



UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE BIOLOGIA

**Maria Luiza Caldas Nogueira**

“Expressão e caracterização de proteínas envolvidas na via da quinase mTOR e na divisão celular bacteriana”

Este exemplar corresponde à redação final da tese defendida pelo(a) candidato (a) Maria Luiza Caldas Nogueira e aprovada pela Comissão Julgadora.

Dissertação apresentada ao Instituto de Biologia para obtenção do Título de Mestre em Biologia Funcional e Molecular, na área de Bioquímica.

Orientadora: Profa. Dra. Ana Carolina de Mattos Zeri

Campinas, 2012

FICHA CATALOGRÁFICA ELABORADA POR  
ROBERTA CRISTINA DAL' EVEDOVE TARTAROTTI – CRB8/7430  
BIBLIOTECA DO INSTITUTO DE BIOLOGIA - UNICAMP

N689e Nogueira, Maria Luiza Caldas, 1984-  
Expressão e caracterização de proteínas envolvidas  
na via da quinase mTOR e na divisão celular bacteriana /  
Maria Luiza Caldas Nogueira. – Campinas, SP: [s.n.],  
2012.

Orientador: Ana Carolina de Mattos Zeri.  
Dissertação (mestrado) – Universidade Estadual de  
Campinas, Instituto de Biologia.

1. Redobramento de proteína. 2. Mapeamento de  
interação de proteínas. 3. Proteínas FtsZ. 4.  
Proteínas MinC. 5. Complexo Ragulator. I. Zeri, Ana  
Carolina de Mattos. II. Universidade Estadual de  
Campinas. Instituto de Biologia. III. Título.

Informações para Biblioteca Digital

**Título em Inglês:** Expression and characterization of proteins involved in the mTOR  
kinase pathway and bacterial cell division

**Palavras-chave em Inglês:**

Protein refolding  
Protein interaction mapping  
FtsZ proteins  
MinC proteins  
Ragulator complex

**Área de concentração:** Bioquímica

**Titulação:** Mestre em Biologia Funcional e Molecular

**Banca examinadora:**

Ana Carolina de Mattos Zeri [Orientador]  
Ana Carolina Migliorini Figueira  
Jörg Kobarg

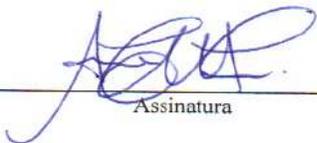
**Data da defesa:** 23-02-2011

**Programa de Pós Graduação:** Biologia Funcional e Molecular

Campinas, 23 de fevereiro de 2012

**BANCA EXAMINADORA**

Profa. Dra. Ana Carolina de Mattos Zeri (Orientadora)

  
Assinatura

Profa. Dra. Ana Carolina Migliorini Figueira

  
Assinatura

Prof. Dr. Jörg Kobarg

  
Assinatura

Profa. Dra. Juliana Helena Costa Smetana

\_\_\_\_\_  
Assinatura

Prof. Dr. Mário Tyago Murakami

\_\_\_\_\_  
Assinatura

Em memória de minha avó, Nona.

## **AGRADECIMENTOS**

À Coordenadoria de Apoio à Pesquisa, CAPES, e à Fundação de Amparo à Pesquisa do Estado de São Paulo, FAPESP, pelo auxílio financeiro concedido a este projeto.

Ao programa de Pós-graduação em Biologia Funcional e Molecular da UNICAMP, em especial à Andréia.

À Eneida de Paula, pela oportunidade de realização de Estágio no Programa de Estágio Docência, PED.

Ao Laboratório Nacional de Biociências, LNBio, e ao Centro Nacional de Pesquisas em Energia e Materiais, CNPEM, pela infra-estrutura oferecida.

Aos pesquisadores, funcionários e alunos do LNBio, pelo suporte técnico, convivência e troca de experiências.

Aos Profs. Ana Carolina Figueira, Carmen Ferreira e Mario Murakami, pela participação no exame de qualificação e avaliação deste trabalho.

