. Am. Vet. Med. Assoc., v.171, n.12, 1251-54.
- Bae, T., and Schneewind, O. (2003). The YSIRK-G/S motif of staphylococcal protein A and its role in efficiency of signal peptide processing. J. Bacteriol. 185, 2910–2919.
- Baird, G.J. (2003). Current perspectives on caseous lymphadenitis. J Postgrad Clin Study. 25:62-8.
- Baird, G.J., Fontaine, M.C. (2007). *Corynebacterium pseudotuberculosis* and its role in ovine caseous lymphadenitis. J Comp Pathol. 137: 179-210.
- Baneyx, F., 1999, Recombinant protein expression in *Escherichia coli*. Current Opinion in Biotechnology, v. 10, p. 411-421.
- Barinov A., Loux V., Hammani A., Nicolas P., Langella P., Ehrlich D., Maguin E., van de Guchte M. (2009). Prediction of surface exposed proteins in *Streptococcus pyogenes*, with a potential application to other Grampositive bacteria. Proteomics. 9:61-73.

- Batey, R.G. (1986) Pathogenesis of caseous lymphadenitis in sheep goats. *Aust Vet J.* 63:269-72.
- Belanger, A. E., Hatfull, G. F. (1999) *J. Bacteriol.* 181, 6670–6678
- Belchior, S.E., Gallardo, A., Abalos, A., Jodor, N., Jensen O. (2006). Actualización sobre linfadenitis caseosa: el agente etiológico y la enfermedad. *Rev Vet Argent.* 23, 258-78.
- Bhavsar A.P., Guttman J.A., Finlay B.B., et al. (2007). Manipulation of host-cell pathways by bacterial pathogens. *Nature.* 449,827-834.
- Biberstein E.L., Knight H.D., Jang S. (1971). Two biotypes of *Corynebacterium pseudotuberculosis*. *Vet. Rec.* 89, 691–692.
- Billington, S.J., Esmay, P.A., Songer, J.G., and Jost, B.H. (2002). Identification and role in virulence of putative iron acquisition genes from *Corynebacterium pseudotuberculosis*. *FEMS Microbiol. Lett.* 208, 41–45.
- Binns, S.H., Bailey, M., Green, L.E. (2002). Postal survey of ovine caseous lymphadenitis in the United Kingdom between 1990 and 1999. *Vet Res.* 150, 263-8.
- Binns, S.H., Green, L.E., Bailey, M.(2007). Development and validation of an ELISA to detect antibodies to *Corynebacterium pseudotuberculosis* in ovine sera. *Vet. Microbiol.* 123, 169-179.
- Blatch, G.L., and Lässle, M. (1999). The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *BioEssays News Rev. Mol. Cell. Dev. Biol.* 21, 932–939.
- Brogden, K.A., Engen, R.L., Songer, J. G., Gallagher, J. (1990). Changes in ovine erythrocyte morphology due to sphingomyelin degradation by *Corynebacterium pseudotuberculosis* phospholipase D. *Microbial Pathogenesis.* 8, 157-62.
- Brown, C.C., Olander, H.J. (1987). Caseous lymphadenitis of goats and sheep: a review. *Vet. Bulletin,* v.57, n.1.
- Bujard, H., Gentz, R., Lanzer, M., Stueber, D., Mueller, M., Ibrahimi, I., Haeuptle, M.T., and Dobberstein, B. (1987). A T5 promoter-based transcription-translation system for the analysis of proteins in vitro and in vivo. *Methods Enzymol.* 155, 416–433.
- Burnside, K., and Rajagopal, L. (2011). Aspects of eukaryotic-like signaling in Gram-positive cocci: a focus on virulence. *Future Microbiol.* 6, 747–761.
- Burrell, D.H. (1978). Experimental induction of caseous lymphadenitis in sheep by intralymphatic inoculation of *Corynebacterium ovis*. *Res. Vet. Sci.* 24, 269-76.
- Burrell, D.H (1980a). A haemolysis inhibition test for detection of antibody to *Corynebacterium ovis* exotoxin. *Res. Vet. Sci.* 28, 190-194.
- Burrell, D. H. (1983). Caseous lymphadenitis vaccine. *NSW Vet.Proc.* 19, 53-57.
- Buschiazzo A., Alzari P.M. (2008) Structural insight into sialic acid enzymology. *Curr Opin Chem Biol* 12, 565–572.
- Buxton, A., Fraser, G. (1977). *Corynebacterium*. In Buxton, A., Fraser, G. *Animal Microbiology*. Blackwell Scientific Publications. 177-183.
- Cameron, H.S., McOmie, W.A. (1940). The agglutination reaction in *Corynebacterium ovis* infection. *Cornell Vet.* 30, 41-46.
- Canova, M.J., Molle, V. (2014). Bacterial Serine/Threonine Protein Kinases in Host-Pathogen Interactions. *J. Biol. Chem.*, 2014 [Epub ahead of print].

- Carminatti, R., Bahia, R., Moura-Costa, L.F., Paule, B.J.A., Vale, V.L., Regis, L., Freire, S.M., Nascimento, I., Schaer, R., Meyer, R. (2003). Determinação da sensibilidade e da especificidade de um teste de ELISA indireto para o diagnóstico de linfadenite caseosa em caprinos. *R. Ci. Med. Biol.* 2, 88-93.
- Carne, H.R. (1932). The diagnosis of caseus lymphadenitis by means of intradermal inoculation of allergic reagents. *Aust. Vet. J.* 8, 42-47.
- Carne, H.R., Onon, E.O. (1978). Action of *Corynebacterium ovis* exotoxin on endothelial cells of blood vessels. *Nature.* 271, 246-8.
- Carne, H.R., Onon, E.O. (1940). The toxin of *Corynebacterium ovis*. *J. Path. Bact.* 51, 199-212.
- Cerdeño-Tárraga, A.M., Efstratiou, A., Dover, L.G., Holden, M.T., Pallen, M., Bentley, S.D., Besra, G.S., Churcher, C., James, K.D., De Zoysa, A., Chillingworth, T., Cronin, A., Dowd, L., Feltwell, T., Hamlin, N., Holroyd, S., Jagels, K., Moule, S., Quail, M.A., Rabinowitsch, E., Rutherford, K.M., Thomson, N.R., Unwin, L., Whitehead, S., Barrell, B.G., Parkhill, J. (2003). The complete genome sequence and analysis of *Corynebacterium diphtheriae* NCTC13129. *Nucleic Acids Res.* 31, 6516-6523.
- Cesari, E. (1930). Sur Le diagnostic de la lymphadénie caéense par l'intradermo-réaction à la Preiz-Nocardine. *BULL. Acad. Vét, France.* 6, 291-295.
- Cetinkaya, B., Karahan, M., Atil, E., kalin, R., De Barre, T., Vaneechoutte, M. (2002). Identification of *Corynebacterium pseudotuberculosis* isolates from sheep and goats by PCR, *Vet. Microbiol.* V. 2359, 1-9.
- Chao, J., Wong, D., Zheng, X., Poirier, V., Bach, H., Hmama, Z., and Av-Gay, Y. (2010). Protein kinase and phosphatase signaling in Mycobacterium tuberculosis physiology and pathogenesis. *Biochim. Biophys. Acta* 1804, 620–627
- Chart, H., Smith, H. r., La Ragione, R. m., and Woodward, M. j. (2000). An investigation into the pathogenic properties of Escherichia coli strains BLR, BL21, DH5 $\alpha$  and EQ1. *J. Appl. Microbiol.* 89, 1048–1058.
- Cole, S.T., Eiglmeier, K., Parkhill, J., James, K.D., Thomson, N.R., Wheeler, P.R., Honoré, N., Garnier, T., Churcher, C., Harris, D., et al. (2001). Massive gene decay in the leprosy bacillus. *Nature* 409, 1007–1011.
- Comfort D., Clubb R.T. (2004). A comparative genome analysis identifies distinct sorting pathways in grampositive bacteria. *Infect Immun.* 72:2710–2722.
- Consortium, T.U. (2014). Activities at the Universal Protein Resource (UniProt). *Nucleic Acids Res.* 42, D191–D198.
- Corrêa, J.I. Avaliação da expressão diferencial de genes envolvidos com a virulência de *Corynebacterium pseudotuberculosis*. 114p. Tese.
- Costa Filho, G.A. (1978). Diagnóstico precoce da linfadenite caseosa dos caprinos através da intradermo-reação. *An. Universidade Federaç Rural de Pernambuco, Recife.* 161-170.
- Costa, M.P., Mcculloch, J.A., Almeida, S.S., Dorella, F.A., Fonseca, C. T., Oliveira, D.M., Teixeira, M.F.S., Laskowska E., Lipinska, B., Meyer, R., Portela, R.W., Oliveira, S.C., Myoshi, A., Azevedo, V. (2011). Molecular characterization of the *Corynebacterium pseudotuberculosis hsp60-hsp1-* operon, and evaluation of the immune response and protective efficacy induced by *hsp60* DNA vaccination in mice. *BMC Research Notes.* 4, 243.
- Coyle, M.B., Lipsky, B.A. (1982). *Coryneform* bacteria in infectious diseases: clinical and laboratory aspects, *Clin. Microbiol. Rev.* 3, 221-247.



- Cowley, S., Ko, M., Pick, N., Chow, R., Downing, K.J., Gordhan, B.G., Betts, J.C., Mizrahi, V., Smith, D.A., Stokes, R.W., et al. (2004). The Mycobacterium tuberculosis protein serine/threonine kinase PknG is linked to cellular glutamate/glutamine levels and is important for growth in vivo. *Mol. Microbiol.* *52*, 1691–1702.
- Cozzone, A.J. (2005). Role of protein phosphorylation on serine/threonine and tyrosine in the virulence of bacterial pathogens. *J. Mol. Microbiol. Biotechnol.*, *9*, 198–213.
- Dalia, A.B., Standish, A.J., Weiser, J.N. (2010). Three surface exoglycosidases from *Streptococcus pneumoniae*, NanA, BgaA, and StrH, promote resistance to opsonophagocytic killing by human neutrophils. *Infect Immun.* *78*, 2108-16.
- Das, A.K. (1998). The structure of the tetratricopeptide repeats of protein phosphatase 5: implications for TPR-mediated protein-protein interactions. *EMBO J.* *17*, 1192–1199.
- De Rose, R., Tennent, J., McWaters, P., Chaplin, P.J., Wood, P.R., Kimpton, W., Cahill, R., and Scheerlinck, J.P.Y. (2002). Efficacy of DNA vaccination by different routes of immunisation in sheep. *Vet. Immunol. Immunopathol.* *90*, 55–63.
- Denis M. (1999). Modulation of *Mycobacterium lepraemurium*: growth in murine macrophages: beneficial effect of tumor necrosis factor  $\alpha$  and granulocyte–macrophage colony stimulating factor. *Infect. Immun.* *59*, 705–707.
- Demain, A.L., and Vaishnav, P. (2009). Production of recombinant proteins by microbes and higher organisms. *Biotechnol. Adv.* *27*, 297–306.
- Dercksen, D.P., Brinkhof, J.M.A., Dekker-Nooren, T., Van Maanen, K., Bode, C.F., Baird, G., Kamp, E.M. (2000). A comparison of four serological tests for diagnosis of caseus lymphadenitis in sheep and goats. *Vet. Microbiol.* *75*, 167-175.
- Dhar, M.S., Gupta, V., and Viridi, J.S. (2013). Detection, distribution and characterization of novel superoxide dismutases from *Yersinia enterocolitica* Biovar 1A. *PLoS One* *8*, e63919.
- Dorella, F.A., Carvalho Pacheco, L., Oliveira, S.C., Miyoshi, A., and Azevedo, V. (2006). *Corynebacterium pseudotuberculosis*: microbiology, biochemical properties, pathogenesis and molecular studies of virulence. *Vet. Res.* *37*, 201–218.
- Dorella, F.A, Pacheco L.G., Seyffert N., Portela R.W., Meyer R., Miyoshi A., Azevedo V. (2009). Antigenes of *Corynebacterium pseudotuberculosis* and prospects for vaccine development. *Expert Reviews Vaccines.* *8*, 205-213.
- Dramsi S., Trieu-Cuot P., Bierne H. (2005). Sorting sortases: a nomenclature proposal for the various sortases of Gram-positive bacteria. *Res Microbiol.* *156*, 289–297.
- Dunn, K.L.R., Farrant, J.L., Langford, P.R., and Kroll, J.S. (2003). Bacterial [Cu,Zn]-Cofactored Superoxide Dismutase Protects Opsonized, Encapsulated *Neisseria meningitidis* from Phagocytosis by Human Monocytes/Macrophages. *Infect. Immun.* *71*, 1604–1607.
- Dussurget O., Stewart G., Neyrolles O., Pescher P., Young D., Marchal G. (2001). Role of *Mycobacterium tuberculosis* copper-zinc superoxide dismutase. *Infect Immun.* *69*:529-533.
- EBDA-Empresa Baiana de Desenvolvimento Agrícola S/A (BR/BA). José Arthur Hage da Silva. Patente de vacina viva contra linfadenite caseosa de caprinos. PI9903524-3, 12 ago. 1999. Disponível em: <<http://www.patentesonline.com.br/patente-de-vacina-viva-contralinfadenite-caseosa-de-caprinos-12162.html>>. Acesso em: 01 jul. 2014.
- Eggleton, D. G., Doidge, C.V., Middleton, H.D., Minty, D.W. (1991). Immunisation against ovine caseous lymphadenitis: efficacy of monocomponent *Corynebacterium pseudotuberculosis* toxoid vaccine and combined clostridial-corynebacterial vaccines. *Aust.Vet. J.* *68*, 320-321.

- Eggleton, D. G., J. A. Haynes, H. D. Middleton, and J. C. Cox. (1991). Immunisation against caseous lymphadenitis: correlation between *Corynebacterium pseudotuberculosis* toxoid content and protective efficacy in combined clostridial-corynebacterial vaccines. *Aust. Vet. J.* 68, 322-325.
- Ellis, J.A., Hawk, D.A., Mills, K.W., Pratt, D.L. (1991). Antigen specificity and activity of ovine antibodies induced by immunisation with *Corunebacterium pseudotuberculosis* culture filtrate. *Vet. Immunol. Immunopathol.* 28, 303-316.
- Exton, J.H. (1990). Signalling through phosphatidylcholine breakdown. *J. Biol. Chem.* 265:1-4.
- Fasano, O., Alfdrich, T., Tamanoi, F., Taparowsky, E., Furth, M., and Wigler, M. (1984) Analysis of the potential of the human H-ras gene by random mutagenesis. *Proc. Natl. Acad. Sci.* 81, 4008.
- Flower, D.R. (2007). Immunoinformatics and the in silico prediction of immunogenicity. An introduction. *Methods Mol. Biol. Clifton NJ* 409, 1–15.
- Foster, T.J., Ho¨o¨k, M. (1998) Surface protein adhesins of *Staphylococcus aureus*. *Trends Microbiol.* 6, 484–488
- Fridovich, I. (1995). Superoxide radical and superoxide dismutases. *Annu. Rev. Biochem.* 64, 97–112.
- Fujita, Y., Nakayama, M., Naito, M., Yamachika, E., Inoue, T., Nakayama, K., Iida, S., Ohara, N. (2014). Hemoglobin receptor protein from *Porphyromonas gingivalis* induces interleukin-8 production in human gingival epithelial cells through stimulation of the mitogen-activated protein kinase and NF-κB signal transduction pathways. *Infect. Immun.* 82, 202-1.
- Ganendren, R., Carter, E., Sorrell, T., Widmer, F., and Wright, L. (2006). Phospholipase B activity enhances adhesion of *Cryptococcus neoformans* to a human lung epithelial cell line. *Microbes Infect.* 8, 1006–1015.
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server; (In) John M. Walker (ed): *The Proteomics Protocols Handbook*, Humana Press. 571-607
- Gaupp, R., Ledala, N., and Somerville, G.A. (2012). Staphylococcal response to oxidative stress. *Front. Cell. Infect. Microbiol.* 2, 33.
- Gieffing, C., Jelencsics, K.E., Gelbmann, D., Senn, B.M., Nagy, E. (2010). The pneumococcal eukaryotic-type serine/threonine protein kinase StkP co-localizes with the cell division apparatus and interacts with FtsZ *in vitro*. *Microbiology.* 156, 1697–1707.
- Gieffing, C., Meinke, A.L., Hanner, M., et al. (2008). Discovery of a novel class of highly conserved vaccine antigens using genomic scale antigenic fingerprinting of pneumococcus with human antibodies. *J Exp Med.* 205, 117–131.
- Goebel, M., and Yanagida, M. (1991). The TPR snap helix: a novel protein repeat motif from mitosis to transcription. *Trends Biochem. Sci.* 16, 173–177.
- Gouy, M., Gautier, C. (1982). Codon usage in bacteria: correlation with gene expressivity. *Nucleic Acids Res.* 10, 7055-7074.
- Gromiha, M.M., Pujadas, G., Magyar, C., Selvaraj, S., and Simon, I. (2004). Locating the stabilizing residues in (α/β)<sub>8</sub> barrel proteins based on hydrophobicity, long-range interactions, and sequence conservation. *Proteins* 55, 316–329.
- Grossman, T.H., Kawasaki, E.S., Punreddy, S.R., Osburne, M.S. (1998). Spontaneous cAMP-dependent derepression of gene expression in stationary phase plays a role in recombinant expression instability. *Gene* 209, 95–103.

- Guimarães, A.S., Seyffert, N., Bastos, B., Portela, R.W., Meyer, R., Carmo, F.B, Cruz, J.C.M., Lage, A.P., Heinemann, M.B., Myoshi, A.; Azevedo, V., Gouveia, A. (2011). Caseus lymphadenitis: epidemiology, diagnosis, and control. *The IIOAB Journal*. 2, 33-43.
- Hanks, S.K., and Quinn, A.M. (1991). Protein kinase catalytic domain sequence database: identification of conserved features of primary structure and classification of family members. *Methods Enzymol*. 200, 38–62.
- Hanning, G., Makrides, S. C. (1998), Strategies for optimizing heterologous protein xpression in *Escherichia coli*, *Trends in Biotechnology*. 16, 54-60.
- Hansson, M., Nygren, P.A., and Ståhl, S. (2000). Design and production of recombinant subunit vaccines. *Biotechnol. Appl. Biochem.* 32 ( Pt 2), 95–107.
- Hard, G.C. (1969). Electron Microscopic Examination of *Corynebacterium ovis*. *J. Bacteriol.* 97, 1480–1485.
- Hayes, C., Bose B., Sauer R. (2002) Stop codons preceded by rare arginine codons are efficient determinants of SsrA tagging in *Escherichia coli*. *Proc Natl Acad Sci USA*, 99, 3440-3445.
- Haynes, J.A., Tkalcevic, J., and Nisbet, I.T. (1992). Production of an enzymatically inactive analog of phospholipase D from *Corynebacterium pseudotuberculosis*. *Gene* 119, 119–121.
- Hodgson, A.L., Bird, P., and Nisbet, I.T. (1990). Cloning, nucleotide sequence, and expression in *Escherichia coli* of the phospholipase D gene from *Corynebacterium pseudotuberculosis*. *J. Bacteriol.* 172, 1256–1261.
- Hodgson, A.L., Krywult, J., Corner, L.A., Rothel, J.S., and Radford, A.J. (1992). Rational attenuation of *Corynebacterium pseudotuberculosis*: potential cheesy gland vaccine and live delivery vehicle. *Infect. Immun.* 60, 2900–2905.
- Hodgson, A.L., Carter, K., Tachedjian, M., Krywult, J., Corner, L.A., McColl, M., and Cameron, A. (1999). Efficacy of an ovine caseous lymphadenitis vaccine formulated using a genetically inactive form of the *Corynebacterium pseudotuberculosis* phospholipase D. *Vaccine* 17, 802–808.
- Holstad, G. (1989). *Corynebacterium pseudotuberculosis* infection in goats.IX. The effect of vaccination agains natural infection. *Acta. Vet. Scand.* 30, 285-293
- Holt, J.G., Krieg, N.R., Sneath, P.H.A., Staley, J.T., Williams, S.T. (1991). Irregular, nosporing Gram-positive Gram-positive rods. In: Holt, J.G., Krieg, N.R., Sneath P.H.A., Staley, J.T., Williams, S.T. *Bergey's manual of determinative bacteriology*. Williams & Wilkins, Baltimore. 593.
- Hsu, T.Y., Renshaw, H.W., Livingston, C.W., Augustine, J.L., Zink, D.L., and Gauer, B.B. (1985). *Corynebacterium pseudotuberculosis* exotoxin: fatal hemolytic anemia induced in gnotobiotic neonatal small ruminants by parenteral administration of preparations containing exotoxin. *Am. J. Vet. Res.* 46, 1206–1211.
- Ibarra, J.A., and Steele-Mortimer, O. (2009). Salmonella - the ultimate insider. *Salmonella virulence factors that modulate intracellular survival*. *Cell. Microbiol.* 11, 1579–1586.
- Ikemura, T. (1981) Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: a proposal for a synonymous codon choice that is optimal for the *E. coli* translational system. *J Mol Biol*, 151, 389-409.
- Ilangovan, U, Ton-That, H, Iwahara, J, Schneewind, O, Clubb, RT. (2001). Structure of sortase, the transpeptidase that anchors proteins to the cell wall of *Staphylococcus aureus*. *Proc Natl Acad Sci USA*. 98, 6056–6061.