Aos membros da banca - Profs. Ana Carolina Figueira e Jorg Kobarg - pela disposição em participar da avaliação final deste trabalho.

À Ana Carolina Zeri, pela amizade e orientação de forma eficiente e amorosa, sendo um importante alicerce no desenvolvimento deste trabalho e de minha formação profissional.

À amiga Juliana Smetana que extra-oficialmente desempenhou o papel de co-orientação deste trabalho, onde esteve extremamente presente e teve papel imprescindível para este trabalho e principalmente para minha formação profissional.

Ao Maurício Sforça, que com muita paciência e dedicação me ensinou muito sobre RMN.

Ao Jorge Neves, companheiro de equipe que aprofundou meus conhecimentos na área de RMN.

A Rodrigo Portugal e a Jefferson Betini, pela oportunidade de colaboração com Microscopia Eletrônica, muito válida para este trabalho.

Ao Alisson Cardoso, pelo auxílio com a reação de cross-linking

A Eduardo M. Cilli, por me receber em seu laboratório e pelos seus ensinamentos na área de síntese de peptídeos de fase sólida.

A Frederico Gueiros Filho, pela oportunidade de colaboração com as proteínas FtsZ e MinC, muito presente neste trabalho.

À amigas Ana Lu, Débora, Cristiane, Cynthia e Pati, companheiras de bancada que participaram assiduamente de minha vida no desenvolvimento deste trabalho. A participação de vocês enriqueceu muito este trabalho.

À amigas Bruna, Desirrè, Roberta e Vivi.

À técnica e amiga Jaqueline Zanella, pelo auxílio e orientação no processo de purificação de proteínas.

Aos muitos amigos que fiz neste Laboratório, pela convivência, troca de experiência, sugestões e críticas, em especial: Alexandre, Américo, Emerson, Eli, Juliana Oliveira e Roberta.

Ao pessoal do LEC que me auxiliaram na interpretação dos dados de CD.

À Adriana P. Leme, pela agilidade e eficiência no auxílio com a Espectrometria de Massas.

Ao meu namorado João Ricardo, por estar sempre presente e me incentivar nos momentos difíceis.

A meus pais que com carinho e dedicação trilharam meu caminho para que chegasse até aqui.

À minhas irmãs, Gabi, Carol e Mari, que foram sempre meu porto seguro, me auxiliando, incentivando e resolvendo meus conflitos com suas “idéias”.

Ao Biel, a Fefe e ao “*baby*” da Carol, por me servirem de incentivo.

E por fim, a Deus, por tornar tudo possível

# SUMÁRIO

ÍNDICE DE FIGURAS .....	XII
ÍNDICE DE TABELAS .....	XIV
RESUMO .....	XV
ABSTRACT .....	XVI
1. CAPÍTULO 1.....	2
1.1. JUSTIFICATIVA.....	2
ESTUDO DO COMPLEXO RAGULATOR .....	2
FTSZ E SEUS MODULADORES .....	2
1.2. OBJETIVO .....	3
2.    CAPÍTULO 2 – O COMPLEXO <i>RAGULATOR</i> .....	3
2.1    INTRODUÇÃO.....	3
A QUINASE MTOR E O CONTROLE DO CRESCIMENTO CELULAR .....	3
REGULAÇÃO DE MTORC1 POR AMINOÁCIDOS.....	5
ESTRUTURA DAS PROTEÍNAS DO COMPLEXO RAGULATOR .....	9
2.2. RESULTADOS OBTIDOS NA CARACTERIZAÇÃO DA PROTEÍNA P18 .....	15
2.2.1. EXPRESSÃO E PURIFICAÇÃO DOS CLONES DE MP1/P14.....	15
2.2.2. ESTRATÉGIAS DE CLONAGEM DA P18 .....	16
2.2.3. PREDIÇÃO DE DESORDEM DA P18 .....	18
2.2.4. EXPRESSÃO DA PROTEÍNA DE MAMÍFERO P18 EM BACTÉRIA .....	18
2.2.5. PADRONIZAÇÃO DAS CONDIÇÕES DE DIÁLISE DA P18 .....	19
2.2.6. PADRONIZAÇÃO DA PURIFICAÇÃO DA PROTEÍNA P18 RENATURADA COM URÉIA .....	21
2.2.7. ESPALHAMENTO DE LUZ DINÂMICO – DLS .....	27
2.2.8. MICROSCOPIA ELETRÔNICA .....	28
2.2.9. DICROISMO CIRCULAR – CD.....	30
2.2.10. CROSS-LINKING .....	32
3. CAPÍTULO 3 – FTSZ E SUA MODULADORA MINC.....	35
3.1. INTRODUÇÃO.....	35

FTSZ E A DIVISÃO CELULAR.....	35
O SISTEMA MINCDE.....	37
3.2 RESULTADOS DA CARACTERIZAÇÃO DA INTERAÇÃO DE FTSZ COM MINC.....	39
3.2.1. EXPRESSÃO DAS PROTEÍNAS MINC E FtsZ.....	39
3.2.2. EXPRESSÃO E PURIFICAÇÃO DAS PROTEÍNAS MINC E FtsZ .....	39
3.2.2.1 <i>Purificação da MinC</i> .....	40
3.2.2.2. <i>Purificação da FtsZ</i> .....	42
3.2.2.3. <i>Purificação dos mutantes de FtsZ</i> .....	44
3.2.3. TESTE DE ESTABILIDADE DE MINC A 40°C .....	45
3.2.4. TITULAÇÃO DE MINC COM FTSZ .....	46
3.2.5. ESPECTROS DE DICROÍSMO CIRCULAR DE MINC, FTSZ E MUTANTES .....	54
4. METODOLOGIA.....	58
4.1. MEIOS DE CULTURA, CEPAS E CLONES DE E. COLI .....	58
4.2. PREPARO DE CÉLULAS DE E. COLI COMPETENTES.....	59
4.2.1. <i>Preparo de E. coli competente para transformação por choque térmico</i> .....	59
4.2.2. <i>Preparo de E. coli eletrocompetente</i> .....	60
4.3. TRANSFORMAÇÃO DE E.COLI .....	60
4.3.3.1. <i>Transformação de E.coli por choque térmico</i> .....	60
4.3.2. <i>Eletroporação de E.coli</i> .....	60
4.4. ESTRATÉGIAS DE CLONAGEM GÊNICA.....	61
4.4.1. <i>Clonagem da p18</i> .....	61
4.4.2. <i>Clones obtidos por colaboradores</i> .....	62
4.5. ELETROFORESE EM GEL DE AGAROSE .....	62
4.6. EXPRESSÃO DE PROTEÍNAS RECOMBINANTES EM E.COLI. ....	62
4.6.1. <i>Expressão do clone de p18 em pET28a</i> .....	62
4.6.2. <i>Diálise e purificação da p18</i> .....	63
4.6.3. <i>Expressão e purificação dos clones de MP1/p14</i> .....	65
4.6.4. <i>Expressão e purificação das proteínas FtsZ e MinC</i> .....	67
4.6.4.1. <i>Purificação da MinC</i> .....	68
4.6.5. <i>Expressão e purificação de proteínas FtsZ e MinC em meio mínimo M9</i> .....	69
4.7. ELETROFORESE EM GEL DE POLIACRILAMIDA (SDS-PAGE). ....	69
4.8. ESTUDOS BIOFÍSICOS .....	70
4.8.1. <i>Espalhamento de Luz Dinâmico – DLS</i> .....	70
4.8.2. <i>Dicroísmo Circular – CD</i> .....	70