- Imlay, J.A., Linn, S. (1988) DNA damage and oxygen radical toxicity. *Science*. *240*, 1302–1309.
- Jenkins, G.M., and Frohman, M.A. (2005). Phospholipase D: a lipid centric review. *Cell. Mol. Life Sci. CMLS* *62*, 2305–2316.
- Johnson, E.H., Oliveira, S.C., Ribeiro, O.C., Silva, J.A. (1987). Serological detection of abscesses caused by *Corynebacterium pseudotuberculosis* using the DIG-ELISA. *Proc. IV Int. Conf. Goats, Brasilia*. 1358.
- Jolly, R.D. (1966). Some observations on surface lipids of virulent and attenuated strains of *Corynebacterium ovis*. *J. Appl. Bacteriol.* *29*, 189-196.
- Kalinowski, J., Bathe, B., Bartels, D., Bischoff, N., Bott, M., Burkovski, A., Dusch, N., Eggeling, L., Eikmanns, B.J., Gaigalat, L., et al. (2003). The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins. *J. Biotechnol.* *104*, 5–25.
- Kane, J., Violand, B., Curran, D., Staten, N., Duffin, K., Bogosian, G. (1992) Novel in-frame two codon translational hop during synthesis of bovine placental lactogen in a recombinant strain of *Escherichia coli*. *Nucleic Acids Res.* *20*, 6707-6712.
- Keith, K.E., and Valvano, M.A. (2007). Characterization of SodC, a Periplasmic Superoxide Dismutase from *Burkholderia cenocepacia*. *Infect. Immun.* *75*, 2451–2460.
- Kestintepe, H. Stabilization of *Corynebacterium ovis* antigens for serum agglutination test (1976). *Firat Univ. Vet. Fak., Dergisi.* *3*, 84-93.
- Kim, S., Oh, D.-B., Kwon, O., and Kang, H.A. (2009). Identification and functional characterization of the NanH extracellular sialidase from *Corynebacterium diphtheriae*. *J. Biochem. (Tokyo)* *147*, 523–533.
- Kim, S., Oh, D.-B., Kwon, O., and Kang, H.A. (2010). Identification and functional characterization of the NanH extracellular sialidase from *Corynebacterium diphtheriae*. *J. Biochem. (Tokyo)* *147*, 523–533.
- Kim, S., Oh, D.-B., Kang, H.A., and Kwon, O. (2011). Features and applications of bacterial sialidases. *Appl. Microbiol. Biotechnol.* *91*, 1–15.
- Kini, R.M., Evans, H.J. (1987). Structure-function relationships of phospholipases: the anticoagulant region of phospholipase A<sub>2</sub>. *The Journal of Biological Chemistry.* *30*, 14402-14407.
- Klein, J., and Sato, A. (2000). The HLA System. *N. Engl. J. Med.* *343*, 702–709.
- Knighton, D., Zheng, J., Ten Eyck, L., Ashford, V., Xuong, N., Taylor, S., and Sowadski, J. (1991). Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* *253*, 407–414.
- Koul, A., Choidas, A., Tyagi, A.K., Drlaca, K., Singh, Y., and Ullrich, A. (2001). Serine/threonine protein kinases PknF and PknG of *Mycobacterium tuberculosis*: characterization and localization. *Microbiology* *147*, 2307–2314.
- Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E.L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol.* *305*, 567-580.
- Kuamazawa, N., and Yanagawa, R. (1972) Chemical properties of the pili of *Corynebacterium renale*. *Infect Immun.* *5*, 27–30.
- Kudo, I. (2004). Diversity of phospholipase A<sub>2</sub> enzymes: foreword. *Biological & Pharmaceutical Bulletin*, *27*, 1157.

- Kurniyati, K., Zhang, W., Zhang, K., Li, C. (2013). A surface-exposed neuraminidase affects complement resistance and virulence of the oral spirochaete *Treponema denticola*. *Mol. Microbiol.* *89*, 842-56.
- Kurland, C., Gallant, J. (1996). Errors of heterologous protein expression. *Curr Opin Biotechnol.* *7*, 489–493
- Ladner, R.C. (2007). Mapping the epitopes of antibodies. *Biotechnol. Genet. Eng. Rev.* *24*, 1–30.
- Lamb, J.R., Tugendreich, S., and Hieter, P. (1995). Tetratricopeptide repeat interactions: to TPR or not to TPR? *Trends Biochem. Sci.* *20*, 257–259.
- Lan, D.T., Makino, S., Shirahata, T., Murakami, M. (2004). Hot topics in phospholipase A<sub>2</sub> field. *Biological & Pharmaceutical Bulletin.* *8*, 1179-1182.
- Langenegger, C.H.; Langenegger, J., Costa, S.G. (1987). Alérgeno para o diagnóstico da linfadenite caseosa em caprinos. *Pesq. Vet. Bras.* *7*, 27-32.
- Leamaster, B.R., Shen, D.T., Gorham, J.R., Leathers, C.W., Wells, H.D. (1987). Efficacy of *Corynebacterium pseudotuberculosis* bacterin for the immunologic protection of sheep against development of caseous lymphadenitis. *Am. J. Vet. Res.* *48*, 869-872.
- Leclere, V., Chotteau-Lelievre, A., Gancel, F., Imbert, M., Blondeau, R. (2001). Occurrence of two superoxide dismutases in *Aeromonas hydrophila*: molecular cloning and differential expression of the *sodA* and *sodB* genes. *Microbiology.* *147*, 3105–3111.
- Lee, E.-Y., Choi, D.-Y., Kim, D.-K., Kim, J.-W., Park, J.O., Kim, S., Kim, S.-H., Desiderio, D.M., Kim, Y.-K., Kim, K.-P., et al. (2009). Gram-positive bacteria produce membrane vesicles: Proteomics-based characterization of *Staphylococcus aureus*-derived membrane vesicles. *PROTEOMICS* *9*, 5425–5436.
- Lefebvre, M., Valvano, M. (2001). In vitro resistance of *Burkholderia cepacia* complex isolates to reactive oxygen species in relation to catalase and superoxide dismutase production. *Microbiology* *147*. 97–109.
- Liljeqvist, S., and Ståhl, S. (1999). Production of recombinant subunit vaccines: protein immunogens, live delivery systems and nucleic acid vaccines. *J. Biotechnol.* *73*, 1–33.
- Linder, R., Bernheimer, A.W. (1978). Effect of sphingomyelin-containing liposomes of phospholipase D from *Corynebacterium ovis* and the cytolysin from *Stoichactis helianthus*. *Biochim. Biophys. Acta.* *530*, 236-46.
- Liu, D.T., Chan, W.M., Fan D.S., Lam, D.S. (2005). An infected hydrogel buckle with *Corynebacterium pseudotuberculosis*, *Br. J. Ophthalmol.*, *89*, 245-246.
- Lund, A., Almlid T., Larsen, H.J., Steine, T. (1982). Antibodies to *Corynebacterium pseudotuberculosis* in adult goats from a naturally infected herd. *Acta Vet. Scand.* *23*, 473-482.
- Lopez, P.J., Marchand, I., Joyce, S.A., Dreyfus, M. (1999). The Cterminal half of RNase E, which organizes the *Escherichia coli* degradosome, participates in mRNA degradation but not rRNA processing in vivo. *Mol. Microbiol.* *33*, 188–199.
- Lundegaard, C., Lund, O., Nielsen, M. (2008). Accurate approximation method for prediction of class I MHC affinities for peptides of length 8, 10 and 11 using prediction tools trained on 9mers. *Bioinformatics.* *24*, 1397-1398.
- Lynch, M., Kuramitsu, H. (2000). Expression and role of superoxide dismutases (SOD) in pathogenic bacteria. *Microbes Infect.* *2*, 1252-1255.
- Madden, D.R. (1995). The Three-Dimensional Structure of Peptide-MHC Complexes. *Annu. Rev. Immunol.* *13*, 587–622.

- Maki, L.R., Shen, S., Bergston, R.C., Stetzwenbach, L.D. (1985). Diagnosis of *Corynebacterium pseudotuberculosis* infection in sheep using an enzyme-linked immunosorbent assay. *Am. J. Vet. Res.* **46**, 212-214;
- Makrides, S.C. (1996). Strategies for Achieving High-Level Expression of Genes in *Escherichia coli*. *Microbiological Reviews.* **60**, 512-538.
- Mandlik, A., Swierczynski, A., Das, A., and Ton-That, H. (2008). Pili in Gram-positive bacteria: assembly, involvement in colonization and biofilm development. *Trends Microbiol.* **16**, 33–40.
- Marchler-Bauer, A., Anderson, J.B., Derbyshire, M.K., DeWeese-Scott, C., Gonzales, N.R., Gwadz, M., Hao, L., He, S., Hurwitz, D.I., Jackson, J.D., et al. (2007). CDD: a conserved domain database for interactive domain family analysis. *Nucleic Acids Res.* **35**, 237-240.
- Marraffini, L.A., Dedent, A.C., Schneewind, O. Sortases and the art of anchoring proteins to the envelopes of gram-positive bacteria. *Microbiol Mol Biol Rev.* **70**,192–221.
- Marshall, S.A., Lazar, G.A., Chirino, A.J., and Desjarlais, J.R. (2003). Rational design and engineering of therapeutic proteins. *Drug Discov. Today* **8**, 212–221.
- Mattos-Guaraldi, A.L., Formiga, L.C., and Andrade, A.F. (1998) Trans-sialidase activity for sialic acid incorporation on *Corynebacterium diphtheriae*. *FEMS Microbiol. Lett.* **168**, 167-172.
- Mazmanian, S.K., Liu, G., Ton-That, H., Schneewind, O. (1999). *Staphylococcus aureus* sortase, an enzyme that anchors surface proteins to the cell wall. *Science.* **285**, 760–763.
- McNamara, P.J., Bradley, G.A., Songer, J.G. (1994). Targeted mutagenesis of the phospholipase D gene results in decreased virulence of *Corynebacterium pseudotuberculosis*. *Mol Microbiol.* **12**, 921-930.
- McNamara, P.J., Cuevas, W.A., Songer, J.G. (1995). Toxic phospholipases D of *Corynebacterium pseudotuberculosis*, *C. ulcerans* and *Arcanobacterium haemolyticum*: cloning and sequence homology. *Gene.* **156**, 113-118.
- McNulty, D., Claffee, B., Huddleston, M., Porter, M., Cavnar, K., Kane, J. (2003). Mistranslational errors associated with the rare arginine codon CGG in *Escherichia coli*. *Protein Expr Purif.* **27**, 365-374.
- Menzies, P. I. (1998). Caseous lymphadenitis of sheep and goats. In AIELLO, S.; MAYS, A. *The Merck Veterinary Manual*, Whitehouse Station, NJ, Merck & Co. **8**, 55.
- Menzies, P.I., Hwang, T.I., Prescott. J.F. (2004). Comparison of an interferon-gamma to a phospholipase D enzyme linked immunorbent assay for diagnosis of *Corynebacterium pseudotuberculosis* infection in experimentally infected goats. *Vet. Microbiol.* **100**, 129-137.
- Merchant, I.A., Packer, R.A. *The Genus Corynebacterium*, in: Merchant, I.A., Packer, R.A. *Veterinary bacteriology and virology*. The Iowa State University Press. 425-440.
- Mergulhao, F.J., Summers, D.K., Monteiro, G.A. (2005). Recombinant protein secretion in *Escherichia coli*. *Biotechnol Adv.* **23**,177–202.
- Meyer, R., Carminati, R., Cerqueira, R. B., Vale, V., Viegas, S., Martinez, T., Nascimento, I., Scaer, R., Silva, J. A. H., Ribeiro, M., Régis, M., Paule, B., Freire, S. M. (2002). Avaliação da resposta imune humoral em caprinos inoculados com uma vacina viva atenuada liofilizada contra *Corynebacterium pseudotuberculosis* *R. Ci. Méd. Biol.* **1**, 42-48.
- Meyer, R. (2003). *Corynebacterium pseudotuberculosis* e o hospedeiro caprino: aspectos da prevalência, do diagnóstico e da imunidade. (Doutorado em Imunologia). Instituto de Ciências da Saúde, Universidade Federal da Bahia, Salvador.

- Miroux, B., and Walker, J.E. (1996). Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J. Mol. Biol.* *260*, 289–298.
- Misawa, S., Kumagai, I. (1999). Refolding of therapeutic proteins produced in *Escherichia coli* as inclusion bodies. *Biopolymers.* *51*, 297-307.
- Modlin, R.L., Melancon-Kaplan, J., Young, S.M.M. *et al.* (1988). Learning from lesions: patterns of tissue inflammation in leprosy. *Proc. Natl Acad. Sci.* *85*, 1213–1217.
- Monti, E., Preti, A., Venerando, B., and Borsani, G. (2002). Recent development in mammalian sialidase molecular biology. *Neurochem. Res.* *27*, 649–663.
- Nakane, A., Numata, A., Chen, Y., Minagawa, T. (1991). Endogenous  $\gamma$  interferon-independent host resistance against *Listeria monocytogenes* infection in CD4+ T cell and asialo GM1+ cell-depleted mice. *Infect. Immun.* *59*, 3439–3445.
- Nain, S.P.S., Garg, D.N., Chandiramani, N.K. (1984). An agar-gel-immunoprecipitation test for detection of *Corynebacterium pseudotuberculosis ovis* antibodies in sheep and goats sera. *Indian J. Comp. Microbiol. Infect. Dis.* *5*, 93-96.
- Nathan, C., Shiloh, M.U. (2000). Reactive oxygen and nitrogen intermediates in the relationship between mammalian hosts and microbial pathogens. *Proc Natl Acad Sci USA.* *97*, 8841-8848.
- Niebisch, A., Kabus, A., Schultz, C., Weil, B., and Bott, M. (2006). Corynebacterial Protein Kinase G Controls 2-Oxoglutarate Dehydrogenase Activity via the Phosphorylation Status of the OdhI Protein. *J. Biol. Chem.* *281*, 12300–12307.
- Nielsen, M., and Lund, O. (2009). NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. *BMC Bioinformatics* *10*, 296.
- Nozaki, C.N., Faria, M.A.R., Machado, T.N.M. (2000). Extirpação cirúrgica dos abscessos da linfadenite caseosa em caprinos. *Arq Inst Biol.* *67*, 187-9.
- O'Hare, H.M., Durán, R., Cerveñansky, C., Bellinzoni, M., Wehenkel, A.M., Pritsch, O., Obal, G., Baumgartner, J., Vialaret, J., Johnsson, K., Alzari, P.M. (2008). Regulation of glutamate metabolism by protein kinases in mycobacteria. *Mol. Microbiol.* *70*, 1408-23.
- Oliveira, G.C. Atividades da Rede Genoma de Minas Gerais no biênio 2006-2007 - apoio à Agroindústria: Sequenciamento do genoma da bactéria *Corynebacterium pseudotuberculosis*. Disponível em: <<http://bioinfo.cpqrr.fiocruz.br/files>> Acesso em: 12 de agosto de 2013.
- Pacheco, L.G.C., Pena, R.R., Castro, T.L.P., Dorella, F.A., Bahia, R.C., Craminati, R., Frota, M.N.L., Oliveira, S.C., Meyer, R., Alves, F.S.F., Myoshi, A., Azevedo, V. (2007). Multiplex PCR Assay for identification of *Corynebacterium pseudotuberculosis* from pure cultures and for rapid detection of this pathogen in clinical samples. *56*, 480-486.
- Pan, S.H., Malcolm, B.A. (2000). Reduced background expression and improved plasmid stability with pET vectors in BL21 (DE3). *Biotechniques.* *29*, 1234–1238.
- Paton, M.W., Mercy, A.R., Wilkinson, F.C., Gardner, J.J., Sutherland, S.S., Ellis, T.M. (1988). The effects of caseous lymphadenitis on wool production and bodyweight in young sheep. *Aust. Vet. J.* *65*, 117-9.
- Paton, M.W., Sutherland, S.S., Rose, I.R., Hart, R.A., Mercy, A.R., Ellis, T.M. (1995). The spread of *Corynebacterium pseudotuberculosis* infection to unvaccinated and vaccinated sheep. *Aust. Vet. J.* *72*, 266–269.
- Paton M.W., Walker S.B., Rose I.R., Watt G.F. (2003). Prevalence of caseous lymphadenitis and usage of caseous lymphadenitis vaccines in sheep flocks, *Aust. Vet. J.* *81*, 91–95.

- Paule, B.J.A., Azevedo, V., Regis, F.F., Carminati, R., Bahia, C.R., Vale, V.L.C., Moura-Costa, L.F., Freire, S.M., Nascimento, I., Schaer, R., Goes, A.M., Meyer, R.. (2003). Experimental *Corynebacterium pseudotuberculosis* primary infection in goats: kinetics of IgG and interferon- $\gamma$  production, IgG avidity and antigen recognition by Western blotting. *Vet. Immunol. Immunopathol.* **96**, 129-139.
- Peel, M.M., Palmer, G.G., Stacpoole, A.M., Kerr, T.G. (1997). Human lymphadenitis due to *Corynebacterium pseudotuberculosis*: report of ten cases from Australia and review. *Clin Infect Dis.* **24**, 185-91.
- Pépin, M., Pittet, J-C., Olivier, M., Gohint, I. (1994). Cellular composition of *Corynebacterium pseudotuberculosis* pyogranulomas in sheep. *J. Leukocyte Biol.* **56**, 666–670.
- Petersen, T.N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786.
- Peti, W., Page, R. (2007). Strategies to maximize heterologous protein expression in *Escherichia coli* with minimal cost. *Protein Expression and Purification.* **51**, 1–10.
- Petri, W.A., Mann, B.J., Huston, C.D. (2010). Microbial adherence. In: Mandell, G.L., Bennett, J.E., Dolin, R. (editors), *Mandell, Douglas, Bennett's. Principles and Practice of Infectious Diseases.* 7. ed. Churchill Livingstone Elsevier, 15.
- Piddington, D.L., Fang, F.C., Laessig, T., Cooper, A.M., Orme, I.M., Buchmeier, N.A. (2001). Cu,Zn superoxide dismutase of *Mycobacterium tuberculosis* contributes to survival in activated macrophages that are generating an oxidative burst. *Infect Immun.* **69**, 4980-4987.
- Piontkowski, M.D., Shivvers, D.W. (1998). Evaluation of a commercially available vaccine against *Corynebacterium pseudotuberculosis* for use in sheep. *J. Am. Vet. Med. Assoc.* **212**, 1765–1768.
- Prisic, S., Dankwa, S., Schwartz, D., Chou, M.F., Locasale, J.W., Kang, C.-M., Bemis, G., Church, G.M., Steen, H., and Husson, R.N. (2010). Extensive phosphorylation with overlapping specificity by *Mycobacterium tuberculosis* serine/threonine protein kinases. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 7521–7526.
- Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., and Lopez, R. (2005). InterProScan: protein domains identifier. *Nucleic Acids Res.* **33**, W116–W120.
- Raggiaschi, R., Gotta, S., and Terstappen, G.C. (2005). Phosphoproteome analysis. *Biosci. Rep.* **25**, 33–44.
- Rammensee, H.G., Friede, T., and Stevanović, S. (1995). MHC ligands and peptide motifs: first listing. *Immunogenetics* **41**, 178–228.
- Rebecchi, M.J., and Pentylala, S.N. (2000). Structure, Function, and Control of Phosphoinositide-Specific Phospholipase C. *Physiol. Rev.* **80**, 1291–1335.
- Renshaw, H.W., Graff, V.P., Gates, N.L. (1979). Visceral lymphadenitis in thin ewe syndrome: Isolation of *Corynebacterium*, *Staphylococcus* and *Moraxella spp* from internal abscesses in emaciated ewes. *Am. J. Vet. Res.* **40**, 1110-1114.
- Ribeiro, M.G., Júnior, J.G.D., Paes, A.C., Barbosa, P.G., Júnior, G.N., Listoni, F.J.P. (2001). Punção aspirativa com agulha fina no diagnóstico de *Corynebacterium pseudotuberculosis* na linfadenite caseosa caprina. *Arq. Inst> Biol.* **68**, 23-28.
- Richmond, G.S., Smith, T.K. (2011). Phospholipases A<sub>1</sub>. *Int. J. Mol. Sci.* **12**, 588-612.
- Rogan, D., and Babiuk, L.A. (2005). Novel vaccines from biotechnology. *Rev. Sci. Tech. Int. Off. Epizoot.* **24**, 159–174.



- Rudolph, R. (1996). Successful protein folding on an industrial scale. In: Cleland, J.L., Craik, C.S. Protein Engineering: Principles and Practice. Wiley-Liss, Inc. 283-298.
- Saha, S., and Raghava, G.P.S. (2006). Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins* 65, 40–48.
- Saha, S., Raghava, G.P.S. (2004). BcePred: Prediction of Continuous B-Cell Epitopes in Antigenic Sequences Using Physico-chemical Properties. In G.Nicosia, V.Cutello, P.J. Bentley, J.Timis. ICARIS LNCS. 3239, 197-204.
- Sambrook, J., Russel, D. W. (2000). Molecular Cloning – A laboratory Manual. Cold Spring Harbor Press, New York.
- San Mateo, L.R., Toffer, K.L., Orndorff, P.E., Kawula, T.H. (1999). Neutropenia restores virulence to an attenuated Cu,Zn superoxide dismutase-deficient *Haemophilus ducreyi* strain in the swine model of chancroid. *Infect Immun.* 67, 5345-5351.
- Sanjay, M.K., Shrideshikan, S.M., Usha, M.S., Philipraj, A., Gaddad, S.M., and Shivannavar, C.T. (2010). Detection, amplification & sequence homology of sodC in clinical isolates of *Salmonella* sp. *Indian J. Med. Res.* 131, 565–570.
- Sansone, A., Watson, P.R., Wallis, T.S., Langford, P.R., and Kroll, J.S. (2002). The role of two periplasmic copper- and zinc-cofactored superoxide dismutases in the virulence of *Salmonella choleraesuis*. *Microbiology* 148, 719–726.
- Santos, A.R., Pereira, V.B., Barbosa, E., Baumbach, J., Pauling, J., Rottger, R., Turk, M.Z., Silva, A., Miyoshi, A., and Azevedo, V. (2013). Mature Epitope Density - A strategy for target selection based on immunoinformatics and exported prokaryotic proteins. *BMC Genomics* 14, S4.
- Santos, R., Franza, T., Laporte, M.L., Sauvage, C., Touati, D., Expert, D. (2001). Essential role of superoxide dismutase on the pathogenicity of *Erwinia chrysanthemi* strain 3937. *Mol Plant Microbe Interact.* 14, 758–767.
- Scarselli, M., Giuliani, M.M., Adu-Bobie, J., Pizza, M., and Rappuoli, R. (2005). The impact of genomics on vaccine design. *Trends Biotechnol.* 23, 84–91.
- Schauer, R. (2000). Achievements and challenges of sialic acid research. *Glycoconj. J.* 17, 485–499
- Scherr, N., Honnappa, S., Kunz, G., Mueller, P., Jayachandran, R., Winkler, F., Pieters, J., and Steinmetz, M.O. (2007). Structural basis for the specific inhibition of protein kinase G, a virulence factor of *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U. S. A.* 104, 12151–12156.
- Scherr, N., Müller, P., Perisa, D., Combaluzier, B., Jenö, P., and Pieters, J. (2009). Survival of Pathogenic *Mycobacteria* in Macrophages Is Mediated through Autophosphorylation of Protein Kinase G. *J. Bacteriol.* 191, 4546–4554.
- Schneewind, O., Model, P., and Fischetti, V.A. (1992). Sorting of protein a to the staphylococcal cell wall. *Cell* 70, 267–281.
- Schneewind, O., Mihaylova-Petkov, D., and Model, P. (1993). Cell wall sorting signals in surface proteins of gram-positive bacteria. *EMBO J.* 12, 4803–4811.
- Selim, A.S. (2001). Oedematous skin disease of buffalo in Egypt, *J. Vet. Med. B Infect. Dis. Vet. Public Health.* 48, 241–258.
- Seyffert, N. *et al.* (2010) High seroprevalence of caseous lymphadenitis in Brazilian goat herds revealed by *Corynebacterium pseudotuberculosis* secreted proteins-based ELISA. *Res. Vet. Sc.* 88, 50-55.

- Shen, D.T, Jen, L.W., Gorham, J.R. (1982). The detection of *Corynebacterium pseudotuberculosis* antibody in goats by the enzyme-linked immunosorbent assay (ELISA). Proc. 3<sup>rd</sup> Int. Conf. Goat Prod. Dis., Tucson, Arizona. 445-448.
- Shen, D.-K., Noodeh, A.D., Kazemi, A., Grillot, R., Robson, G., and Brugère, J.-F. (2004). Characterisation and expression of phospholipases B from the opportunistic fungus *Aspergillus fumigatus*. FEMS Microbiol. Lett. 239, 87–93.
- Shigidi, M.T.A. (1978). An indirect haemagglutination test for the serodiagnosis of *Corynebacterium ovis* infection in sheep. Res. Vet. Sci. 24, 57-60.
- Shigidi, M.T.A. (1974). Antigenic relationship of various isolates of *Corynebacterium pseudotuberculosis*. Bull. Epizootic Dis. Africa. 22, 263-269.
- Shigidi, M.T.A. (1979). A comparison of five serological tests for the diagnosis of experimental *Corynebacterium ovis* infection in sheep. Brit. Vet. J. 135, 172-177.
- Sibbald, M.J.J.B., van Dij, J.M.L. (2009). Secretome Mapping in Gram-Positive Pathogens. In Karl Wooldridge (ed.), Bacterial Secreted Protein: Secretory Mechanisms and Role in Pathogenesis. Caister Academic Press.193-225.
- Simeone, R., Bottai, D., Brosch, R., et al. (2009). ESX/type VII secretion systems and their role in host-pathogen interaction. Curr Opin Microbiol. 12, 4-10.
- Simmons, C.P., Hodgson, A.L., Strugnell, R.A. (1997). Attenuation and vaccine potential of *aroQ* mutants of *Corynebacterium pseudotuberculosis*. Infect. Immun. 65, 3048–3056.
- Smith, S.G., Wilson, T.J., Dow, J.M., Daniels, M.J. (1996). A gene for superoxide dismutase from *Xanthomonas campestris* pv. *campestris* and its expression during bacterial-plant interactions. Mol Plant Microbe Interact. 9, 584–593.
- Song, J.K., and Rhee, J.S. (2001). Enhancement of stability and activity of phospholipase A1 in organic solvents by directed evolution. Biochim. Biophys. Acta BBA - Protein Struct. Mol. Enzymol. 1547, 370–378.
- Songer, J.G. (1997). Bacterial phospholipases and their role in virulence. Trends Microbiol. 5, 156–161.
- Songer, J.G., Libby, S.J., Iandolo, J.J., Cuevas, W.A. (1990). Cloning and expression of the phospholipase D gene from *Corynebacterium pseudotuberculosis* in *Escherichia coli*. Infect. Immun. 58, 131-6.
- Sørensen, H.P., and Mortensen, K.K. (2005). Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. J. Biotechnol. 115, 113–128.
- Stanford K., Brogden K.A., McClelland L.A., Kozub G.C., Audibert F. (1998). The incidence of caseous lymphadenitis in Alberta sheep and assessment of impact by vaccination with commercial and experimental vaccines, Can. J. Vet. Res. 62, 38–43.
- Stavrinides, J., McCann, H.C., Guttman, D.S., et al. (2008). Host-pathogen interplay and the evolution of bacterial effectors. Cell Microbiol. 10, 285-292.
- Stout, T.J., Foster, P.G., and Matthews, D.J. (2004). High-throughput structural biology in drug discovery: protein kinases. Curr. Pharm. Des. 10, 1069–1082.
- Studier, F.W. (1991) Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. J Mol Biol. 219, 37–44

- Tachedjian, M., Krywult, J., Moore, R.J., Hodgson, A.L. (1995). Caseous lymphadenitis vaccine development: site-specific inactivation of the *Corynebacterium pseudotuberculosis* phospholipase D gene. *Vaccine*. 13, 1785-92.
- Tai, S.S. (2006). Streptococcus pneumoniae protein vaccine candidates: properties, activities and animal studies. *Crit. Rev. Microbiol.* 32, 139-153.
- Tashjian, J.J., Campbell, S.G. (1983). Interaction between caprine macrophages and *Corynebacterium pseudotuberculosis*: an electron microscopy study, *Am. J. Vet. Res.* 44, 690-693.
- Taylor, G. (1996). Sialidases: structures, biological significance and therapeutic potential. *Curr. Opin. Struct. Biol.* 6, 830\_837.
- Terpe, K. (2006). Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.* 72, 211–222.
- Thompson, J. D., Higgins, D. G., Gibson, T. J. (1994). Clustal W: Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weigthing, Position-Specific Gap Penalties and Weigth Matrix Choice”, *Nucleic Acids Researchs.* 22, 4673-4680.
- Tonello, F., and Zornetta, I. (2012). Bacillus anthracis Factors for Phagosomal Escape. *Toxins* 4, 536–553.
- Ton-That, H., and Schneewind, O. (2003). Assembly of pili on the surface of *Corynebacterium diphtheriae*. *Mol. Microbiol.* 50, 1429–1438.
- Ton-That, H., and Schneewind, O. (2004). Assembly of pili in Gram-positive bacteria. *Trends Microbiol.* 12, 228–234.
- Trost, E., Ott, L., Schneider, J., Schröder, J., Jaenicke, S., Goesmann, A., Husemann, P., Stoye, J., Dorella, F.A., Rocha, F.S., Soares, S.C., D'afonseca, V., Miyoshi, A., Ruiz, J., Silva, A.; Azevedo, V., Burkovski, A., Guiso, N., Join-lambert, O.F., Kayal, S., Tauch, A. (2010). The complete genome sequence of *Corynebacterium pseudotuberculosis* FRC41 isolated from a 12-year-old girl with necrotizing lymphadenitis reveals insights into gene-regulatory networks contributing to virulence. *BMC Genomics.* 11, 728.
- Ulbrich-Hofmann, R., Lerchner, A., Oblozinsky, M., and Bezakova, L. (2005). Phospholipase D and its application in biocatalysis. *Biotechnol. Lett.* 27, 535–544.
- Unanian, M. M., Feliciano-Silva, A.E.D., Pant, K. P. (1985). Abscesses and caseous lymphadenitis in goats in tropical semi-arid northeast Brazil. *Trop. Anim. Health Prod.* 17, 57-62.
- Vale, V.L.C., Freire, S., Ribeiro, M., Regis, L., Bahia, R., Carminati, R., Paule, B.J.A., Nascimento, I., Meyer, R. (2003). Reconhecimento de antígenos por anticorpos de caprinos naturalmente infectados ou imunizados contra *Corynebacterium pseudotuberculosis*. *Revista de Ciências Médicas e Biológicas, Salvador.* 2, 192-198.
- Valli, V.E., Parry, B.W. (2007). Caseous lymphadenitis. In: Juby KVF, Kennedy PC, Palmer N, editors. *Pathology of domestic animals.* Academic Press. 3, 238-40.
- Varki, A. (1999) Sialic acids. In: Varki A, Cummings R, Esko J, Freeze, Hart G, Marth J (eds) *Essential of glycobiology.* Cold Spring Harbor Laboratory, Cold Spring Harbor. 195–209
- Vimr, E.R., Lichtensteiger, C. (2002). To sialylate, or not to sialylate: that is the question. *Trends Microbiol* 10, 254–257

- Vimr, E.R., Kalivoda, K.A., Deszo, E.L. (2004). Steenberg SM: Diversity of microbial sialic acid metabolism. *Microbiol Mol Biol Rev.* 68:132-153.
- Walburger, A., Koul, A., Ferrari, G., Nguyen, L., Prescianotto-Baschong, C., Huygen, K., Klebl, B., Thompson, C., Bacher, G., and Pieters, J. (2004). Protein kinase G from pathogenic mycobacteria promotes survival within macrophages. *Science* 304, 1800–1804.
- Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., and Barton, G.J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189–1191.,
- Walker, J., Jackson, H., Brandon, M.R., Meeusen, E.(1991). Lymphocyte subpopulations in pyogranulomas of caseous lymphadenitis. *Clin. Exp. Immunol.* 86, 13–18.
- Wang, J., Kolko, M. (2010). Phospholipases A2 in ocular homeostasis and diseases. *Biochimie.* 92, 611-619.
- Whitmore, S.E., Lamont, R.J. (2012). Tyrosine phosphorylation and bacterial virulence. *Int. J. Oral Sci.* 4, 1-6.
- Wilks, K.E., Dunn, K.L., Farrant, J.L., Reddin, K.M., Gorringer, A.R., Langford, P.R., Kroll, J.S. (1998). Periplasmic superoxide dismutase in meningococcal pathogenicity. *Infect Immun.* 66,213-217.
- Williamson, L.H. (2001). Caseous lymphadenitis in small ruminants. *Vet. Clin. North Am. Food Anim. Pract.* 17, 359–371.
- Wooldridge. K. *et al.* (2009). Bacterial secreted proteins secretory mechanisms and role in pathogenics, 512p.
- Yamada, M, Nakane, A. (1999). Tumor necrosis factor  $\alpha$  and  $\gamma$  interferon are required for the development of protective immunity to secondary *Corynebacterium pseudotuberculosis* infection in mice. *J. Vet. Med. Sci.* 61, 1203–1208..
- Yanagawa, R., and Honda, E. (1976). Presence of pili in species of human and animal parasites and pathogens o the genus *Corynebacterium*. *Infect Immun.* 13, 1293–1295.
- Youn, H.D., Kim, E.J., Roe, J.H., Hah, Y.C., and Kang, S.O. (1996). A novel nickel-containing superoxide dismutase from *Streptomyces* spp. *Biochem. J.* 318, 889–896.
- Yozwiak, M.L., and Songer, J.G. (1993). Effect of *Corynebacterium pseudotuberculosis* phospholipase D on viability and chemotactic responses of ovine neutrophils. *Am. J. Vet. Res.* 54, 392–397.
- Yu, N.Y., Wagner, J.R., Laird, M.R., Melli, G., Rey, S., Lo, R., Dao, P., Sahinalp, S.C., Ester, M., Foster, L.J., et al. (2010). PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinforma. Oxf. Engl.* 26, 1608–1615.

## 10. ANEXOS

Anexo 1 – Sequências das *ORF's pld, pkng, spac, sodc e nanh* otimizadas com códons preferenciais de *E. coli*.

Anexo 2 - Predição de peptídeo sinal das proteínas SpaC, SodC e NanH.

Anexo 3 – Alinhamentos de sequência.

Anexo 4 – Representação esquemática das plasmídeos de expressão

## **ANEXO 1**

**Sequências das *ORF's pld, pkg, spac, sodc e nanh* otimizadas com códons preferenciais de *E. coli*.**

### Sequência otimizada – ORF pld

ATGATCATCATGGCCATTATGCTGCCGGTTGGCAACGCAGCGGCAGCTCCGGTTGTGCATAATCCGGCTTCCACCGGAATCG  
CCCGGCTACGCGATTGCCCTCGCGCTTCTGACCACCCAGGGTGTGGACGACGCCGTGGCGATCGGTGCGAACGCACTGGAGA  
TCGACTTTACCGCGTGGGGCCGTGGTTGGTGGCGAGACCACGACGGCATCCCGACCTCTGCCGGTGCAGCGCCGAAGAGATT  
TTCAAGCACATCGCAGACAAACGTAACAGGGCGCGAACATTACTTTTTACCTGGCTGGATATCAAAAATCCGGACTATTGTCTG  
TGATGCGCGTAGCGTCTGCAGCATTAAATGCGCTGCGCGACCTGGCGCGTAAGTATCTGGAACCAGCGGGCGTGCCTGTTCTGT  
ACGGCTTACAAAACCGTCCGGTGGTCCGGCATGGAAAACGATTACGGCAGATCTGCGTGATGGTGAGGCTGTGCGTTGAGC  
GGTCCGGCTCAAGACGTGCTGAACGACTTCGCACGCAGCGAGAACAAGATTCTGACCAAGCAGAAAATCCGGGATTACGGTTA  
TTACAACATTAACCAGGGTTTCGGTAATTGCTACGGTACCTGGAACCGTACGTGTGATCAATTGCGTAAAAGCAGCGAAGCCC  
CGCATCAAGGTAAGCTGGTAAAACGTTGGCTGGACCATCGCGACCGGTCAAGATGCACGTGCAATGATTTACTGGGCAAG  
GCAATATGCGATGGCTGATTTTTGGCTTCAAAATCACTCATTTCTATCGTCACGCGGACACGAGAACAGCTTTAAGGCTAT  
TAAGCGCTGGGTAGACAAGCATAGCGCGACCACCTGGCGACGGTGGCCGATAATCCTTGGTAAGGT

### Sequência otimizada – ORF pkng

ATGGAAGATGATCTGAGCGGTCTGCTGAACGACCTGAGCAGCATTCAAAAGGACACTGATACCGACGCCCGTTTCGCGCGAGAA  
ATCCATTAGCACGTTCCGCTCTCGTCTGGCACCAACCGTGATGACCGTACCGTTGCGAATGGTATGGTCACCCCTGCCGTTCA  
TTCTGTCTCGACGGCAGAGGAAATGCTGAAAGACGATGCCTACATTGAGCAAAAGGGCCCTGGAGAAGCCGCTGCTGCACCCG  
GGTGATATCCTGGCAAACCAATACGAAGTGCAGCGGTGTGATTTGCCATGGTGGCATGGGCTGGATCTATGTTGCGCTCTGACCG  
TAACGTGAGCGGTCTGCTGCTGTTGAAAGGCATGCGCGATAAAAGCGAACGCGAATCCGCACGATTACGGTGCAGCCGTTG  
CGGAGCGTGAGTTTCTGGCGGACATCACGCACCCTGGTATCGTAAAAGCTACAATTTTATTGACGACCACAGAGTTGAGGGT  
GGCTTCAATTGTTATGGAGTATGTTCCAGGTCGCTCCTTGGCAGACCGTGCAGGAACAGCCTGGCGGCTTCTCAGCATCGA  
TATTGCGATTGCTTACATCCTGGAGATTCTGCCGGCATTTGGACTATCTGCATAGCCCGGGTGTGCTGTACAATGATCTGAAAC  
CGGACAATATCATTGTGCTGAGGACCAAGTGAAGCTGATTGACCTGGGTGCTGTAAGCGGCATCGGTGCGTTTGGTTACATT  
TACAGCACGAAAGGCTTTCAGGCACCGGAAGTTGCGACCGATGGTCCGACCATCGCAAGCGACATCTATACCATTGGTCTGTAC  
GCTGGCGGCTCTGACCGTTGATCTGCCGAGCGCGGATGGTGTCTATCTGCCGGCATCCCGTCCCCGGACGAAGAGAAACTGT  
TCGCCAGAAATTTGAGCTTCTATCGCCTGCTGCGTGTATGAGCAGCGATCCGGCGAAGCGTTTTGCGAGCATCCGTGAA  
TTAGAAACGCAACTGTTTGGCATTCTGCGTGAATACTTAGCACTGGCCGAGCACCGCCAATTCGCCGACAGCACAGCCTCTA  
TAGCCCGCAGCGTAGCACCTTTGGTACGAAACACATGGTCTTTTCGTACGGACAGCTGATCGATGGCATTGAGCGCAACGTTT  
GTATTACCAGCGAGGAAGTTAACGCCCGGTTGCCGGTCCCGCTGCTGGATCGTACCGATCCGGGTGCGATTCTGATTAGCGGT  
GCCAGCTACGCGGAACCGAGCCAGACCTTGCAAACGCTGCGCGACGCGATGGCACAAGAAGAATTTGCGAATAGCAAAGAAT  
CCCCTTGGGTGTTGTGCGTACTCTGCTGGACCTGGGTTTTATCGATGAAGCAGCAGCTGGCTGGACACTTGGATGCAACCC  
TGTCTGACGACTGGCGCCATCAGTGGTATTCTGGTGTGACGAGCCTCTTCTTTCAGCAGTACGTTGAGCGCAGCGCCATTTT  
AATGAAGTTTACAATATCCTGCCGGGTGAGGCGCACCGAAACTGGCGCGTGCAGGATGCTGCTGCAAGGCAT  
GGGCATGGATAGCACCCACTGCTGTCCCGGATACCGCCGTTAGCGCAGCGGAGCTGAAGGTAACACCATCGGTATGTGGT  
CAGAGCTGACCAAGGACCCGGAGACTTTGCGCTTCAAGGCATTGTACCTGTATGCGCTGGTGTGGCGTACCAACCCGACC  
GTGTCTCCGCTTTCCGGCTGGCGCGTACGTTGGTGGGAGAACCAAGTGGACCTGGCTGTTAGCACCCCTGGACCAAGTGC  
GCAGAACTCCACGATCGCCGATGGCCGAGCTGACCGGATTTTGCACCTGATCAGCGGTGACCTGAGCGAGAGCCGCGATCC  
GTCGTGCCGACGTCGCTGGAGTCGATCCCGACGAATGACCCGCTTTCTTGCAGATCAAGATCGCGATTATGAATGCGGCA  
CTGACCTGGCTGCGTCAGAGCCGTTTGAAGCTGCTGCCGCCGATAACGATCTGTTTACTACCGTTTACCCAGGTTGGCCT  
GCGTACCGGCTGTATGACAGCCTGCGTCTGCTGGCGGTTCTGCGCCAAATGTTTCATCACCGCTATACCTTGGTGGATATGG  
CGAACCAAGTCCGTCCGATGACGTGGTTTTAAGGT

### Sequência otimizada – ORF spac

ATGAAAGTCGAGATTCGTTTTTCAGACGGGCAGCAAAATCAGCACTCCGCTTACCCCGAGCGTGTATGTTAATGGCTGCGTCTGA  
AGAATTCGAGAATACCGAGCCGATCGCCCAATTCGCCGGTAGGTCGCTCGGTTGATCCGACGGTTTTGTAGCCCGAACGCCG  
AGGCAAAAGTCTGGATCGGCATGAGCCAACACGAATCGTGGTGAACGTTTAAACGACCGCAATAGCACCGATCTCTACGTT  
CAATCGTTCAATCGCAATACCGGTCTGGATGAGTTCAAACCGGTGGGTGCGCGCACGCCGTGGGTTTACAATGCAATGGCGTA  
TAATCCGAAAGAGGGCTACATCTATGCCATTAGCCAGGGCCGTTTGA AAAACCCCTGCAGTCTAGCAAGCTGCGCATCTACGACG  
AAGATCCGAACATCTCGGGTCACTGCTGCAAAATTTCCCGGTCAATGCCGGCGTGAAGATCTGGGCGCTATTATTCGTT  
TTGAATGGTCCCGCTGACTGCTGATGGCCGAACGACTGTGGGTGGTATGACGTCGGTATCATTGACGGTAATGCGCTTTA  
TCTGGTATCCAACAGCAGCCAGAGCGGTACCCACAACCTTGTACACGCTGGACCTGGACACCCGCTTGGCGACCGTGGTGCCT  
CTAACACCGGTTTTTCGAACGATTATACCAGCACTGGTAAAACGACAGCAACTATGTTTGGGGTATTAAGAATAGCTCCAAC  
CCAGCCGCTGCTGGAGCGTATCGATGTGCGGATGGCAGCCGTAAGAATTTAGCCTGGACGGTGTCAAGGATCCTCTGGGCCA  
GAACGTGGAGAAAGGTTTACGGCACGGCGTGGACTTACGGTAACGGTAGCCTGGGTTTTCCGGTAATAACCGACGGGTAGCG  
TGTACCAGATCGACATTTGGCAACGAGTCCGGTACCACATCGACGATCTGCATCTGAAGATTGTGGCGAAGCGTAAAGGCCCG  
ACCAGCAAAAACATGATGCAACGCTAATGGTATTTTGGACCCGGTATCAGGACTTAAAAGTTACCAAGAACTGGAGAA  
GATCGAGGTAACCAAGTTAGCTGGACGATTACCGTTGAGAACGCTGGCCCGTGGCCGAGCAGCGGTTTTACCTGCAAGATG  
TTGTCCAGAGGGCTACGCTGAAGTTAAGGGTATCCGCAGAGCAATGGTGGTATCAAGATATCAGCGTCAACGGTAACGTT

ATCAACGCCAGCCATGGCCCGATTGCTGTCAATGAGAAAGCGACGTACAAGGTTACCGCAAACCAGAGCATTAGCAATAATGA  
GAAGTGTCTGCAAAATACCGCGAGCATCTACGCCAATGAAAAAGATCTGATTGAGGACAATAACGTCGCGTCTGATGGTGCCT  
GTATTCGGCAATCACAAAACCGTTGTGGACCAGAATGGTGACGAGAAGATTGATGGTCAAGACGGCAGCGTGGCTGCGGGC  
AACGGCCTGCGTAAGGTGACCTACAAAATTGAAGTAAAAACCCGAAAGGTTTCCGGAAACTAAATACAGCTTGACGGACAC  
CCCGCAGTTCGCGGATAGCGTGAAGCTGGAGCGCTGAAAGTCATCAGCGACTATGGTAAGAAGAATCAAGAAGTGCAGGCAG  
CTGACATCTCTGTCAACGGCTTGGTGTGGCCGAAGATGGCACCCCGATCGAGGGCAACGCAAGCAAAGATAACCGTGCACACC  
TACATGGTCGAGGCGTTTTACAGCGGTCCGAAAAGCTCAGCACACGACTCGACGGACGAATGCAAGGACTCCACGCCGAAATT  
CGGTCTGTCAACAGCGCAAGCTGAAGTCCGTTGAGCGGAAAAGACCAGCGAGGGCTGCGCCCAATCGTTCGCAATGTGC  
CGATTAAGATTAGCTGAAAAAAGTTGACGCGGAGAACAAGGAAACCGAGCTGCAAGCAACCTTCGAGCTGTACCGTGTCTCT  
GATGGCGCAAGGTCGCGACCCTGCAACCTCGCCCAAACAGCAAGAGCAGACCGAAGTCGAACCGGGCAAGTATCGTTTGAT  
CGAAACCAGAGCCCGTTGGTTACCAGCTGCTGCCCGCTCTGTGGAGTTTGAAGTTACGCGTACGGCGGACGGTAAGGCCG  
ACGTGACCATCGACGCGGCATCCTTTCGCTGAGCAAGTCCGACAGCAGGGTAAGGACCCGAACTGGTTATTCTTACCCTC  
GCTGATGTTCCGTTGGTACTCTGCCGAAAACCGGTGGCCACGGTGTGGCGATCTATCTGGTTATGGGCGTGTCTGGTTCT  
GGTTGGTGTAGCTGGAGCTGTATCGTAATCAACTGATTAGCAGCTAAGGT