4.9. MICROSCOPIA ELETRÔNICA.....	71
4.10. ESPECTROS DE RMN .....	71
4.11. CROSS-LINKING DO RAGULATOR .....	72
5. CONCLUSÕES .....	74
O COMPLEXO RAGULATOR.....	74
A PROTEÍNA FTSZ E SUA MODULADORA MINC.....	75
6. PERSPECTIVAS .....	76
O COMPLEXO RAGULATOR.....	76
A PROTEÍNA FTSZ E SUA MODULADORA MINC.....	76
6. REFERÊNCIAS BIBLIOGRÁFICAS.....	78
ANEXO I – DECLARAÇÃO CIBIO/IB .....	82
ANEXO II – ARTIGO .....	84

## ÍNDICE DE ABREVIATURAS

<b>4E-BP1</b>	<i>Eukaryotic translation initiation factor 4E-binding protein 1</i>
<b>Akt/PKB</b>	<i>Protein kinase B</i>
<b>AMPK</b>	<i>AMP-activated protein kinase</i>
<b>ARG/GLU</b>	Arginina e Glutamina
<b>Avo3/TSC11</b>	<i>Temperature-sensitive suppressors of Csg2 mutants</i>
<b>DCM</b>	Diclorometano
<b>Deptor</b>	<i>Dep-domain-containing mTOR interacting protein</i>
<b>DIC</b>	Diisopropilcarbodiimida
<b>DivIVA</b>	<i>cell-division initiation protein DivIVA</i>
<b>DMF</b>	Dimetilformamida
<b>D.O.<sub>600</sub></b>	Densidade Óptica a 600nm
<b>DTT</b>	1,4-Ditiotreitól
<b>EDTA</b>	Ácido Etilenodiaminotetraacético
<b>EGO</b>	<i>exit from growth arrest</i>
<b>ERK</b>	<i>Extra cellular signal- regulated protein kinase</i>
<b>Kog1</b>	<i>Kontroller of Growth 1 ortologo da Raptor</i>
<b>FtsA</b>	<i>Filament temperature sensitive A</i>
<b>FtsZ</b>	<i>Filament temperature sensitive Z</i>
<b>GSE</b>	<i>Gap1p sorting in the endosome</i>
<b>Gtr1p</b>	<i>GTP-binding protein Ressemblance 1</i>
<b>Gtr2p</b>	<i>GTP-binding protein Ressemblance 2</i>
<b>HOBt</b>	1-hidroxibenzotriazol
<b>HsVam6</b>	<i>Homo sapiens VAcuolar Morphogenesis 6</i>
<b>IPTG</b>	Isopropyl β-D thiogalactoside
<b>KCl</b>	Cloreto de Potássio
<b>LB</b>	Meio Luria Bertani ( <i>Lysogen Broth</i> )
<b>LST8</b>	<i>lethal with sec 13 protein 8</i>
<b>MAPKs</b>	<i>Mitogen-activated protein kinase</i>
<b>MEK</b>	<i>MAPK/ERK kinase</i>
<b>MinC</b>	Cell-division inhibitor C
<b>MinD</b>	<i>ATP-binding protein MinD</i>
<b>MinE</b>	Cell division topological specificity factor

<b>MP1</b>	<i>mitogen-activated protein kinase scaffold protein 1</i>
<b>MreB</b>	<i>ROD SHAPE-DETERMINING PROTEIN MREB</i>
<b>mTOR</b>	Mammalian Target of Rapamycin
<b>mTORC1</b>	<i>mTOR Complex 1</i>
<b>mTORC2</b>	<i>mTOR Complex 2</i>
<b>NHSQC</b>	<sup>15</sup> N – Heteronuclear Single Quantum Correlation
<b>Noc</b>	<i>nucleoid occlusion protein</i>
<b>P14</b>	<i>Endosomal adaptor protein p14</i>
<b>P18</b>	<i>Ragulator complex protein LAMTOR1</i>
<b>PCR</b>	Reação em cadeia da polimerase
<b>PDK1</b>	<i>3-phosphoinositide-dependent protein kinase 1</i>
<b>PIK3</b>	<i>Phosphoinositide 3-kinase</i>
<b>PPAR<math>\gamma</math></b>	<i>Peroxisome proliferator-activated receptor – <math>\gamma</math></i>
<b>Rag</b>	Família de 4 proteínas membro da super família das GTPases <i>Ras</i>
<b>Ragulator</b>	Complexo formado pelas proteínas MP1, p14 e p18
<b>Raptor</b>	<i>Regulatory-associated protein of mTOR</i>
<b>Rheb</b>	<i>Ras-related GTPbinding protein</i>
<b>Rictor</b>	<i>Rapamycin-insensitive companion of mTOR</i>
<b>S6K 1</b>	<i>ribosomal S6 kinase 1</i>
<b>SlmA</b>	<i>sterol regulatory element binding protein 1</i>
<b>SREBP1</b>	<i>Sterol regulatory element binding protein</i>
<b>TAE</b>	Solução tampão Tris-Acetato-EDTA
<b>TCEP</b>	<i>Tris 2-carboxietil phosphina</i>
<b>TFA</b>	Ácido Trifluoroacético
<b>Torin 1</b>	mTOR inhibitor
<b>TRAP 1</b>	<i>Tumor necrosis factor-associated protein 1</i>
<b>TSC1/TSC2</b>	Complexo de Esclerose Tuberosa
<b>Vam6</b>	<i>VAcuolar Morphogenesis 6</i>
<b>ZipA</b>	Z-interacting protein A

## ÍNDICE DE FIGURAS

FIGURA 1 – ATIVAÇÃO DO COMPLEXO mTORC1 POR AMINOÁCIDOS.....	6
FIGURA 2 – VIAS DE REGULAÇÃO DO COMPLEXO mTORC1 .....	7
FIGURA 3 – ESTRUTURA DO HETERODÍMERO MP1/P14.....	10
FIGURA 4 – COMPARAÇÃO ENTRE AS ESTRUTURAS DE GSE1P E O HETERODÍMERO MP1/P14.....	11
FIGURA 5 – ESQUEMA DE INTERAÇÃO DA PROTEÍNA P18 COM SEUS LIGANTES.....	13
FIGURA 6 – COMPARAÇÃO ENTRE A GSE2P E A PROTEÍNA HUMANA P18 .....	14
FIGURA 7 - CROMATOGRAMA DE AFINIDADE DO HETERODÍMERO MP1/P14.....	15
FIGURA 8 - CROMATOGRAMA DE EXCLUSÃO POR TAMANHO MOLECULAR DE MP1/P14.....	16
FIGURA 9 - CLONAGEM DA PROTEÍNA P18.....	17
FIGURA 10 – PREDIÇÃO DE DESORDEM DA PROTEÍNA P18.....	18
FIGURA 11. FRAÇÃO INSOLÚVEL DA P18.....	19
FIGURA 12 – CROMATOGRAMA DE AFINIDADE DA PROTEÍNA P18 DIALISADA NAS CONDIÇÕES DE 1 A 6 DA TABELA 1.....	21
FIGURA 13. CROMATOGRAMA DA EXCLUSÃO POR TAMANHO MOLECULAR DA P18.....	22
FIGURA 14. TESTE DE INTERAÇÃO DO COMPLEXO <i>RAGULATOR</i> .....	23
FIGURA 15 – AFINIDADE DA DIÁLISE DA PROTEÍNA P18 SOB A CONDIÇÃO DE TAMPÃO 7 DA TABELA 1, COM E SEM ADIÇÃO DE MP1/P14.....	25
FIGURA 16 – CROMATOGRAFIA DE EXCLUSÃO POR TAMANHO MOLECULAR, EM COLUNA SUPERDEX 200 16/60®, DA DIÁLISE DA PROTEÍNA P18 SOB A CONDIÇÃO DE TAMPÃO 7 DA TABELA 1, COM E SEM ADIÇÃO DE MP1/P14.....	26
FIGURA 17. GRÁFICOS DE REGULARIZAÇÃO DE ESPALHAMENTO DINÂMICO DE LUZ (DLS) DA P18 .....	28
FIGURA 18. MICROSCOPIA ELETRÔNICA DA PROTEÍNA P18 RENATURADA .....	29
FIGURA 19 – MICROSCOPIA ELETRÔNICA DOS PONTOS DE DIÁLISE DA P18.....	30
FIGURA 20 – ESPECTRO DE CD DAS PROTEÍNAS DO COMPLEXO <i>RAGULATOR</i> .....	31
FIGURA 21 – GEL SDS-PAGE DO <i>CROSS-LINKING</i> DO RAGULATOR .....	32
FIGURA 22 – ESTRUTURAS CRISTALOGRAFICAS DA PROTEÍNA FtsZ DE 5 ESPÉCIES DIFERENTES .....	36
FIGURA 23 – ESQUEMA DE INIBIÇÃO DA DIVISÃO BACTERIANA EM <i>B. SUBTILLIS</i> .....	38
FIGURA 24 – CROMATOGRAMA DA PROTEÍNA MINC MARCADA COM <sup>15</sup> N E PURIFICADA POR AFINIDADE.....	40
FIGURA 25. CROMATOGRAFIA DA PROTEÍNA MINC PURIFICADA POR FILTRAÇÃO EM GEL NA COLUNA SUPERDEX 200.....	41
FIGURA 26. ESPECTRO UNIDIMENSIONAL DE HIDROGÊNIO DA PROTEÍNA MINC.....	42
FIGURA 27. CROMATOGRAFIA DE EXCLUSÃO POR TAMANHO MOLECULAR DA PROTEÍNA FtsZ PURIFICADA POR PRECIPITAÇÃO EM SULFATO DE AMÔNIO.....	43