### Sequência otimizada – ORF sodc

ATGCTGCGTGTCTGCTGTTGCTCTGGTTGGTGTCTTCTTCGATCGCGGTTCTGTCCGCGTGTAGCAGCAGCACTACCACGAAGGA  
TAGCGCAGACAAAGCGATGACCAGCGCAGCGAGCACCGCGAGCAACAGCCACGCATTTCGCGACCGCGGAGCTGAAGAATCAGA  
GCGGCGAGGGTGTGGGCATGGTTCGAGTTTTCCAAAACGGCAGCCGGCAGCAAGATTACGGTGGACGCAAGGGGCTGACCCCT  
GGTTTTCAATGGCTTTTACCTGCACGACAAAGGTGTCTGCCAGGGTAACCTCAAAGCGCGGGTGGCCATCTGCATGGCAAAAA  
TCAACCACACGGTGTGGTTATCACCCGACGCGGGTGTCTGCCGAGCCTGTGGCTAATGCAGACGGCACCGCCCATCTGG  
AAGTGGTTACCGATGCGGTGATGAAGATCTGCTGTTCCGTTGCGGACGGCACCAGCCTGATTGTCCACGAGGGTCCGACAAAC  
TACGCGAATATCCCGCAGCGTTACGCCCGAAGCGTCCGGACGAAGAAACGTTGAAAACGGGTGACGCGGGTGCACGCTTCG  
GTGCGGTGTATACCAAGAAGTAAGGT

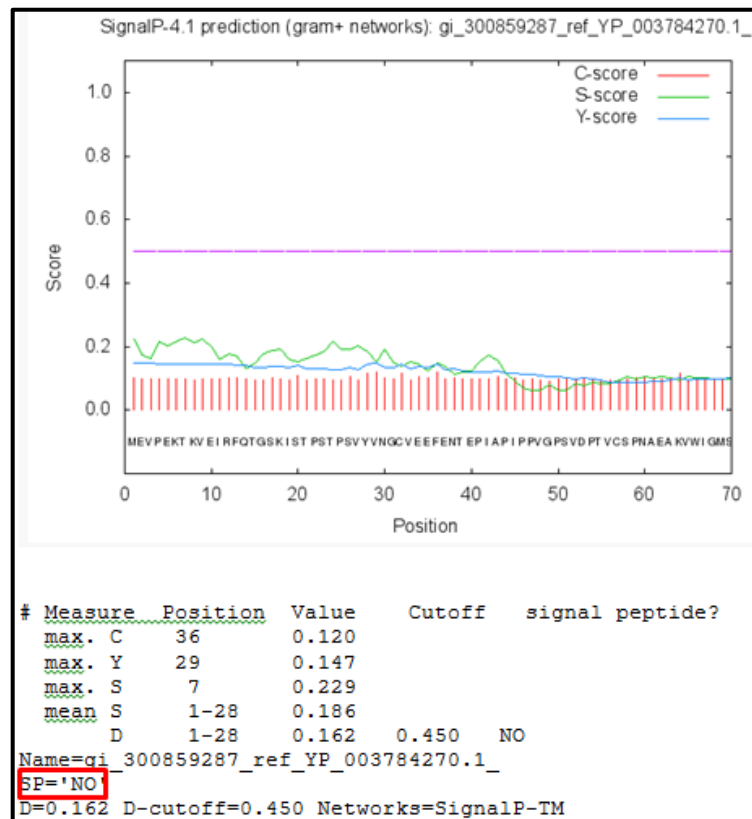
### Sequência otimizada - ORF nanh

ATGGCACTGGTCAACCTGACTGCAGCACTGGCACTGACGACCCAAAGGCTGACCGCGCCAGCAGCTCTCGCTGAAGAGGGCAG  
CACCATTGTCCAGCACCAGTGCACAACAATGGTCTGTTTGTATGCAGCCCCCGCGCACCGGTCGCGCGTGGCGCGGTGGGTG  
AGGGTAAACTGCCTGACCCGGTACCAGCGAGTTTTTCGATAGCAAGGTTATTCGTGACGTTGACCCGGCGGGTACGCGTTGT  
TTCCGTATCCCGCGATCGCGACCGCTGTTGACGGCACCCGCTGGTTCGCTGATAATCGCTATGGTCCGACGCTCAATC  
ACGCACCTGGTGCCTGATGCCCGTATGAGAATACCAAGCGTAATAGCCGTGAGAAAACAGACGGACATTGAGATCTACCGTT  
CCCACGACAGCGGTAAGAGCTTCGAGGAAGGCGACTTTATTGCTAAAGGTACCAGGATCCGCGGATTTGAGCTACACCGAC  
CCGAGCATCGTGGTTGATCAGAAAACGGGCAAGATTTTTGCGGTGTTTCGTGCGTGGTTACGATTACCGTTTTCTCGATGCGAC  
GGCGGGTGTAAACCGCGGCTCCGTTGAGGATCCGATCCAAAACGACGCGTGCAGGACACGGTATCATTGAGAGCAACGACA  
ATGGCCAGACGTGGGGTAACATGAAACTGCTGAGCGAGTTGACCGAGAAGATCAAGGTACGCCATAAGCCGGGTCAAGAGCTG  
GCGGGTCCGCGTCTGTTGCGACGCTGTTGATGGCATCCAGCTGAAGTATGGTGAGAATGCCGCTGCTGTTGTTCCGAT  
GAGCGTTCACTACAGCCCGAACGGTGCAGTCCGCAACGTCGCTCTGACTCCGACGATCACGGTGTACCTGGAACG  
TCGGTCAAGGTGTGGCGGTGAGGGCGAAACTAATGGCGACGAGAACAAGCTTGTGGAGCTGTCTGACGGCCGATATTATGATG  
AATAGCAAGAGCACGGCTCTAACCCGAGCAATCCGCTGGTGGCGAGCGTTGGCGTAGCTATAGCGAGGATCAGGGCGAAAC  
CTGGTCAAACCGGAGCGGTGATCGTTGCCCTCCGACGACCCGCTGCAGCGCAACACCGGTATTAACGTGACCTGATTA  
AGGCGTATCCGAGCGCGCCAGAGGGTAGCGCAGCCGACAGTGTATTGTTACTCGGCACCGATTGATACTCGTATCAAAGGC  
GAGTACGGTCAAGACGGCCGTAGCAATGGTTGGGTGATGGCGAGCTGTGACGATGGTTCGTAGCTGGACCCATGGCCGCCAGAT  
TGAGGCCGTCGTTTTCAATATAGCGTACGACCCGATGGCGGACGGCAACATTGGTATGGTTTACGAGAGCGGCGATAAAG  
AGCGTGGTATGACGCTGAAATTCGCCAAGTTAACATGGCGTGGCTGGGTGCGGACTGCTTGAACAACAAGCGCTGGGTATT  
ACCGATCTGCCGACCCAGCGCTGGCAAAGCTATCGAAGATGCCAAAGCGGCCACCGCAAAGCGGAAGAAGCGACGGCGAA  
TGTCTGAAAGTACGAAAAGAACTGGAAGAGGCAGAGCTGGAGAATGAAGAACTGGCCCAGGCTCTGGCAGAGGCGAAGTCTG  
CCGCGCAAGACGCTATCGCGGACGCCAAGAAGCGAATGCCGCGCTTAAAGCAGCCGAAGGTGCAACCATCGAAGCGGAAGAA  
GCGGCAATGAAGGCAGAGAATGAAGCAAAGCGCTGGCGGAGCAACTGGCCAAAGGTTGAAGCGGAACTGGCAAACAGCCAGGA  
CCAGGCTAAAGCTTTGGCTGAGGCCAAGAAGCGGCGGAGATTGCGCGCAAAGCAGCAGAAGAAGCCCTGAAACTGGAAAAAG  
AGAAGTCCGGCAAGGCTGGCGGTACGGATAACACCGAAAACAAGGGCTTCTGGCAAGAAGTGTGCGCATTTTTCCCTGGTTTT  
GCGCCGATCTTACGCTTCTGGCGTCCATCTGGGGCGGTATGCAGAAGTTGCTGGCGTTTTAAGGT

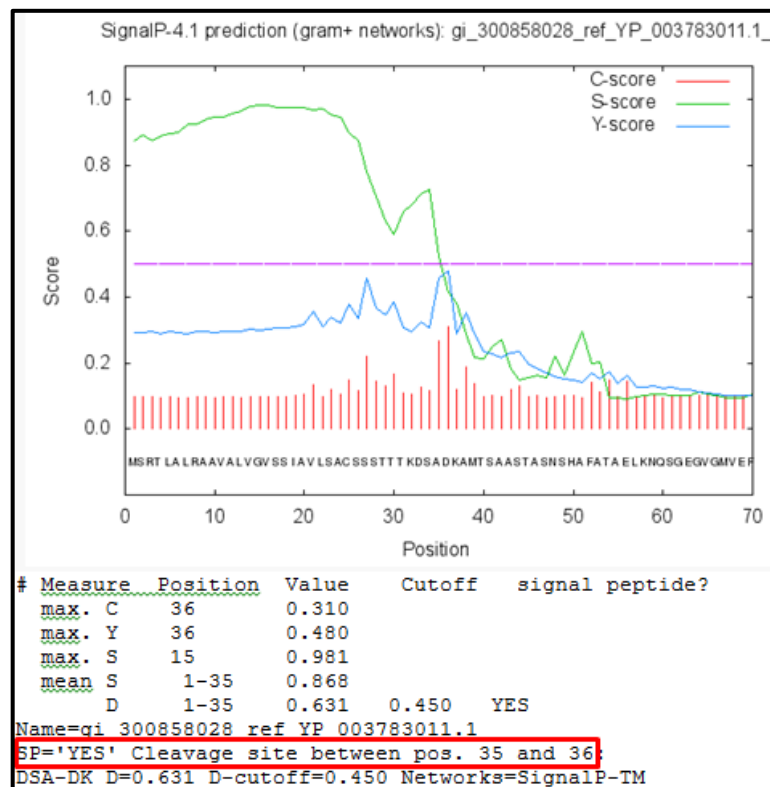


## **ANEXO 2**

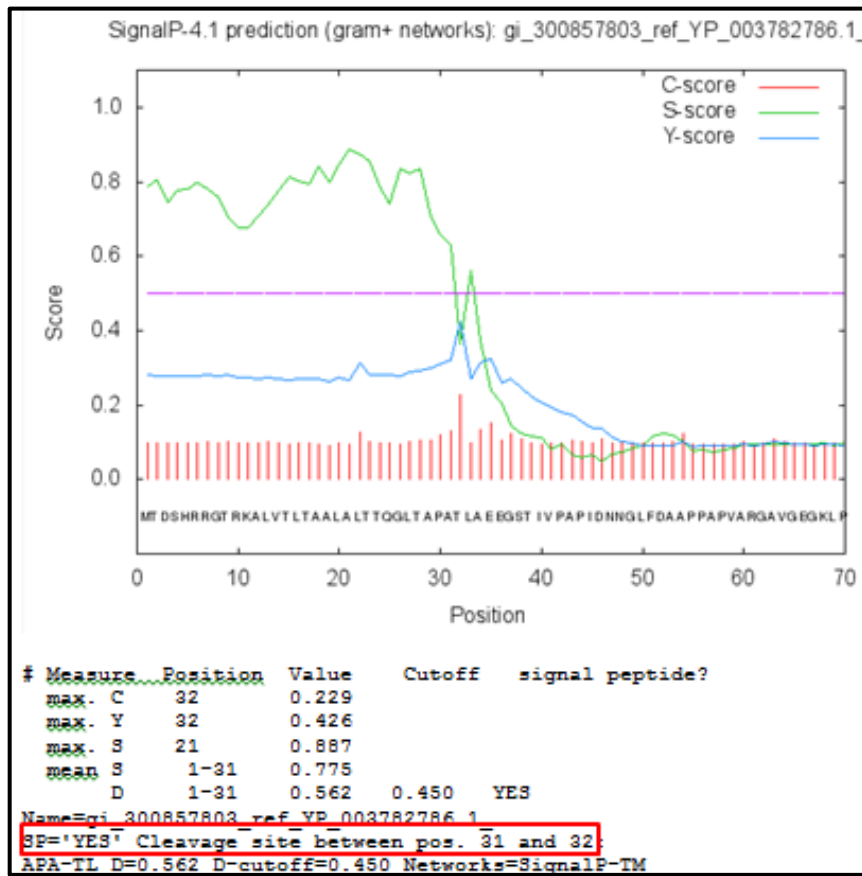
**Predição de peptídeo sinal das proteínas SpaC, SodC e NanH**



Predição de peptídeo sinal da proteína SpaC através do SignalP 4.0.



Predição de peptídeo sinal da proteína SodC através do SignalP 4.0.

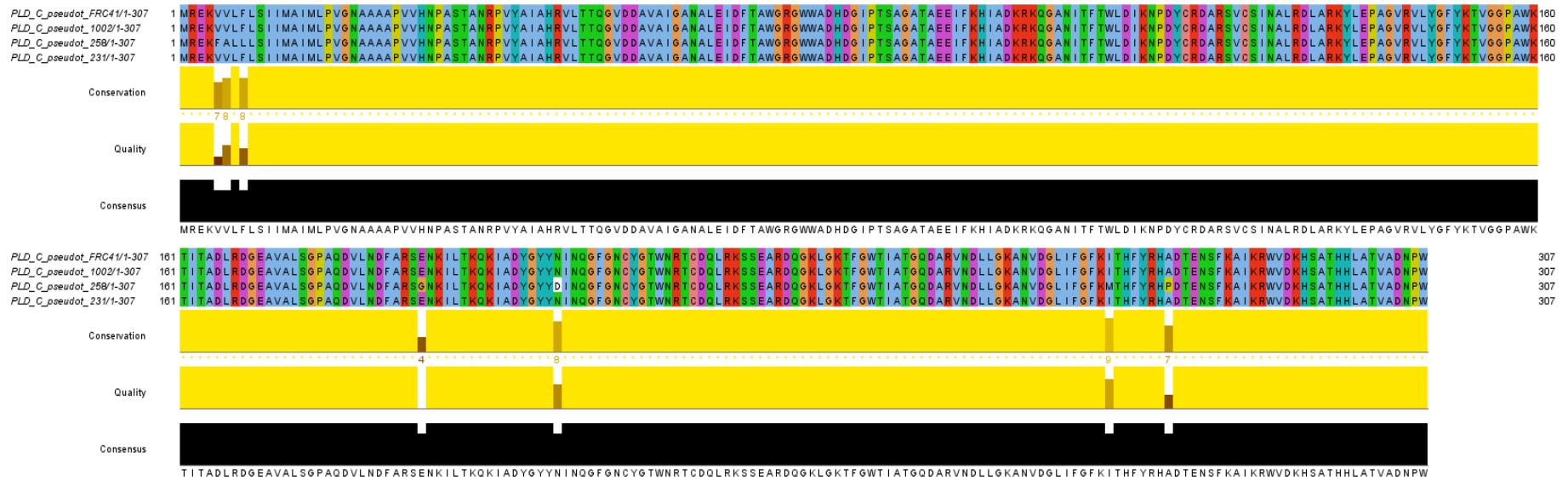


Predição de peptídeo sinal da proteína NanH através do SignalP 4.0.

## **ANEXO 3**

### **Alinhamentos de sequências**

## Alinhamento da proteína PLD de *C.pseudotuberculosis* FRC41 com sequências de microrganismos do grupo CMNR

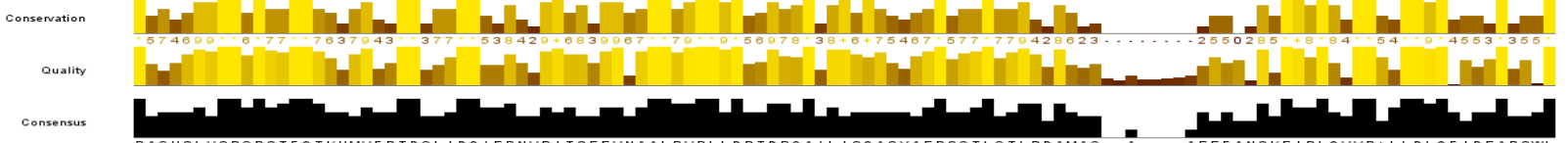




*PlnG\_C\_pseudot\_FRC41/1-749* 255 IDLGAVSGIGAFGYIYGTGKGFQAPVAVTDDGSIASDIYTGRTLAALTVDLPSADGVYLPGIPSPDEEKLFAQNLVSYRLLRRRMSDPAKRFASIRELETQLFGVLYREYLALAEHRQF 373  
*PlnG\_C\_pseudot\_1002/1-749* 255 IDLGAVSGIGAFGYIYGTGKGFQAPVAVTDDGSIASDIYTGRTLAALTVDLPSADGVYLPGIPSPDEEKLFAQNLVSYRLLRRRMSDPAKRFASIRELETQLFGVLYREYLALAEHRQF 373  
*PlnG\_C\_pseudot\_258/1-749* 255 IDLGAVSGIGAFGYIYGTGKGFQAPVAVTDDGSIASDIYTGRTLAALTVDLPSADGVYLPGIPSPDEEKLFAQNLVSYRLLRRRMSDPAKRFASIRELETQLFGVLYREYLALAEHRQF 373  
*PlnG\_C\_pseudot\_C231/1-749* 255 IDLGAVSGIGAFGYIYGTGKGFQAPVAVTDDGSIASDIYTGRTLAALTVDLPSADGVYLPGIPSPDEEKLFAQNLVSYRLLRRRMSDPAKRFASIRELETQLFGVLYREYLALAEHRQF 373  
*PlnG\_C\_dipht\_HC02/1-720* 231 IDLGAVSGIGAFGYIYGTGKGFQAPVAVTDDGSIASDIYTGRTLAALTLNDLNGAGAGKDKLPTADEEKLFAQNLVSYRLLRRRMSDPAKRFASIRELETQLFGVLYREYLALAEHRQF 348  
*PlnG\_C\_glut\_ATCC14067/1-822* 317 IDLGAVSGIGAFGYIYGTGKGFQAPVAVTDDGSIASDIYTGRTLAALMPLVVDGVLAPGIPSPKNSLRLRRHLSFYRLLRRRMSDPAKRFASIRELETQLFGVLYREYLALAEHRQF 435  
*PlnG\_M\_tuberc\_SUMu007/1-750* 202 IDLGAVSRINRFGYLYTGFQAPVAVTDDGSIASDIYTGRTLAALTLDPFRNGRVRDGLP...EDDVRKTYDYSYRLLRRRMSDPAKRFASIRELETQLFGVLYREYLALAEHRQF 408  
*PlnG\_Nocardia\_farc\_IFM10152/1-783* 311 IDLGAVSLAEAYNLYYGTGKGFQAPVAVTDDGSIASDIYTGRTLAALTLDPFRNGRVRDGLP...EDDVRKTYDYSYRLLRRRMSDPAKRFASIRELETQLFGVLYREYLALAEHRQF 429  
*PlnG\_Rhodococcus\_pyrri\_SB3094/1-769* 293 IDLGAVALIERYGYLYYGTGKGFQAPVAVTDDGSIASDIYTGRTLAALTLDPFRNGRVRDGLP...EDDVRKTYDYSYRLLRRRMSDPAKRFASIRELETQLFGVLYREYLALAEHRQF 411



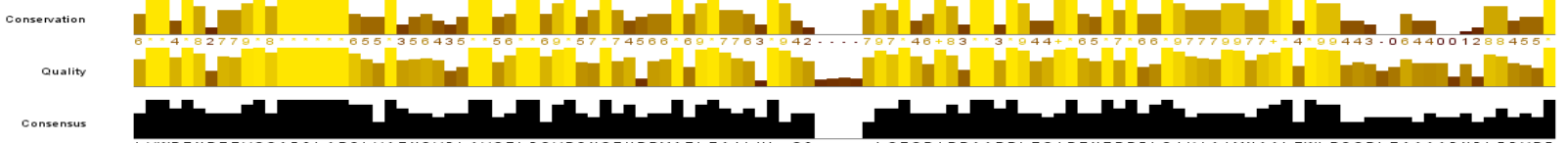
*PlnG\_C\_pseudot\_FRC41/1-749* 374 PAQHSLYSQPRSTFGTKHMVFRDQDLIDGIERNVRIITSEEVNAALPVPVLLDRDPPGAILISGASYAEPSTLQTLRDAMAQ...EAFANKEIPLGVVRLTLDLGFIDEARSWL 484  
*PlnG\_C\_pseudot\_1002/1-749* 374 PAQHSLYSQPRSTFGTKHMVFRDQDLIDGIERNVRIITSEEVNAALPVPVLLDRDPPGAILISGASYAEPSTLQTLRDAMAQ...EAFANKEIPLGVVRLTLDLGFIDEARSWL 484  
*PlnG\_C\_pseudot\_258/1-749* 374 PAQHSLYSQPRSTFGTKHMVFRDQDLIDGIERNVRIITSEEVNAALPVPVLLDRDPPGAILISGASYAEPSTLQTLRDAMAQ...EAFANKEIPLGVVRLTLDLGFIDEARSWL 484  
*PlnG\_C\_pseudot\_C231/1-749* 374 PAQHSLYSQPRSTFGTKHMVFRDQDLIDGIERNVRIITSEEVNAALPVPVLLDRDPPGAILISGASYAEPSTLQTLRDAMAQ...EAFANKEIPLGVVRLTLDLGFIDEARSWL 484  
*PlnG\_C\_dipht\_HC02/1-720* 349 PAQHSLYSQPRSTFGTKHMVFRDQDLIDGIERNVRIITSEEVNAALPVPVLLDRDPPGAILISGASYAEPSTLQTLRDAMAQ...EAFANKEIPLGVVRLTLDLGFIDEARSWL 459  
*PlnG\_C\_glut\_ATCC14067/1-822* 436 PAQHSLYSQPRSTFGTKHMVFRDQDLIDGIERNVRIITSEEVNAALPVPVLLDRDPPGAILISGASYAEPSTLQTLRDAMAQ...EAFANKEIPLGVVRLTLDLGFIDEARSWL 546  
*PlnG\_M\_tuberc\_SUMu007/1-750* 409 PGLSTIFSPSSTFGTKHMVFRDQDLIDGIERNVRIITSEEVNAALPVPVLLDRDPPGAILISGASYAEPSTLQTLRDAMAQ...EAFANKEIPLGVVRLTLDLGFIDEARSWL 525  
*PlnG\_Nocardia\_farc\_IFM10152/1-783* 430 PQLSTIFSPSSTFGTKHMVFRDQDLIDGIERNVRIITSEEVNAALPVPVLLDRDPPGAILISGASYAEPSTLQTLRDAMAQ...EAFANKEIPLGVVRLTLDLGFIDEARSWL 547  
*PlnG\_Rhodococcus\_pyrri\_SB3094/1-769* 412 PGLSTIFSPSSTFGTKHMVFRDQDLIDGIERNVRIITSEEVNAALPVPVLLDRDPPGAILISGASYAEPSTLQTLRDAMAQ...EAFANKEIPLGVVRLTLDLGFIDEARSWL 530

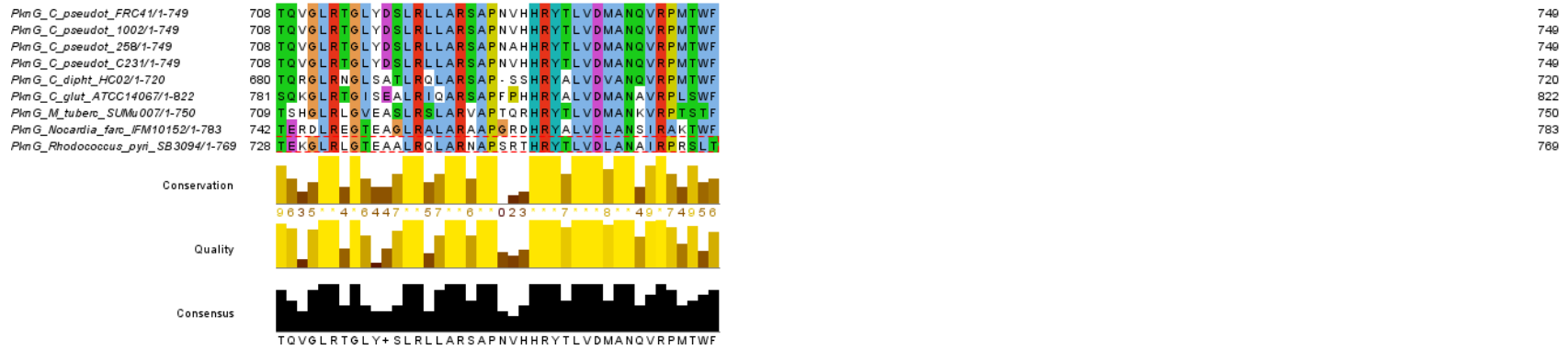


*PlnG\_C\_pseudot\_FRC41/1-749* 485 DTLDATLSDDWRHQWYSGVTSFLFDDYVMAQRHFNEVYNIIPGEAAPKLARAACVEMLLQGMGMDSTPLLSPDTAVSAEELKQ...NTIG-MWSELTKDPETLRFKALYLYA 592  
*PlnG\_C\_pseudot\_1002/1-749* 485 DTLDATLSDDWRHQWYSGVTSFLFDDYVMAQRHFNEVYNIIPGEAAPKLARAACVEMLLQGMGMDSTPLLSPDTAVSAEELKQ...NTIG-MWSELTKDPETLRFKALYLYA 592  
*PlnG\_C\_pseudot\_258/1-749* 485 DTLDATLSDDWRHQWYSGVTSFLFDDYVMAQRHFNEVYNIIPGEAAPKLARAACVEMLLQGMGMDSTPLLSPDTAVSAEELKQ...NTIG-MWSELTKDPETLRFKALYLYA 592  
*PlnG\_C\_pseudot\_C231/1-749* 485 DTLDATLSDDWRHQWYSGVTSFLFDDYVMAQRHFNEVYNIIPGEAAPKLARAACVEMLLQGMGMDSTPLLSPDTAVSAEELKQ...NTIG-MWSELTKDPETLRFKALYLYA 592  
*PlnG\_C\_dipht\_HC02/1-720* 460 DLVLRIGNEWRHQWYSGVTSFLFDDYVMAQRHFNEVYNIIPGEAAPKLARAACVEMLLQGMGMDSTPLLSPDTAVSAEELKQ...NTIG-MWSELTKDPETLRFKALYLYA 568  
*PlnG\_C\_glut\_ATCC14067/1-822* 547 ELLGRIGDDWRHQWYSGVTSFLFDDYVMAQRHFNEVYNIIPGEAAPKLARAACVEMLLQGMGMDSTPLLSPDTAVSAEELKQ...NTIG-MWSELTKDPETLRFKALYLYA 605  
*PlnG\_M\_tuberc\_SUMu007/1-750* 526 DDLAERVGRWRWLVWYRAVAELTLDYDYSATKHFTEVLDLTPGELAPKLALAAATAELAGN...TDEH...KIFYG 593  
*PlnG\_Nocardia\_farc\_IFM10152/1-783* 548 ARLDAG...DWRVHWYTGLAELREQHYESAFAAFEDVLRVLPGEIAPKLALAAATAEVLDL...HWEADPEQWRAYAEKYE 623  
*PlnG\_Rhodococcus\_pyrri\_SB3094/1-769* 531 TRLERE...DDWRHQWYSGVTSFLFDDYVMAQRHFNEVYNIIPGEAAPKLARAACVEMLLQGMGMDSTPLLSPDTAVSAEELKQ...NTIG-MWSELTKDPETLRFKALYLYA 609



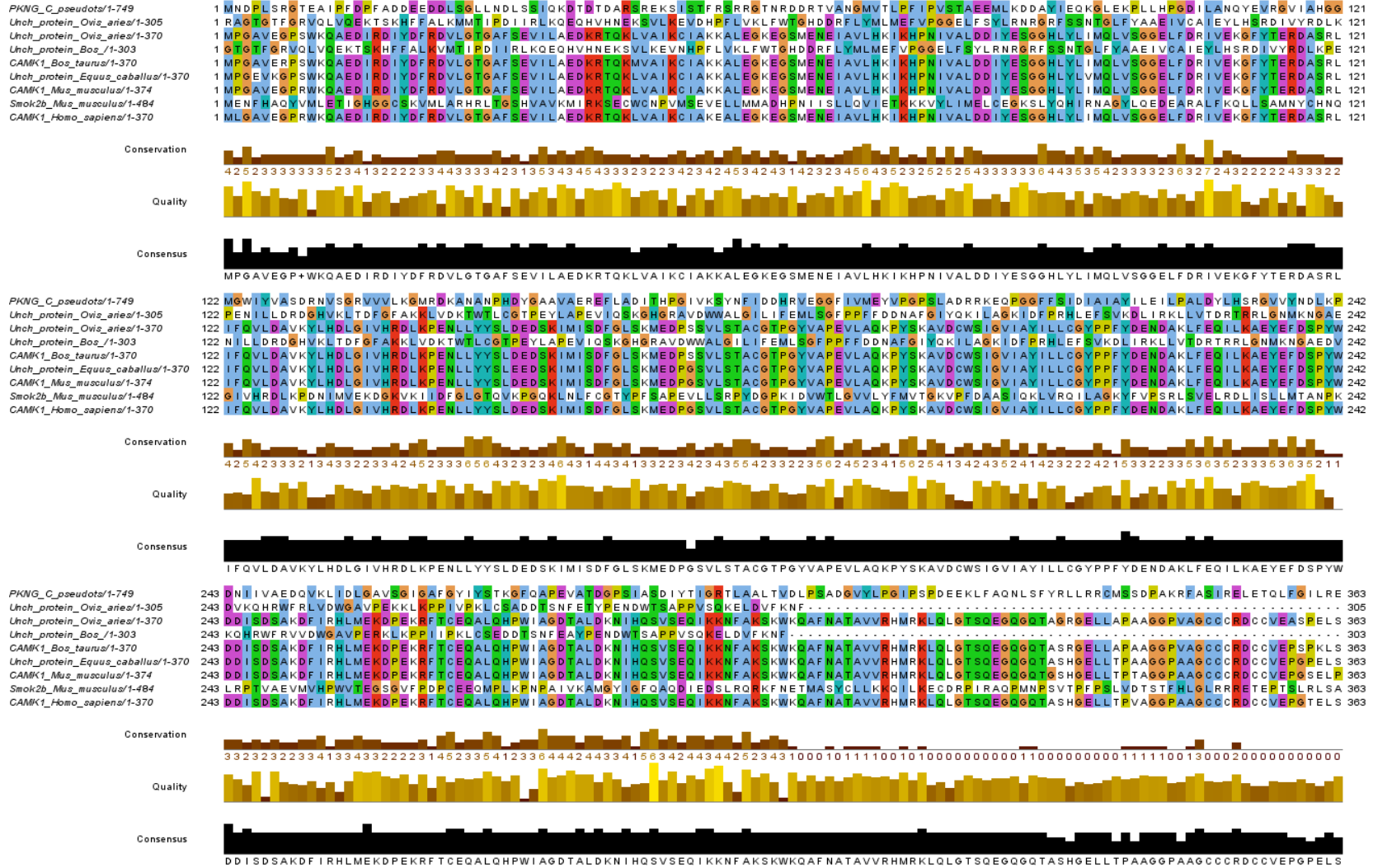
*PlnG\_C\_pseudot\_FRC41/1-749* 593 LVWRNTPPTVSSAFGLARQLVAENQVDLAVSTLDQVPQNSHRRMAELTALHLISG...LSESRIRRAARLESIPITNEPRFLQIKIAIMNAALTWLRQSRLEAAADNDLFYRF 707  
*PlnG\_C\_pseudot\_1002/1-749* 593 LVWRNTPPTVSSAFGLARQLVAENQVDLAVSTLDQVPQNSHRRMAELTALHLISG...LSESRIRRAARLESIPITNEPRFLQIKIAIMNAALTWLRQSRLEAAADNDLFYRF 707  
*PlnG\_C\_pseudot\_258/1-749* 593 LVWRNTPPTVSSAFGLARQLVAENQVDLAVSTLDQVPQNSHRRMAELTALHLISG...LSESRIRRAARLESIPITNEPRFLQIKIAIMNAALTWLRQSRLEAAADNDLFYRF 707  
*PlnG\_C\_pseudot\_C231/1-749* 593 LVWRNTPPTVSSAFGLARQLVAENQVDLAVSTLDQVPQNSHRRMAELTALHLISG...LSESRIRRAARLESIPITNEPRFLQIKIAIMNAALTWLRQSRLEAAADNDLFYRF 707  
*PlnG\_C\_dipht\_HC02/1-720* 569 LVWRNTPPTVSSAFGLARQLVAENQVDLAVSTLDQVPQNSHRRMAELTALHLISG...LSESRIRRAARLESIPITNEPRFLQIKIAIMNAALTWLRQSRLEAAADNDLFYRF 679  
*PlnG\_C\_glut\_ATCC14067/1-822* 666 LVWATNPTTVSSAFGLARQLVAENQVDLAVSTLDQVPQNSHRRMAELTALHLISG...LSESRIRRAARLESIPITNEPRFLQIKIAIMNAALTWLRQSRLEAAADNDLFYRF 780  
*PlnG\_M\_tuberc\_SUMu007/1-750* 594 TVWSTDQVIVSAFGLARQLVAENQVDLAVSTLDQVPQNSHRRMAELTALHLISG...LSESRIRRAARLESIPITNEPRFLQIKIAIMNAALTWLRQSRLEAAADNDLFYRF 708  
*PlnG\_Nocardia\_farc\_IFM10152/1-783* 624 TVWSTDQVIVSAFGLARQLVAENQVDLAVSTLDQVPQNSHRRMAELTALHLISG...LSESRIRRAARLESIPITNEPRFLQIKIAIMNAALTWLRQSRLEAAADNDLFYRF 741  
*PlnG\_Rhodococcus\_pyrri\_SB3094/1-769* 610 VVWRIDAVVSAFGLARQLVAENQVDLAVSTLDQVPQNSHRRMAELTALHLISG...LSESRIRRAARLESIPITNEPRFLQIKIAIMNAALTWLRQSRLEAAADNDLFYRF 727







# Alinhamento da proteína PknG de *C.pseudotuberculosis* FRC41 com sequências de organismos eucariotos



*PKNG\_C\_pseudots/1-749* 364 Y L A L A E H R Q F P A Q H S L Y S P Q R S T F G T K H M V F R T D Q L I D G I E R N V R I T S E E V N A A L P V R L L D R T D P G A I L I S G A S Y A E P S Q T L Q T L R D A M A Q E E F A N S K E I P L G V V R T L L D L G F I D E A R S W L 484  
*Unch\_protein\_Ovis\_aries/1-305* .....  
*Unch\_protein\_Ovis\_aries/1-370* 364 P S L H P Q L ..... 370  
*Unch\_protein\_Bos/1-303* .....  
*CAMK1\_Bos\_taurus/1-370* 364 P S L H P Q L ..... 370  
*Unch\_protein\_Equus\_caballus/1-370* 364 P S L P S Q L ..... 370  
*CAMK1\_Mus\_musculus/1-374* 364 P A P P S G R A M D ..... 374  
*Smok2b\_Mus\_musculus/1-484* 364 N R Q M S V C G R S T S K K R D R F S W E S V S G R P L H T T T H M D H T H T R T R S V P C I Y S M F C T I Q E N S S D D S T E G H T S A S A E D K P V R S R G W E R G I K G W T R K I G N V M R K L C C C I P S K E T S H V G H S R V S P K K 484  
*CAMK1\_Homo\_sapiens/1-370* 364 P T L P H G L ..... 370

Conservation



Quality



Consensus

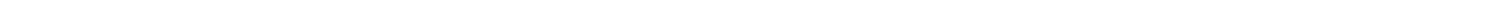
P S L P P Q L R ..... D ..... G ..... S ..... S .....  
 P S L P P Q L R .....

*PKNG\_C\_pseudots/1-749* 485 D T L D A T L S D D W R H Q W Y S G V T S L F L D D Y V M A G R H F N E V Y N I L P G E A A P K L A R A A V C E M L L G G M G M D S T P L L S P D T A V S A A E L K G N T I G M W S E L T K D E T L R F K A L Y L Y A L V W R T N P T T V S S A 605  
*Unch\_protein\_Ovis\_aries/1-305* .....  
*Unch\_protein\_Ovis\_aries/1-370* .....  
*Unch\_protein\_Bos/1-303* .....  
*CAMK1\_Bos\_taurus/1-370* .....  
*Unch\_protein\_Equus\_caballus/1-370* .....  
*CAMK1\_Mus\_musculus/1-374* .....  
*Smok2b\_Mus\_musculus/1-484* .....  
*CAMK1\_Homo\_sapiens/1-370* .....

Conservation



Quality

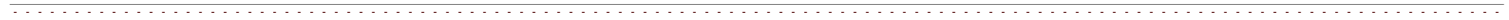


Consensus

.....

*PKNG\_C\_pseudots/1-749* 606 F G L A R Q L V A E N Q V D L A V S T L D Q V P Q N S T H R R M A E L T A I L H L I S G D L S E S R I R R A A R R L E S I P T N E R F L Q I K I A I M N A A L T W L R Q S R L E A A A A D N D L F D Y R F T Q V G L R T G L Y D S L R L L A R S 726  
*Unch\_protein\_Ovis\_aries/1-305* .....  
*Unch\_protein\_Ovis\_aries/1-370* .....  
*Unch\_protein\_Bos/1-303* .....  
*CAMK1\_Bos\_taurus/1-370* .....  
*Unch\_protein\_Equus\_caballus/1-370* .....  
*CAMK1\_Mus\_musculus/1-374* .....  
*Smok2b\_Mus\_musculus/1-484* .....  
*CAMK1\_Homo\_sapiens/1-370* .....

Conservation



Quality



Consensus

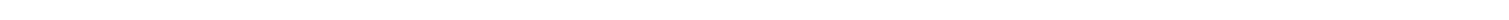
.....

*PKNG\_C\_pseudots/1-749* 727 A P N V H H R Y T L V D M A N Q V R P M T W F 749  
*Unch\_protein\_Ovis\_aries/1-305* .....  
*Unch\_protein\_Ovis\_aries/1-370* .....  
*Unch\_protein\_Bos/1-303* .....  
*CAMK1\_Bos\_taurus/1-370* .....  
*Unch\_protein\_Equus\_caballus/1-370* .....  
*CAMK1\_Mus\_musculus/1-374* .....  
*Smok2b\_Mus\_musculus/1-484* .....  
*CAMK1\_Homo\_sapiens/1-370* .....

Conservation



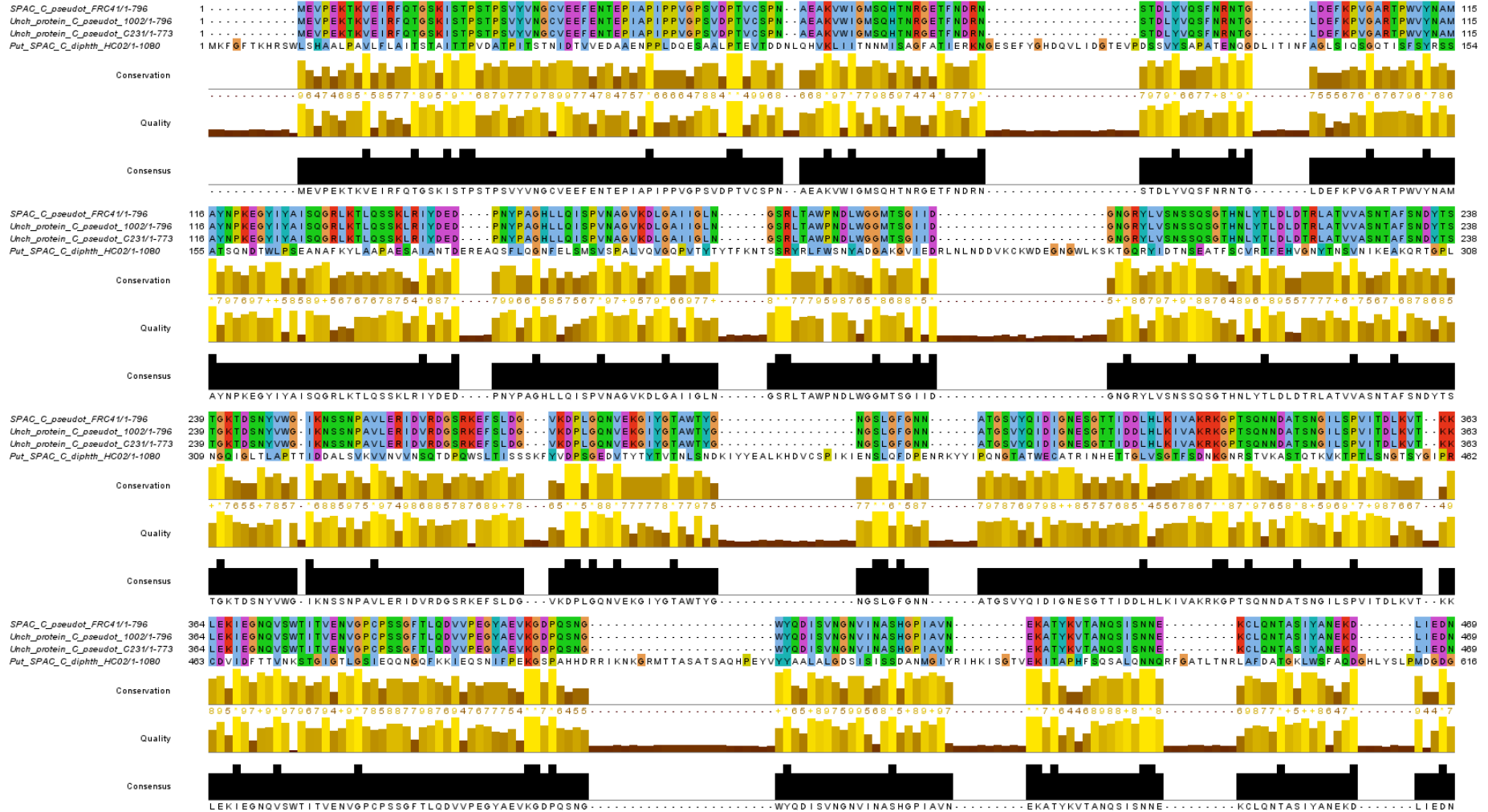
Quality

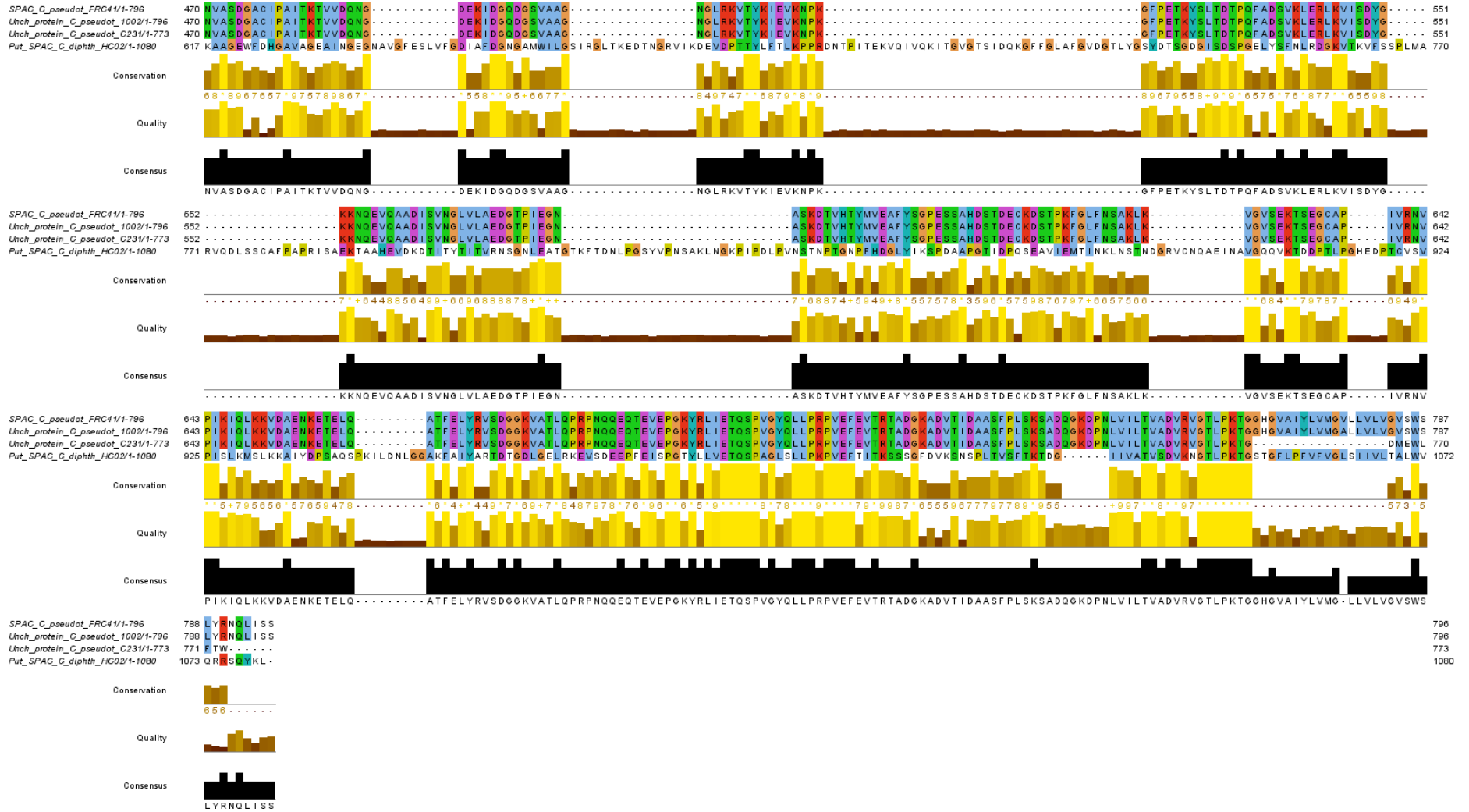


Consensus

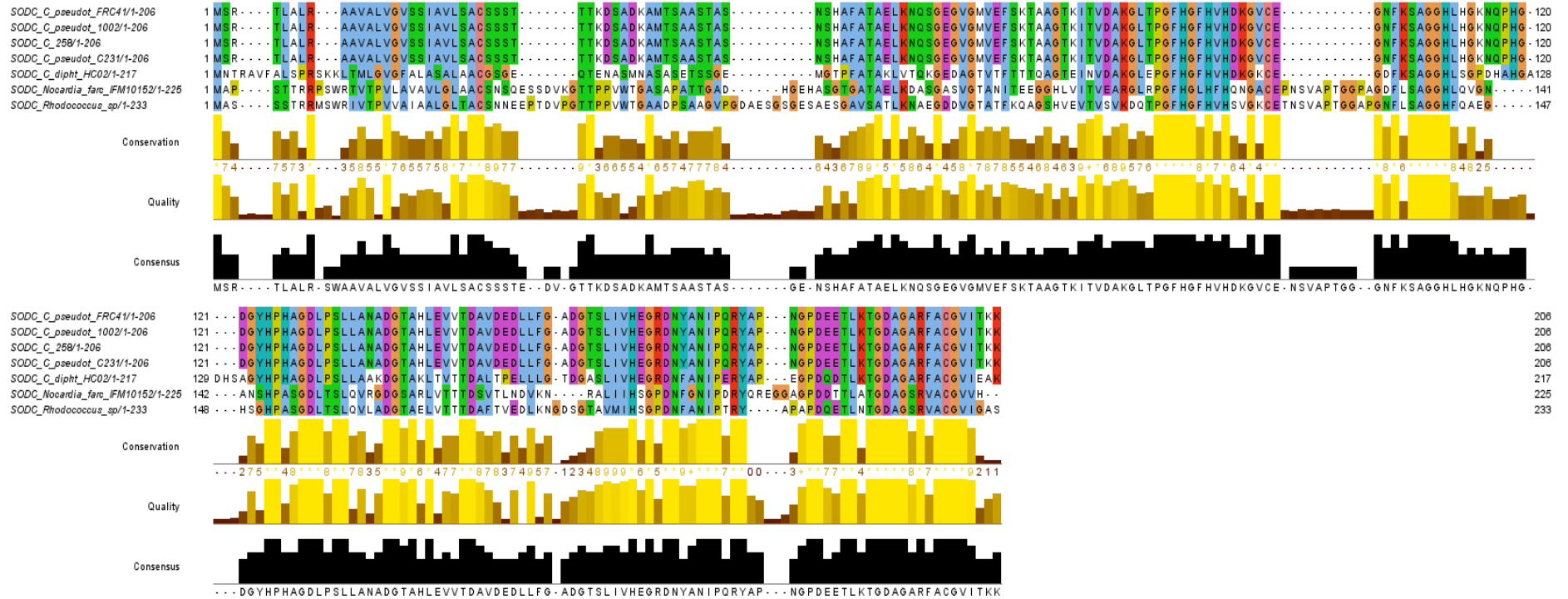
.....