FIGURA 28 – MUTANTES DE FtsZ.....	44
FIGURA 29 – CROMATOGRAMA DA FILTRAÇÃO EM GEL DO MUTANTE 255 SEM A CAT .....	45
FIGURA 30 – ESPECTRO <sup>15</sup> N-HSQC DO TESTE DE ESTABILIDADE DA MINC .....	46
FIGURA 31 – ESPECTRO <sup>15</sup> N-HSQC DA TITULAÇÃO DE MINC COM FtsZ NA PROPORÇÃO 50:1.....	47
FIGURA 32 – ESPECTRO <sup>15</sup> N-HSQC DA TITULAÇÃO DE MINC COM FtsZ NA PROPORÇÃO 10:1.....	48
FIGURA 33 – ESPECTRO <sup>15</sup> N-HSQC DA TITULAÇÃO DE MINC COM FtsZ NA PROPORÇÃO 5:1.....	49
FIGURA 34 – ESPECTRO <sup>15</sup> N-HSQC DA TITULAÇÃO DE MINC COM FtsZ NA PROPORÇÃO 2,5:1.....	50
FIGURA 35 – ESPECTRO <sup>15</sup> N-HSQC DA TITULAÇÃO DE MINC COM FtsZ NA PROPORÇÃO 1:1.....	51
FIGURA 36 – SOBREPOSIÇÃO DOS ESPECTROS DA TITULAÇÃO DE MINC NA PROPORÇÃO 5:1 DE FtsZ, 2,5:1 E 1:1 .....	52
FIGURA 37 – ESPECTRO DE <sup>15</sup> N-HSQC DE FtsZ. ....	53
FIGURA 38 – ESPECTRO DE CD DA PROTEÍNA FtsZ E MUTANTES.....	55