# Alinhamento da proteína SpaC de *C. pseudotuberculosis* FRC41 com sequências de microrganismos do grupo CMNR



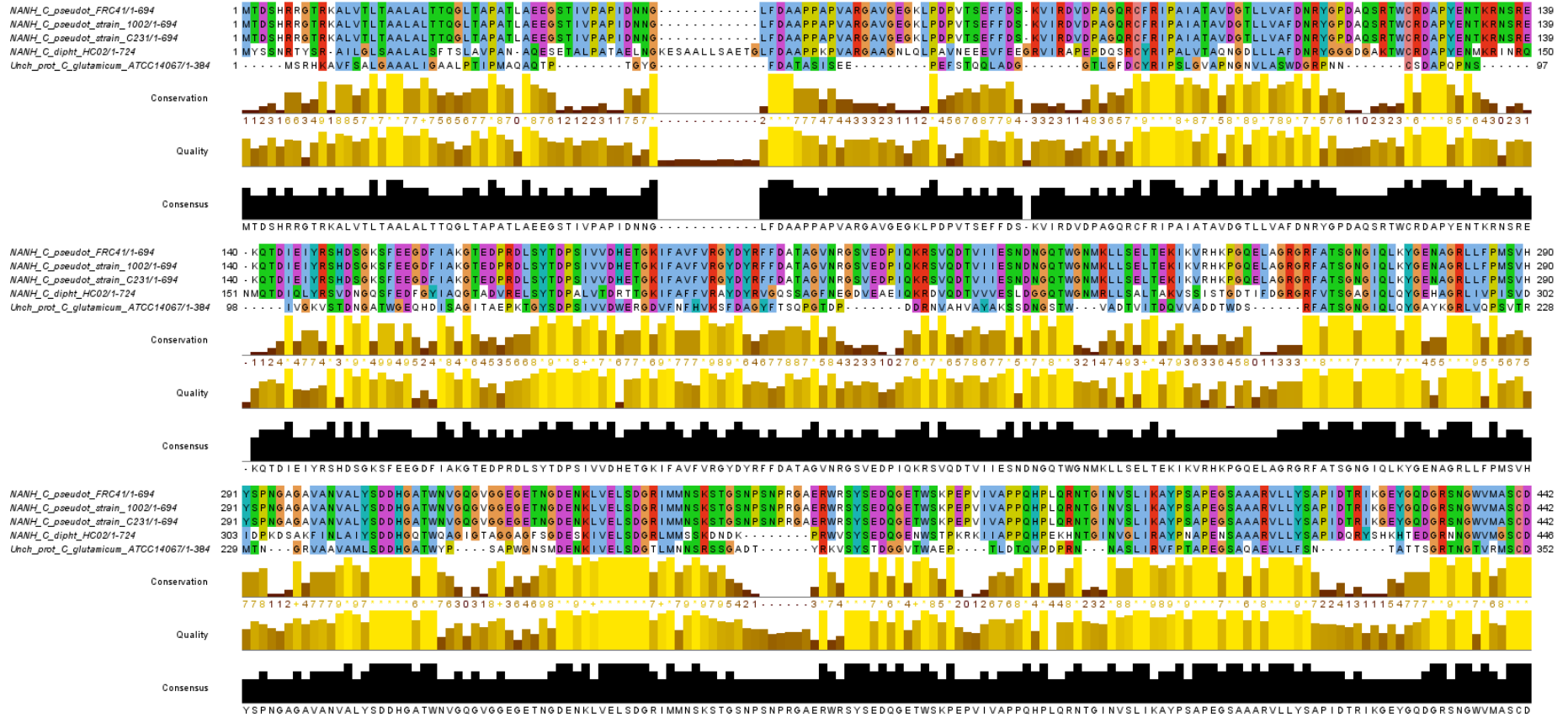


## Alinhamento da proteína SodC de *C.pseudotuberculosis* FRC41 com sequências de microrganismos do grupo CMNR

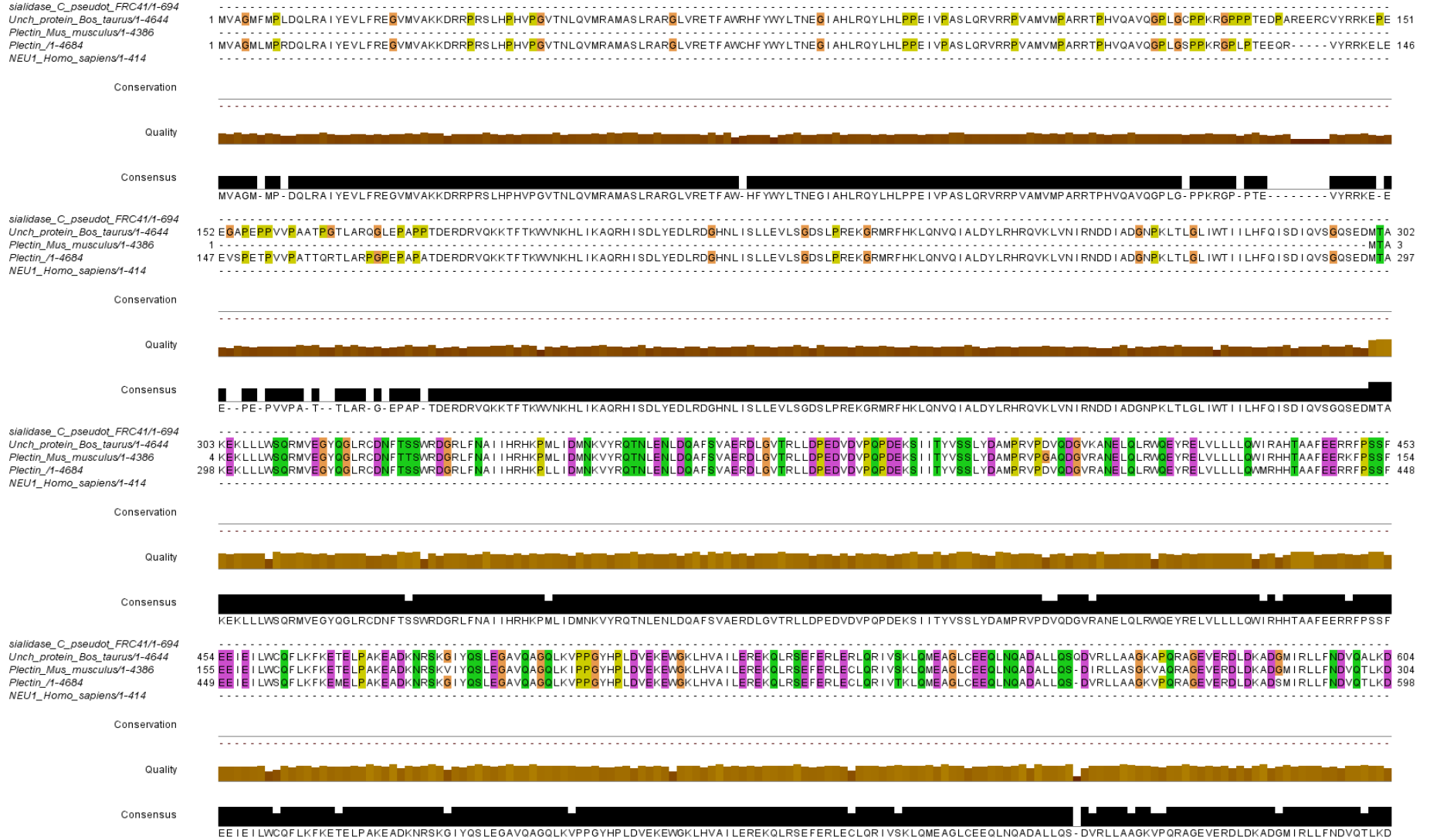




# Alinhamento da proteína NanH de *C.pseudotuberculosis* FRC41 com sequências de microrganismos do grupo CMNR



# Alinhamento da proteína NanH de *C.pseudotuberculosis* FRC41 com seqüências de organismos eucariotos





*sialidase\_C\_pseudot\_FRC411/1-694*  
*Unch\_protein\_Bos\_taurus/1-4644*  
*Plectin\_Mus\_musculus/1-4386*  
*Plectin\_1-4684*  
*NEU1\_Homo\_sapiens/1-414*

605 GRHPQGEQMYRRVYRLHERLVAIRTEYNLRLRQ-----TRRHPLEDSTLRYLQDLLAWVEENQRRLDGAEWGVLDLPSVEALGSHRGLHQSVEEFRAKIERARDEGQLSPATRGAYRDCGLGRLDLQYAKLLNSSKARL 739  
 305 GRHPQGEQMYRRVYRLHERLVAIRTEYNLRLKAGVQAP--VTQVTLQSTQRRPELEDSTLRYLQDLLAWVEENQRRIDSAEWGVLDLPSVEALGSHRGMHQSIEEFRAKIERARDEGQLSPATRGAYRDCGLGRLDLQYAKLLNSSKARL 452  
 599 GRHPQGEQMYRRVYRLHERLVAIRTEYNLRLKAGVAAATQVAQVTLQSVORRPELEDSTLRYLQDLLAWVEENQHRVDGAEWGVLDLPSVEALGSHRGLHQSIEEFRAKIERARDEGQLSPATRGAYRDCGLGRLDLQYAKLLNSSKARL 749

Conservation

Quality

Consensus

GRHPQGEQMYRRVYRLHERLVAIRTEYNLRLKAGV-AP---V-QVTLQSTQRRPELEDSTLRYLQDLLAWVEENQRR-DGAEWGVLDLPSVEALGSHRGLHQSIEEFRAKIERAR-DEGQLSPATRGAYRDCGLGRLDLQYAKLLNSSKARL

*sialidase\_C\_pseudot\_FRC411/1-694*  
*Unch\_protein\_Bos\_taurus/1-4644*  
*Plectin\_Mus\_musculus/1-4386*  
*Plectin\_1-4684*  
*NEU1\_Homo\_sapiens/1-414*

740 RRLSLHGFVAAATKELMWLSEKKEEEVGFQWSDRNTNMAAKKESYSALMRELELKEKKIKEIQNTGDRLLREDHPARPTVESFQAALQQTQWQWMLQLCCCEIAHLKENTAYFFQFSDVREAEELQKLETLRRKYSCDRITVTRLEDL 890  
 453 RRLSLHGFVAAATKELMWLSEKKEEEVGFQWSDRNTNMAAKKESYSALMRELELKEKKIKEIQNTGDRLLREDHPARPTVESFQAALQQTQWQWMLQLCCCEIAHLKENTAYFFQFSDVREAEELQKLETLRRKYSCDRITVTRLEDL 653  
 750 RRLSLHGFVAAATKELMWLSEKKEEEVGFQWSDRNTNMAAKKESYSALMRELELKEKKIKEIQNTGDRLLREDHPARPTVESFQAALQQTQWQWMLQLCCCEIAHLKENTAYFFQFSDVREAEELQKLETLRRKYSCDRITVTRLEDL 900

Conservation

Quality

Consensus

RRLSLHGFVAAATKELMWLSEKKEEEVGFQWSDRNTNMAAKKESYSALMRELELKEKKIKEIQNTGDRLLREDHPARPTVESFQAALQQTQWQWMLQLCCCEIAHLKENTAYFFQFSDVREAEELQKLETLRRKYSCDRITVTRLEDL

*sialidase\_C\_pseudot\_FRC411/1-694*  
*Unch\_protein\_Bos\_taurus/1-4644*  
*Plectin\_Mus\_musculus/1-4386*  
*Plectin\_1-4684*  
*NEU1\_Homo\_sapiens/1-414*

891 LQDAQDEKEQLNEYKGHLSGLAKRAKAIIVLQKPRNP AHP-RGRVPLLA VCDYKQVEVTVHKGDCEQLVGP AQP SHMKVLSGSGSEAAVPSVCFVLP PPNOEAQAEAVARLEA QHQAALVTLWHQLHVMKSL LAWQSLSRD VQLIRSWLVTF 1041  
 604 LQDAQDEKEQLNEYKGHLSGLAKRAKAIIVLQKPRNP AHPVRGHVPLI AVCDYKQVEVTVHKGDCEQLVGP AQP SHMKVLSGSGSEAAVPSVCFVLP PPNOEAQAEAVARLEA QHQAALVTLWHQLHVMKSL LAWQSLSRD VQLIRSWLVTF 754  
 901 LQDAQDEKEQLNEYKGHLSGLAKRAKAVVLLKPRHP AHPMRGRPLLA VCDYKQVEVTVHKGDCEQLVGP AQP SHMKVLSGSGSEAAVPSVCFVLP PPNOEAQAEAVARLEA QHQAALVTLWHQLHVMKSL LAWQSLSRD VQLIRSWLVTF 1051

Conservation

Quality

Consensus

LQDAQDEKEQLNEYKGHLSGLAKRAKAIIVLQKPRNP AHP-RGRVPLLA VCDYKQVEVTVHKGDCEQLVGP AQP SHMKVLSGSGSEAAVPSVCFVLP PPNOEAQAEAVARLEA QHQAALVTLWHQLHVMKSL LAWQSLSRD VQLIRSWLVTF

*sialidase\_C\_pseudot\_FRC411/1-694*  
*Unch\_protein\_Bos\_taurus/1-4644*  
*Plectin\_Mus\_musculus/1-4386*  
*Plectin\_1-4684*  
*NEU1\_Homo\_sapiens/1-414*

1-----MTDSHRRG-----TKALVLTAAALTTQGLTAPATLAEEGSTIVPAPIDNNGLFDAAPPARVARG 62  
 1042 RLLKPEEQRALRSLLELHYQAF LRDSQDAGGF GPEDRLQAEREYGC SRHYQQLLSLEGG---KCGQCQ-----XLDKPEARECAQR IAEQKAAAEVEGLGKGVARLSAEAEKVLA LPEPSPAAPTLRS 1164  
 755 RLLKPEEQRALRNLLELHYQAF LRDSQDAGGF GPEDRLVAEREYGC SRHYQQLLSLEGGQEESRCQRCISELKD IRLQLEACETRTRVHRLRLP LDKPEARECAQR IAEQKAAAEVEGLGKGVARLSAEAEKVLA LPEPSPAAPTLRS 905  
 1052 RLLKPEEQRALHSLLELHYQAF LRDSQDAGGF GPEDRLMAEREYGC SHHYQQLLSLEGGQEESRCQRCISELKD IRLQLEACETRTRVHRLRLP LDKPEARECAQR IAEQKAAAEVEGLGKGVARLSAEAEKVLA LPEPSPAAPTLRS 1202  
 1-----MTGERPSTALPDRRWGPR I LGFWGGCRWVFAAIFLLLSLAAWSKAEENDFG 52

Conservation

Quality

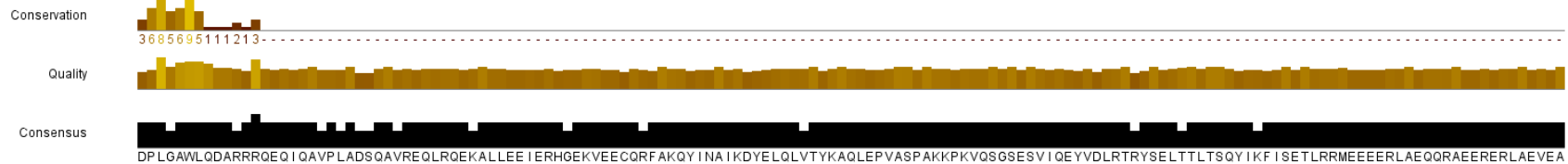
Consensus

RLLKPEEQRALRSLLELHYQAF LRDSQDAGGF GPEDRL- AEREYGC SRHYQQLLSLEGG- QEESRCQRCISELKD IRLQLEACETRTRVHRLRLP LDKPEARECAQR IAEQKAAAEVEGLGKGVARLSAEAEKVLA LPEPSPAAPTLRS

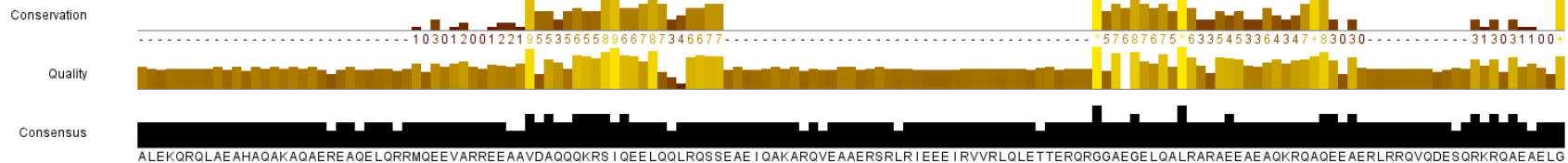
*sialidase\_C\_pseudot\_FRC41/1-694* 63 AVGE--GKLPDP-----VTEFFDskVIKDVDPAGQRFRIPAIATAVDGTLVAFDNRVYGPDAQS-----RTWCRDAPYENTKRNSEKTDIETIRSHDSGKSFE--GDFIAKGTEDPRDLSYDPSIVVDHETG 186  
*Unch\_protein\_Bos\_taurus/1-4644* 1165 ELETLGKLEQVRSLSAIYLEKTKTISLVIRSTGGAEVLAHEEQLKEAQAQVPAALPELEATKAAMKLLRAQAEAOQPVFDALRDELRGAEVGERLQQRHGERDVEVERWRERVLTQLLERWQAVLAQTQDVRRELEQLGRQLRYRESA 1315  
*Plectin\_Mus\_musculus/1-4386* 906 ELETLGKLEQVRSLSAIYLEKTKTISLVIRSTGGAEVLAHEEQLKEAQAQVPAALPELEATKAAMKLLRAQAEAOQPVFNTLRDELRGAEVGERLQQRHGERDVEVERWRERVLTQLLERWQAVLAQTQDVRRELEQLGRQLRYRESA 1056  
*Plectin/1-4684* 1203 ELETLGKLEQVRSLSAIYLEKTKTISLVIRSTGGAEVLAHEEQLKEAQAQVPAALPELEATKAAMKLLRAQAEAOQPTFDALRDELRGAEVGERLQQRHGERDVEVERWRERVLTQLLERWQAVLAQTQDVRRELEQLGRQLRYRESA 1353  
*NEU1\_Homo\_sapiens/1-414* 53 LVQP-----LVIMEQLLWVSGRIQSGDITFRIPAIATRGRILLAFEAEL-----KMSSSDEGAKFIALLRSMDOGSVWSP--TAFIVNDQDVPDGLNLS--AVVSDVETG 149



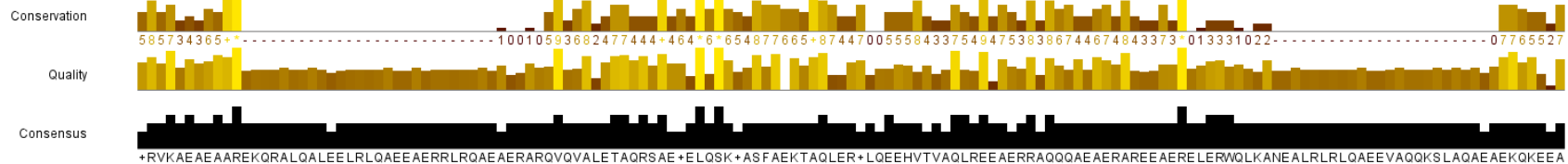
*sialidase\_C\_pseudot\_FRC41/1-694* 187 KIFAVFVRGVDYF-----DPLGAWLQDARRRQEQIQAVPLADSQAVREQLRQEKALLEEIERHGEKVEEQRFAKQYINAIKDYELQLVITYKAQLEPVASPAKKPKVDSGSESVIQEYVDLRTRYSELTTLTQYIKFISETLRRMEEEERLAEQORAEERERLAEVEA 199  
*Unch\_protein\_Bos\_taurus/1-4644* 1316 DPLGAWLQDARRRQEQIQAVPLADSQAVREQLRQEKALLEEIERHGEKVEEQRFAKQYINAIKDYELQLVITYKAQLEPVASPAKKPKVDSGSESVIQEYVDLRTRYSELTTLTQYIKFISETLRRMEEEERLAEQORAEERERLAEVEA 1466  
*Plectin\_Mus\_musculus/1-4386* 1207 DPLGAWLQDARRRQEQIQAVPLADSQAVREQLRQEKALLEEIERHGEKVEEQRFAKQYINAIKDYELQLVITYKAQLEPVASPAKKPKVDSGSESVIQEYVDLRTRYSELTTLTQYIKFISETLRRMEEEERLAEQORAEERERLAEVEA 1207  
*Plectin/1-4684* 1354 DPLGAWLQDARRRQEQIQAVPLADSQAVREQLRQEKALLEEIERHGEKVEEQRFAKQYINAIKDYELQLVITYKAQLEPVASPAKKPKVDSGSESVIQEYVDLRTRYSELTTLTQYIKFISETLRRMEEEERLAEQORAEERERLAEVEA 1504  
*NEU1\_Homo\_sapiens/1-414* 150 VVFLFYF-----

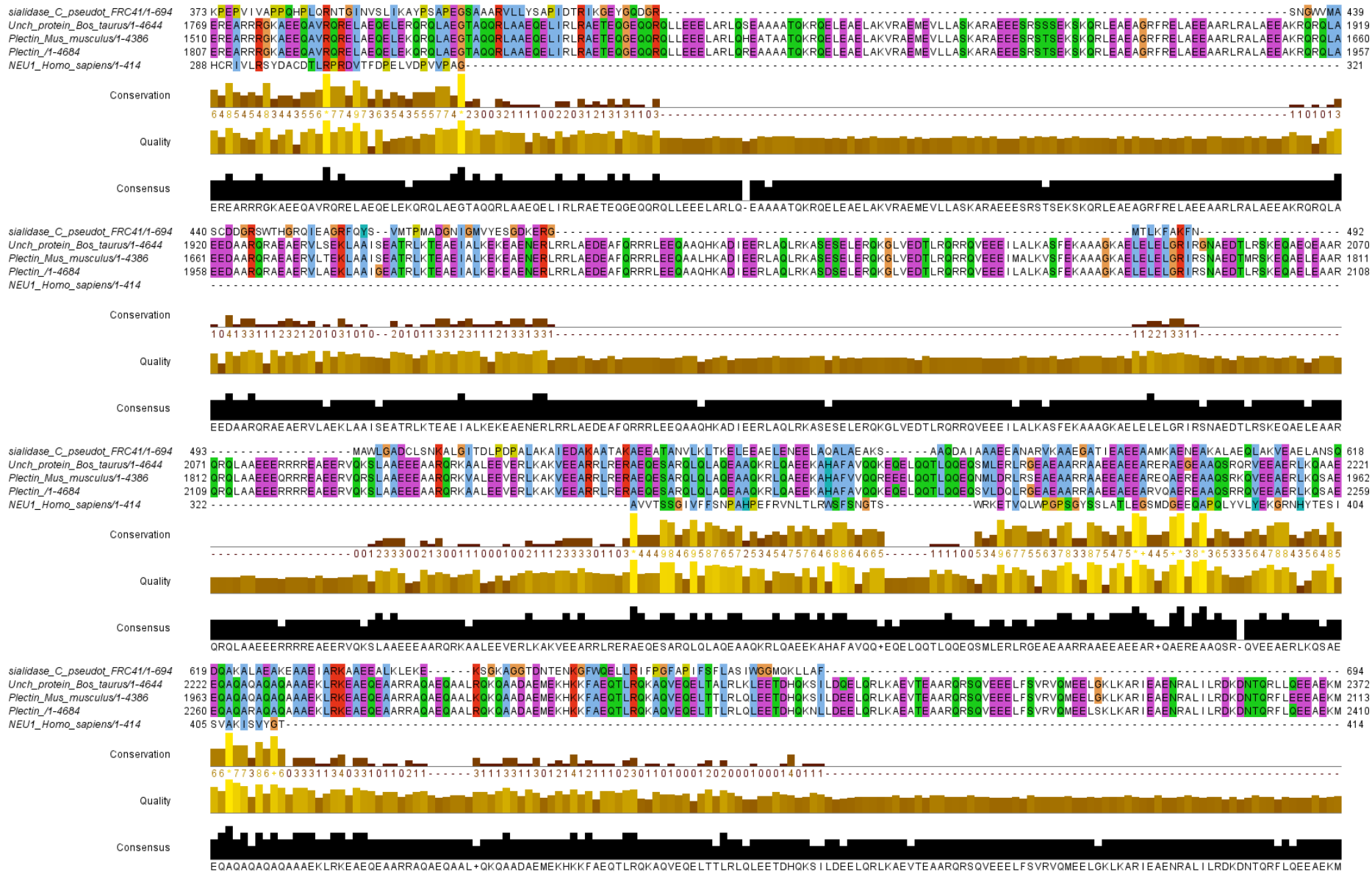


*sialidase\_C\_pseudot\_FRC41/1-694* 200 -----FFDATAGVNRSGVEDPILKRSVDDTWIIESNDN-----GQTWGNMKLSELTEKIKVRHKGPELAG-----RGRFATSNG 271  
*Unch\_protein\_Bos\_taurus/1-4644* 1467 ALEKQRLAEAHAQAKAQAEREAEQLRRMQEEVARREEAAVDAQQKRSIQEELQQLRQSSAEIQAKARQVEAAERSLRLEEIRVVRLQLETTERRGGAEGLQALRARAEEAEAKRQAEAEERLRRQVQDESQRKRAEAELA 1617  
*Plectin\_Mus\_musculus/1-4386* 1208 ALEKQRLAEAHAQAKAQAEREAEQLRRMQEEVARREEAAVDAQQKRSIQEELQQLRQSSAEIQAKARQVEAAERSLRLEEIRVVRLQLETTERRGGAEGLQALRARAEEAEAKRQAEAEERLRRQVQDESQRKRAEAELA 1358  
*Plectin/1-4684* 1505 ALEKQRLAEAHAQAKAQAEREAEQLRRMQEEVARREEAAVDAQQKRSIQEELQQLRQSSAEIQAKARQVEAAERSLRLEEIRVVRLQLETTERRGGAEGLQALRARAEEAEAKRQAEAEERLRRQVQDESQRKRAEAELA 1655  
*NEU1\_Homo\_sapiens/1-414* 157 -----LCAHRAGCQVASTMLVWSKDD-----GVSWTFRNLSLDIGTEVAPQPGS-----G 203



*sialidase\_C\_pseudot\_FRC41/1-694* 272 IOLKYGENAGR-----LFFPMSVHYSPNGAGAVANVALYSDDHGATWNVGGVGGEGETNGDENKLVLESDGRIMMNSKSTGNSNPNRGAERWRSYS-----EIOGETW 372  
*Unch\_protein\_Bos\_taurus/1-4644* 1618 LRVKAEAEAREKQRALQALELRLQAEAEERLRLQAEERARQVQVALETAQRSAAEELQSKASFAEKTAQLERTLEEEHYTVVQUREEATRREQQQAEERAREEAEERLERWQLKANEALRLRLQAEVAAQKSLAQAEAEKQKEEA 1768  
*Plectin\_Mus\_musculus/1-4386* 1359 LRVKAEAEAREKQRALQALELRLQAEAEERLRLQAEERARQVQVALETAQRSAAEELQSKASFAEKTAQLERTLEEEHYTVVQUREEATRREQQQAEERAREEAEERLERWQLKANEALRLRLQAEVAAQKSLAQAEAEKQKEEA 1509  
*Plectin/1-4684* 1656 LRVKAEAEAREKQRALQALELRLQAEAEERLRLQAEERARQVQVALETAQRSAAEELQSKASFAEKTAQLERTLEEEHYTVVQUREEATRREQQQAEERAREEAEERLERWQLKANEALRLRLQAEVAAQKSLAQAEAEKQKEEA 1806  
*NEU1\_Homo\_sapiens/1-414* 204 IQKREPFRKGR-----LIVGGHGLERDGAFCLLDDHGASWRYGSGVSG--IPYGGKKEENFNDECGYELPDGSSVYINAE-----NQNNYHC 287





*sialidase\_C\_pseudot\_FRC411/1-694*  
*Unch\_protein\_Bos\_taurus/1-4644* 2373 KVAEEAARLSVAAQEAARLRQLAEEDLAQQRALAEKMLKEKMQAVQEAATRLKAEAEELLQOQKELAQEQARRLQEDKEQMAQQLAAEETGGFQRTLEAERQRQLEMSAEAEERLRLRVAEMSRARAEEDAQRFRKQAEI I SAKLHRTLELAT 2523  
*Plectin\_Mus\_musculus/1-4386* 2114 KVAEEAARLSVAAQEAARLRQLAEEDLAQQRALAEKMLKEKMQAVQEAATRLKAEAEELLQOQKELAQEQARRLQEDKEQMAQQLAAEETGGFQRTLEAERQRQLEMSAEAEERLRLRVAEMSRARAEEDAQRFRKQAEI I GEKLRHTELAT 2264  
*Plectin/1-4684* 2411 KVAEEAARLSVAAQEAARLRQLAEEDLAQQRALAEKMLKEKMQAVQEAATRLKAEAEELLQOQKELAQEQARRLQEDKEQMAQQLAAEETGGFQRTLEAERQRQLEMSAEAEERLRLRVAEMSRARAEEDAQRFRKQAEI I GEKLRHTELAT 2561  
*NEU1\_Homo\_sapiens/1-414*

Conservation

Quality

Consensus

KQVAEEAARLSVAAQEAARLRQLAEEDLAQQRALAEKMLKEKMQAVQEAATRLKAEAEELLQOQKELAQEQARRLQEDKEQMAQQLAAEETGGFQRTLEAERQRQLEMSAEAEERLRLRVAEMSRARAEEDAQRFRKQAEI I GEKLRHTELAT

*sialidase\_C\_pseudot\_FRC411/1-694*  
*Unch\_protein\_Bos\_taurus/1-4644* 2524 QEKVTLVQTLLEIQRQQSDHDAERLREAIAELEREKELKQKAEALLQKSEEMQTVQQEQLLQETQALQQSFLSEKDSLLQRRERFIEQEKAKLQQLFQDEVAKAQQLREEQQRQQQMEQEKQ-L-ASMEEARRRQREAEQVRRKQEEEL 2674  
*Plectin\_Mus\_musculus/1-4386* 2265 QEKVTLVQTLLEIQRQQSDHDAERLREAIAELEREKELKQKAEALLQKSEEMQTVQQEQLLQETQALQQSFLSEKDSLLQRRERFIEQEKAKLQQLFQDEVAKAQQLREEQQRQQQMEQEKQELMASMEARRRQREAEQVRRKQEEELH 2415  
*Plectin/1-4684* 2562 QEKVTLVQTLLEIQRQQSDHDAERLREAIAELEREKELKQKAEALLQKSEEMQTVQQEQLLQETQALQQSFLSEKDSLLQRRERFIEQEKAKLQQLFQDEVAKAQQLREEQQRQQQMEQEKQRLVASMARRRQREAEQVRRKQEEELQ 2712  
*NEU1\_Homo\_sapiens/1-414*

Conservation

Quality

Consensus

QEKVTLVQTLLEIQRQQSDHDAERLREAIAELEREKELKQKAEALLQKSEEMQTVQQEQLLQETQALQQSFLSEKDSLLQRRERFIEQEKAKLQQLFQDEVAKAQQLREEQQRQQQMEQEKQ-L-ASMEEARRRQREAEQVRRKQEEELQ-

*sialidase\_C\_pseudot\_FRC411/1-694*  
*Unch\_protein\_Bos\_taurus/1-4644* 2675 LEQQRQQEQLLAEENRRLRERLQLEEEHRAALAHSEEITAAQAAATRALPNGODATDGP--AAEPEHAFEGLRQKVPAAQLQEAAGILSVEEVRVGHHTTVAELTQREQVRRYLQGRSSIAGLLLKPTNEKLSVYALQRLQLLSPGTA 2823  
2416 LEQQRQQEQLLAEENRRLRERLQLEEEHRAALAHSEEITAAQAAATRALPNGODATDGP--AAEPEHAFEGLRQKVPAAQLQEAAGILSVEEVRVGHHTTVAELTQREQVRRYLQGRSSIAGLLLKPTNEKLSVYALQRLQLLSPGTA 2565  
*Plectin/1-4684* 2713 LEQQRQQEQLLAEENRRLRERLQLEEEHRAALAHSEEITAAQAAATRALPNGODATDGP--AAEPEHAFEGLRQKVPAAQLQEAAGILSVEEVRVGHHTTVAELTQREQVRRYLQGRSSIAGLLLKPTNEKLSVYALQRLQLLSPGTA 2863  
*NEU1\_Homo\_sapiens/1-414*

Conservation

Quality

Consensus

LEQQRQQE-LLAEENRRLRERLQ-LEEEHRAALAHSEE-ITAAQAAATRALPNGODATDGP--AAEPEHAFEGLRQKVPAAQLQEAAGILS-VEEVRVGHHTTVAELTQREQVRRYLQGRSSIAGLLLKPTNEKLSVYALQRLQLLSPGTA

*sialidase\_C\_pseudot\_FRC411/1-694*  
*Unch\_protein\_Bos\_taurus/1-4644* 2824 VILLEAQAASGFLLDPVRRRLRVYNEAVKEGVVGPVPELHHKLLSAERAVTGYKDPYTGEEISLFDAMKKDLIVREHGIRLLEAQIATGGVDPVHSHRVPVAVYRQYFDEEMNRVLDPSDDTKGFFDPNTHENLTYLQLLRCVDEPPEI 2974  
2566 LILLEAQAASGFLLDPVRRRLRVYNEAVKEGVVGPVPELHHKLLSAERAVTGYKDPYTGEEISLFDAMKKDLIVREHGIRLLEAQIATGGVDPVHSHRVPVAVYRQYFDEEMNRVLDPSDDTKGFFDPNTHENLTYLQLLRCVDEPPEI 2716  
*Plectin/1-4684* 2864 LILLEAQAASGFLLDPVRRRLRVYNEAVKEGVVGPVPELHHKLLSAERAVTGYKDPYTGEEISLFDAMKKDLIVREHGIRLLEAQIATGGVDPVHSHRVPVAVYRQYFDEEMNRVLDPSDDTKGFFDPNTHENLTYLQLLRCVDEPPEI 3014  
*NEU1\_Homo\_sapiens/1-414*

Conservation

Quality

Consensus

LILLEAQAASGFLLDPVRRRLRVYNEAVKEGVVGPVPELHHKLLSAERAVTGYKDPYTGEEISLFDAMKKDLIVREHGIRLLEAQIATGGVDPVHSHRVPVAVYRQYFDEEMNRVLDPSDDTKGFFDPNTHENLTYLQLLRCVDEPPEI

*sialidase\_C\_pseudot\_FRC411/1-694*  
*Unch\_protein\_Bos\_taurus/1-4644* 2975 GLRLLP L T D Q A A K G G E L V Y T D S E A R D V F E K A T V S A P F G K F G G K T V T I W E L I N S E Y F T A E Q R R D L L R Q F R T G - I T V E K I I K I V I I V I E E H E Q K Q L C F E G L R A L V P A A E L L E S G V I D W L F R L L Q L C E R S V Q E V A V E A V R R A L R G S G V I A G 3125  
*Plectin\_Mus\_musculus/1-4386* 2717 GLRLLP L T D K A A K G G E L V Y T D T E A R D V F E K A T V S A P F G K F G G R T V T I W E I I N S E Y F T A E Q R R D L L Q F R T G H I T V E K I I K I V I I V V E E H E R K Q L C F E G L R A L V P A A E L L D S G V I S H E L Y Q L Q R G E R S V R E V A E A D S V R Q A L R G T N V I A G 2867  
*Plectin/1-4684* 3015 GLCLL P L T D K A A K G G E L V Y T D S E A R D V F E K A T V S A P F G K F G G K T V T I W E I I N S E Y F T A E Q R R D L L R Q F R T G R I T V E K I I K I I I V V E E Q E Q K R L C F E G L R S L V P A A E L L S R V I D R E L Y Q L L Q R G E R V R D V A E V D T V R R A L R G A N V I A G 3165  
*NEU1\_Homo\_sapiens/1-414*

Conservation

Quality

Consensus

GLRLLP L T D K A A K G G E L V Y T D S E A R D V F E K A T V S A P F G K F G G K T V T I W E I I N S E Y F T A E Q R R D L L R Q F R T G - I T V E K I I K I V I I V V E E H E Q K Q L C F E G L R A L V P A A E L L E S G V I D - E L Y Q L Q R G E R S V R E V A E V D - V R R A L R G - N V I A G

*sialidase\_C\_pseudot\_FRC411/1-694*  
*Unch\_protein\_Bos\_taurus/1-4644* 3126 W L E E A G K L S I Y E A L K K L L Q P E A A V A L L E A A G T G H I I D P A T S A R L T V D E A V R A G L V G P E L H E K L L S A E K A V T G Y R D P Y S G Q S V S L F A L K K G L I P R E Q G L R L L D A Q L S T G G I V D P S K S H R V P L D V A C A R G Y L D K E T S T A L T A P R D D A K 3276  
*Plectin\_Mus\_musculus/1-4386* 2868 W L E E A G K L S I Y E A L K K D L L Q P E A V A L L E A A G T G H I I D P A T S A R L T V D E A V R A G L V G P E L H E K L L S A E K A V T G Y R D P Y S G Q S V S L F A L K K G L I P R E Q G L R L L D A Q L S T G G I V D P S K S H R V P L D V A Y A R G Y L D K E T N R A L T S P R D D A R 3018  
*Plectin/1-4684* 3166 W L E E A G K L S I Y N A L K K D L L S D M A V A L L E A A G T G H I I D P A T S A R L T V D E A V R A G L V G P E F H E K L L S A E K A V T G Y R D P Y T G Q S V S L F A L K K G L I P R E Q G L R L L D A Q L S T G G I V D P S K S H R V P L D V A C A R G C L D E E T S R A L S A P R A D A K 3316  
*NEU1\_Homo\_sapiens/1-414*

Conservation

Quality

Consensus

W L E E A G K L S I Y E A L K K D L L Q P E - A V A L L E A A G T G H I I D P A T S A R L T V D E A V R A G L V G P E L H E K L L S A E K A V T G Y R D P Y S G Q S V S L F A L K K G L I P R E Q G L R L L D A Q L S T G G I V D P S K S H R V P L D V A C A R G Y L D K E T S R A L T A P R D D A K

*sialidase\_C\_pseudot\_FRC411/1-694*  
*Unch\_protein\_Bos\_taurus/1-4644* 3277 T Y Y N P R I W E P A T Y S Q L Q Q R C R P D Q L T G L S L L P L S E K A A R A R Q E E L Y S E L Q A R E T F E K A - V E V P V G S F K G R A V T W E L I S S E Y F T A E Q R Q E L L R Q F R T G K V T V E K V I K I I I I V E E V E T T R R Q E R L S F S G L R A P V P A S E L L A S G I L S R A Q F E Q 3427  
*Plectin\_Mus\_musculus/1-4386* 3019 V Y H D P S T Q E P Y T Y S Q L Q R C R S D Q L T G L S L L P L S E K A V R A R Q E E V Y S E L Q A R E T L E Q A K V E V P V G S F K G R A M T W E L I S S E Y F T E E Q R Q E L L R Q F R T G K V T V E K V I K I V I I I V E E V E T R R Q E R L S F S G L R A P V P A S E L L D A K I L S R A Q F D Q 3169  
*Plectin/1-4684* 3317 A Y S D P S T Q E P A T Y S Q L Q R C R P D Q L T G L S L L P L S E K A A R A R Q E E L Y S E L Q A R E T F E K T V E V P V G G F K G R T V T W E L I S S E Y F T A E Q R Q E L L R Q F R T G K V T V E K V I K I I I I V E E V E T L R Q E R L S F S G L R A P V P A S E L L A S G V L S R A Q F E Q 3467  
*NEU1\_Homo\_sapiens/1-414*

Conservation

Quality

Consensus

- Y - D P S T - E P A T Y S Q L Q Q R C R P D Q L T G L S L L P L S E K A A R A R Q E E L Y S E L Q A R E T F E K A - V E V P V G S F K G R A V T W E L I S S E Y F T A E Q R Q E L L R Q F R T G K V T V E K V I K I - I T I V E E V E T - R Q E R L S F S G L R A P V P A S E L L A S G I L S R A Q F E Q

*sialidase\_C\_pseudot\_FRC411/1-694*  
*Unch\_protein\_Bos\_taurus/1-4644* 3428 L K D G K T S V K D L S E L D V R T L L Q G S G C L A G I Y L E D I K E K V T I Y E A M R R G L L R P S T A I L L L E A Q A A T G F L V D P V R N Q R L Y V H E A V K A G V V G P E L H E K L L S A E K A V T G Y R D P Y S G S T I S L F Q A M K K G L V V R H G I R L L E A Q I A T G G I I D P V H S H 3578  
*Plectin\_Mus\_musculus/1-4386* 3170 L K D G K T S V K E L S E V G S V R T L L Q G S G C L A G I Y L E D S K E K V T I Y E A M R R G L L R P S T A T L L L E A Q A A T G F L V D P V R N Q R L Y V H E A V K A G V V G P E L H E K L L S A E K A V T G Y R D P Y S G N T I S L F Q A M K K G L V L R D H A I R L L E A Q A V A T G G I I D P V H S H 3320  
*Plectin/1-4684* 3468 L K D G K T S V K D L S E L G S V R T L L Q G S G C L A G I Y L E D I K E K V S I Y E A M R R G L L R A T A A L L L E A Q A A T G F L V D P V R N Q R L Y V H E A V K A G V V G P E L H E Q L L S A E K A V T G Y R D P Y S G S T I S L F Q A M Q K G L V L R Q H G I R L L E A Q I A T G G I I D P V H S H 3618  
*NEU1\_Homo\_sapiens/1-414*

Conservation

Quality

Consensus

L K D G K T S V K D L S E L G S V R T L L Q G S G C L A G I Y L E D I K E K V T I Y E A M R R G L L R P S T A - L L L E A Q A A T G F L V D P V R N Q R L Y V H E A V K A G V V G P E L H E K L L S A E K A V T G Y R D P Y S G S T I S L F Q A M K K G L V L R - H G I R L L E A Q I A T G G I I D P V H S H

*sialidase\_C\_pseudot\_FRC411/1-694*

*sialidase\_C\_pseudot\_FRC411/1-694*  
*Unch\_protein\_Bos\_taurus/1-4644*  
*Plectin\_Mus\_musculus/1-4386*  
*Plectin/1-4684*  
*NEU1\_Homo\_sapiens/1-414*

```

3579 RVPVDVAYQRGYFDEEMNRVLDAPSDDTKGFFDPNTHENLTYRQLLERCVEDPETGLRLLPLKGAEKAEVVEITQVYTEETRRAFEETQIDIPGGGSHGGSTMSLWEVMQSDLIPEEQRAQLMADFAGRVTKERMIIIIEIEKTEIV 3729
3321 RLPVDVAYQRGYFDEEMNRVLDAPSDDTKGFFDPNTHENLTYRQLLERCVEDPETGLRLLPLKGAEKAEVVEITQVYTEETRRAFEETQIDIPGGGSHGGSTMSLWEVMQSDLIPEDQARLMAADFAGRVTKERMIIIIEIEKTEII 3471
3619 RVPVDVAYQRGYFSEEMNRVLDAPSDDTKGFFDPNTHENLTYRQLLERCVEDPETGLRLLPLKGAEKAEVVEITQVYTEETRRAFEETQIDIPGGGSHGGSTMSLWEVMQSDLIPEEQRAQLMADFAGRVTKERMIIIIEIEKTEII 3769
  
```

Conservation

Quality

Consensus

RVPVDVAYQRGYFDEEMNRVLDAPSDDTKGFFDPNTHENLTYRQLLERCVEDPETGLRLLPLKGAEKAEVVEITQVYTEETRRAFEETQIDIPGGGSHGGSTMSLWEVMQSDLIPEEQRAQLMADFAGRVTKERMIIIIEIEKTEII

*sialidase\_C\_pseudot\_FRC411/1-694*  
*Unch\_protein\_Bos\_taurus/1-4644*  
*Plectin\_Mus\_musculus/1-4386*  
*Plectin/1-4684*  
*NEU1\_Homo\_sapiens/1-414*

```

3730 RQQNLSYDYVRRRLTAEDLHEARVIRREYSLLREGRSLRLEVAESAWRYLYGTGCVAGVYLPGRQTLTIYQALKKGLLSAEVARLLLEAQAATGFLLDPVKGERLTVDEAVRKGVLVGPPELHDLRLSAERAVTGYRDPYTEQTIISLF 3880
3472 RQQNLSYDYVRRRLTAEDLYEARIIISLETYNLFRGKKNLREVAESAWRYLYGTGAVAGVYLPGRQTLTIYQALKKGLLSAEVARLLLEAQAATGFLLDPVKGERLTVDEAVRKGVLVGPPELHDLRLSAERAVTGYRDPYTEQTIISLF 3622
3770 RQQGLSYDYVRRRLTAEDLFEARIIISLETYNLFRGKKNLREVAESAWCYLYGTGSAVAGVYLPGRQTLTIYQALKKGLLSAEVARLLLEAQAATGFLLDPVKGERLTVDEAVRKGVLVGPPELHDLRLSAERAVTGYRDPYTEQTIISLF 3920
  
```

Conservation

Quality

Consensus

RQQNLSYDYVRRRLTAEDL-EARIIISLETYNLFRGKKNLREVAESAWRYLYGTG-VAGVYLPGRQTLTIYQALKKGLLSAEVARLLLEAQAATGFLLDPVKGERLTVDEAVRKGVLVGPPELHDLRLSAERAVTGYRDPYTEQTIISLF

*sialidase\_C\_pseudot\_FRC411/1-694*  
*Unch\_protein\_Bos\_taurus/1-4644*  
*Plectin\_Mus\_musculus/1-4386*  
*Plectin/1-4684*  
*NEU1\_Homo\_sapiens/1-414*

```

3881 QAMKKDLIPAEALRLLDAQLATGGIVDPRLGFHLPLEVAYQRGYLNKDTHDQLSEPSVRSYVDPSTDERLSYQQLRRRDRDPSGQMLLLSDARKLIFRGLRQITVEELVRSQVMDEATLQQLQEGLTSIEEVTKNLQKFLGTS 4031
3623 QAMKKELIPAEALRLLDAQLATGGIVDPRLGFHLPLEVAYQRGYLNKDTHDQLSEPSVRSYVDPSTDERLSYQQLRRRDRDPSGQMLLLSDARKLIFRGLRQITVEELVRSQVMDEATLQQLQEGLTSIEEVTKNLQKFLGTS 3773
3921 QAMKKELIPAEALRLLDAQLATGGIVDPRLGFHLPLEVAYQRGYLNKDTHDQLSEPSVRSYVDPSTDERLSYQQLRRRDRDPSGQMLLLSDARKLIFRGLRQITVEELVRSQVMDEATLQQLQEGLTSIEEVTKNLQKFLGTS 4071
  
```

Conservation

Quality

Consensus

QAMKKELIPAEALRLLDAQLATGGIVDPRLGFHLPLEVAYQRGYLNKDTHDQLSEPSVRSYVDPSTDERLSYQQLRRRDRDPSGQMLLLSDARKLIFRGLRQITVEELVRSQVMDEATLQQLQEGLTSIEEVTKNLQKFLGTS

*sialidase\_C\_pseudot\_FRC411/1-694*  
*Unch\_protein\_Bos\_taurus/1-4644*  
*Plectin\_Mus\_musculus/1-4386*  
*Plectin/1-4684*  
*NEU1\_Homo\_sapiens/1-414*

```

4032 IAGVLDATKERLSVYQAMKGIIRPGTAFELLEAAATGYVDPKGLKLTVEEAVRMGIVGPEFKDKLLSAERAVTGYKDPYSGKLIISLFAAMKKGILKDHGIRLLEAQIATGGIDPEESHRLPVEVAYKRGFLDEEMNEILTDPSD 4182
3774 IAGVFDATKERLSVYQAMKGIIRPGTAFELLEAAATGYVDPKGLKLTVEEAVRMGIVGPEFKDKLLSAERAVTGYKDPYSGKLIISLFAAMKKGILKDHGIRLLEAQIATGGIDPEESHRLPVEVAYKRGFLDEEMNEILTDPSD 3924
4072 IAGVFDATKERLSVYQAMKGIIRPGTAFELLEAAATGYVDPKGLKLTVEEAVRMGIVGPEFKDKLLSAERAVTGYKDPYSGKLIISLFAAMKKGILKDHGIRLLEAQIATGGIDPEESHRLPVEVAYKRGFLDEEMNEILTDPSD 4222
  
```

Conservation

Quality

Consensus

IAGVFDATKERLSVYQAMKGIIRPGTAFELLEAAATGYVDPKGLKLTVEEAVRMGIVGPEFKDKLLSAERAVTGYKDPYSGKLIISLFAAMKKGILKDHGIRLLEAQIATGGIDPEESHRLPVEVAYKRGFLDEEMNEILTDPSD

*sialidase\_C\_pseudot\_FRC411/1-694*  
*Unch\_protein\_Bos\_taurus/1-4644* 4183 DTKGFDPNTEENLYLQLMERCVIDPQTGLRLLP LKEKKRERKTSKSSVRKRRVIVDPETGKEMSVYEA YRKGLIDHQTYLELSEQCEWEEITISSSDGVVKSMIIDRRSGRQYDIDDAIAKSLIDRSALDQYRAGTLSITTEFADML 4333  
*Plectin\_Mus\_musculus/1-4386* 3925 DTKGFDPNTEENLYLQLMERCVIDPQTGLRLLP LKEKKRERKTSKSSVRKRRVIVDPETGKEMSVYEA YRKGLIDHQTYLELSEQCEWEEITISSSDGVVKSMIIDRRSGRQYDIDDAITKNLIDRSALDQYRAGTLSITTEFADML 4075  
*Plectin/1-4684* 4223 DTKGFDPNTEENLYLQLMERCVIDPQTGLRLLP LKEKKRERKTSKSSVRKRRVIVDPETGKEMSVYEA YRKGLIDHQTYLELSEQCEWEEITISSSDGVVKSMIIDRRSGRQYDIDDAIAKNLIDRSALDQYRAGTLSITTEFADML 4373  
*NEU1\_Homo\_sapiens/1-414*

Conservation

Quality

Consensus

DTKGFDPNTEENLYLQLMERCVIDPQTGLRLLP LKEKKRERKTSKSSVRKRRVIVDPETGKEMSVYEA YRKGLIDHQTYLELSEQCEWEEITISSSDGVVKSMIIDRRSGRQYDIDDAIAKSLIDRSALDQYRAGTLSITTEFADML

*sialidase\_C\_pseudot\_FRC411/1-694*  
*Unch\_protein\_Bos\_taurus/1-4644* 4334 SGNAGGFRSRSSVGSSSYPISP AVSRTQLASWSDPTTEETGPVAGILDTEILEKVSITEAMHRNLVDNITGQRLLAEACTGGIDPNTGERFPVTEAVNKGLVDKIMVDRINLAQKAFCGFEDPRTKTKMSAAQALKKGLWLYYEAQGRF 4484  
*Plectin\_Mus\_musculus/1-4386* 4076 SGNAGGFRSRSSVGSSSYPISP AVSRTQLASWSDPTTEETGPVAGILDTEILEKVSITEAMHRNLVDNITGQRLLAEACTGGIDPSTGERFPVTEAVNKGLVDKIMVDRINLAQKAFCGFEDPRTKTKMSAAQALKKGLWLYYEAQGRF 4326  
*Plectin/1-4684* 4374 SGNAGGFRSRSSVGSSSYPISP AVSRTQLASWSDPTTEETGPVAGILDTEILEKVSITEAMHRNLVDNITGQRLLAEACTGGIDPSTGERFPVTEAVNKGLVDKIMVDRINLAQKAFCGFEDPRTKTKMSAAQALKKGLWLYYEAQGRF 4524  
*NEU1\_Homo\_sapiens/1-414*

Conservation

Quality

Consensus

SGNAGGFRSRSSVGSSSYPISP AVSRTQLASWSDPTTEETGPVAGILDTEILEKVSITEAMHRNLVDNITGQRLLAEACTGGIDPNTGERFPVTEAVNKGLVDKIMVDRINLAQKAFCGFEDPRTKTKMSAAQALKKGLWLYYEAQGRF

*sialidase\_C\_pseudot\_FRC411/1-694*  
*Unch\_protein\_Bos\_taurus/1-4644* 4485 LEVQYLTGGLIEPDTTPGRVPLDEALQRGTVDARTAQKLRDVSAYS KYLTCPKTKLKI SYKDALDRSMVEEGTGLRLLLEAAADSSKGYYSFYSSVSGSGSTAGSRTGSRGSRAGSRRGSFDATGSGFSMTFSSSSYSSSGYGRRYASGPSAS 4635  
*Plectin\_Mus\_musculus/1-4386* 4227 LEVQYLTGGLIEPDTTPGRVPLDEALQRGTVDARTAQKLRDVSAYS KYLTCPKTKLKI SYKDALDRSMVEEGTGLRLLLEAAADSSKGYYSFYSSVSGSGSTAGSRTGSRGSRAGSRRGSFDATGSGFSMTFSSSSYSSSGYGRRYASGPSAS 4377  
*Plectin/1-4684* 4525 LEVQYLTGGLIEPDTTPGRVPLDEALQRGTVDARTAQKLRDVSAYS KYLTCPKTKLKI SYKDALDRSMVEEGTGLRLLLEAAADSSKGYYSFYSSVSGSGSTAGSRTGSRGSRAGSRRGSFDATGSGFSMTFSSSSYSSSGYGRRYASGPSAS 4675  
*NEU1\_Homo\_sapiens/1-414*

Conservation

Quality

Consensus

LEVQYLTGGLIEPDTTPGRVPLDEALQRGTVDARTAQKLRDVSAYS KYLTCPKTKLKI SYKDALDRSMVEEGTGLRLLLEAAADSSKGYYSFYSSVSGSGSTAGSRTGSRGSRAGSRRGSFDATGSGFSMTFSSSSYSSSGYGRRYASGPSAS

*sialidase\_C\_pseudot\_FRC411/1-694* -----  
*Unch\_protein\_Bos\_taurus/1-4644* 4636 LGGPES TAA 4644  
*Plectin\_Mus\_musculus/1-4386* 4378 LGGPES AVA 4386  
*Plectin/1-4684* 4676 LGGPES AVA 4684  
*NEU1\_Homo\_sapiens/1-414* -----

Conservation

Quality

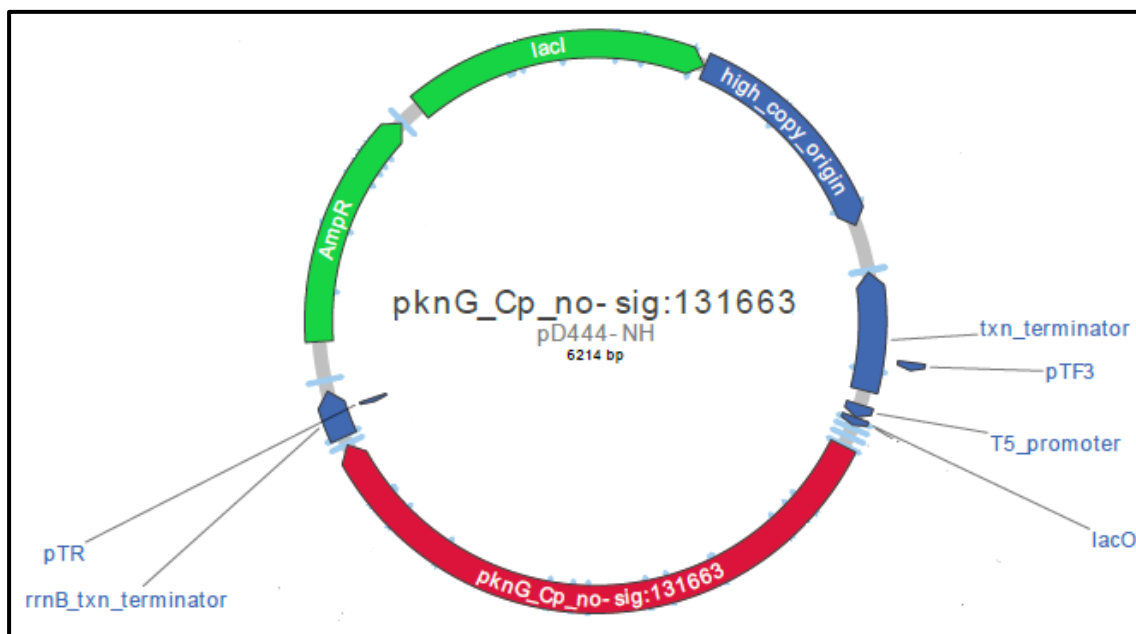
Consensus

LGGPESAVA

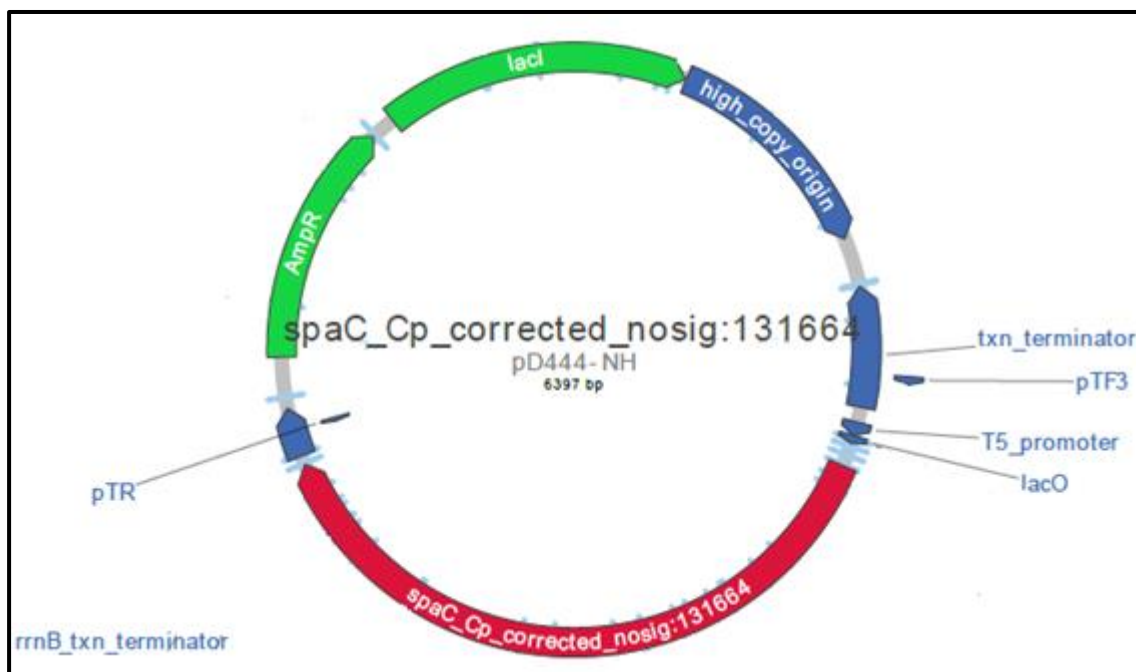
## **ANEXO 4**

**Mapa dos plasmídeos pD444-NH:*pkng*, pD444-NH:*spac*, pD444-NH:*sodc*  
e pD444-NH:*nanh***

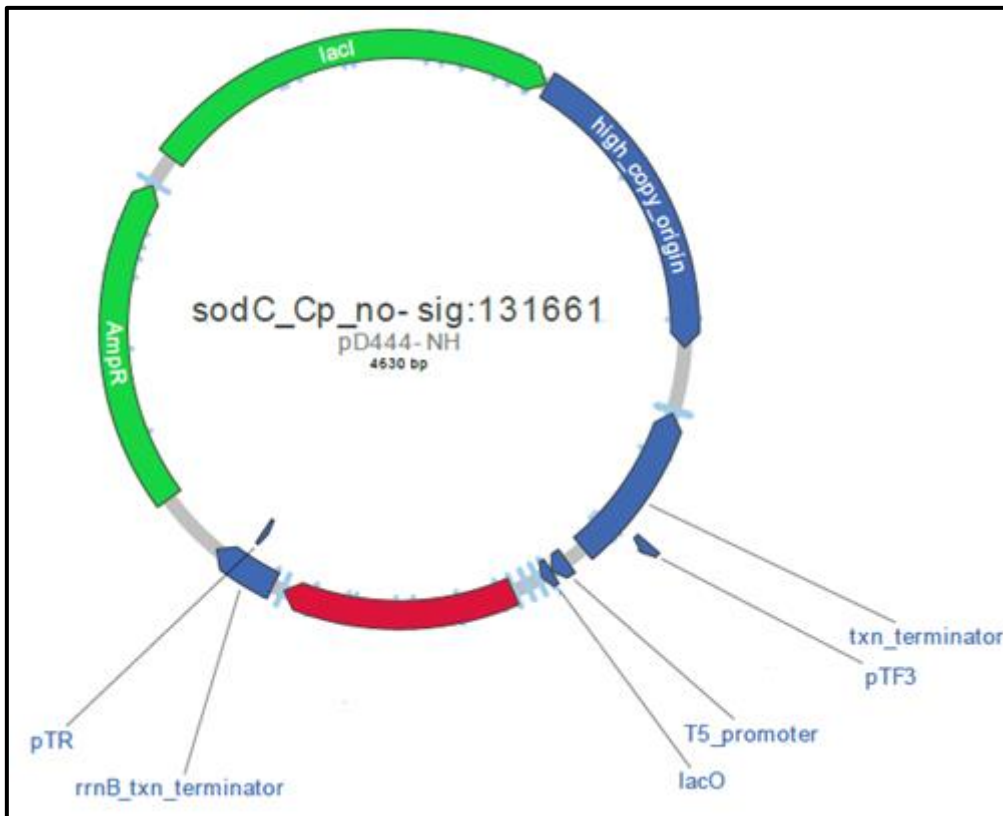




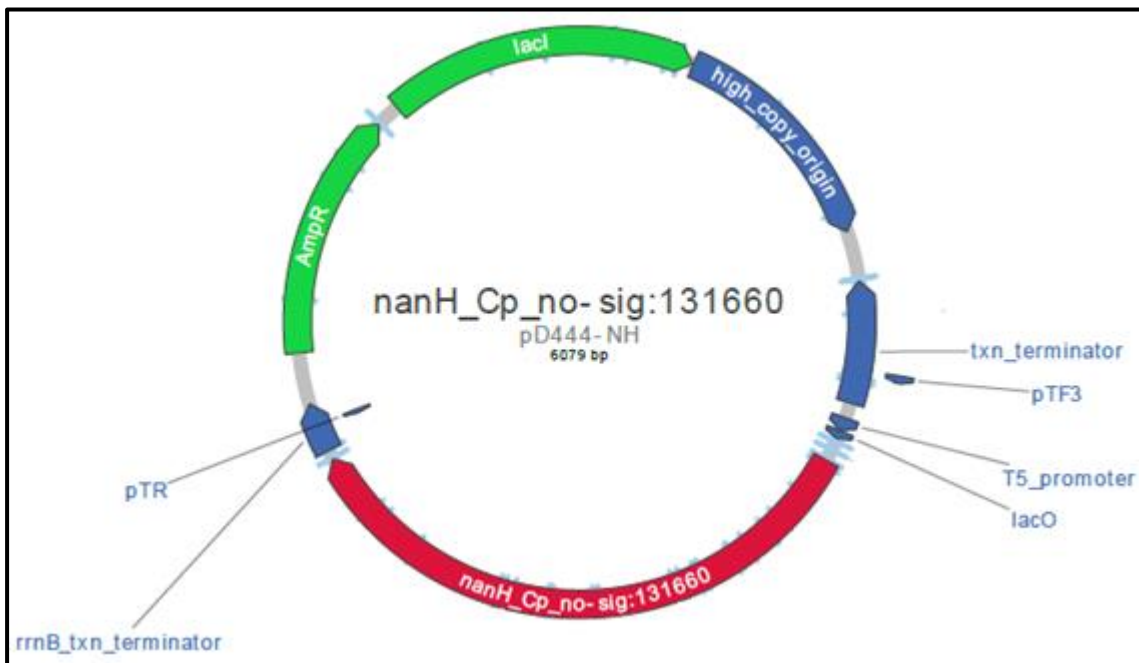
Representação esquemática do vetor pD444-NH contendo o inserto *pkng* (pD444-NH:*pkng*).



Representação esquemática do vetor pD444-NH contendo o inserto *spaC* (pD444-NH:*spaC*).



Representação esquemática do vetor pD444-NH contendo o inserto *sodC* (pD444-NH:*sodC*).



Representação esquemática do vetor pD444-NH contendo o inserto *nanH* (pD444-NH:*nanH*).