## ÍNDICE DE TABELAS

TABELA 1 – CONDIÇÕES DE TAMPÃO VARIADAS NA DIÁLISE DA P18 A 4°C. ....	20
TABELA 2 – VARIAÇÃO DA TEMPERATURA MÉDIA DE DESENOVELAMENTO TÉRMICO DE FTSZ SELVAGEM E MUTANTES COM MINC .....	56
TABELA 3 – CEPAS E ANTIBIÓTICOS UTILIZADOS NESTE TRABALHO .....	59
TABELA 4. DESENHO DOS OLIGONUCLEOTÍDEOS UTILIZADOS NA ESTRATÉGIA DE CLONAGEM DO cDNA DA PROTEÍNA P18 NO VETOR PET28A. ....	61
TABELA 5 – CONDIÇÕES DE EXPRESSÃO TESTADAS PARA O CLONA DA P18 (P18 + PET28A ).....	63
TABELA 6 – CONDIÇÕES DE TAMPÃO VARIADAS NA DIÁLISE DA P18 A 4°C .....	64
TABELA 7 – CLONES UTILIZADOS NA EXPRESSÃO DOS CLONES DE MP1/P14 .....	66
TABELA 8 – CLONES UTILIZADOS NA EXPRESSÃO DE FTSZ E MINC.....	67

## RESUMO

A mTOR é uma via de sinalização muito conservada que controla o crescimento celular em resposta à presença de nutrientes e fatores de crescimento. A desregulação dessa via em humanos está relacionada a doenças como câncer e diabetes. A quinase TOR é ativada na presença de aminoácidos e recentemente descobriu-se que as pequenas GTPases da família Rag são mediadoras da sinalização por Leucina. Essas GTPases são ancoradas na superfície do lisossomo por meio da interação com um complexo de três proteínas denominado *Ragulator*. Esse complexo também ancora um braço da via das MAPKs (MEK-ERK) aos lisossomos. O entendimento deste complexo pode nos ajudar a elucidar doenças em que a via da mTOR se encontra desregulada. Neste trabalho obtivemos o complexo *Ragulator*, através da expressão da proteína p18 em corpos de inclusão e sua renaturação através da adição de suas parceiras Mp1/p14 à diálise. Foram realizados estudos biofísicos com a intenção de caracterização do complexo, entretanto o alto grau de dissociação do mesmo resultou em certa dificuldade para caracterizá-lo completamente. Neste trabalho caracterizamos os agregados formados pela p18 e conseguimos reduzir sua formação através de diálise contendo agente redutor e suas proteínas parceiras. A renaturação da p18 na presença de MP1/p14 favoreceu seu rendimento, indicando a interação entre estas proteínas, porém não foi possível estabilizar o complexo *Ragulator*.

O estudo da divisão bacteriana é centralmente dependente de FtsZ, um homólogo procariótico das tubulinas. FtsZ desencadeia a divisão ao formar o “anel Z”, uma estrutura supramolecular constituída por polímeros de FtsZ que circunda o interior da célula e funciona como arcabouço do aparato de divisão. A formação do anel Z é regulada por moduladores, proteínas que afetam tanto negativamente como positivamente a capacidade de FtsZ polimerizar-se. A proteína MinC é um inibidor da polimerização de FtsZ, recrutada por MinD para a face interna da membrana plasmática, onde o complexo MinCD exerce sua função. MinCD representa um inibidor sítio-específico da polimerização da FtsZ, prevenindo a formação do anel Z nos pólos das células mas permitindo que isto aconteça na região central. A elucidação deste processo seria de grande valia para o desenho racional de inibidores da divisão bacteriana. Neste trabalho, comprovamos a interação entre MinC e FtsZ por Ressonância Magnética Nuclear. Estas proteínas não se encontravam em sua forma monomérica e o alto peso molecular do complexo impossibilitou a identificação dos aminoácidos envolvidos nesta interação, devido a limites da técnica <sup>15</sup>NHSQC. No momento, a proteína MinC está sendo expressa em presença de deutério, o que aumenta significativamente o limite da técnica de <sup>15</sup>NHSQC. Foram realizados ainda estudos biofísicos com intuito de caracterização da interação.

## ABSTRACT

The mTOR signaling pathway is a very well conserved pathway that controls cell growth in response to the presence of nutrients and growth factors. Deregulation of this pathway in humans is related to diseases like cancer and diabetes. The TOR kinase is activated in the presence of amino acids and it was recently discovered that the Rag small GTPases family are mediators of signaling by Leucine. These GTPases are anchored on the surface of the lysosome through interactions with a complex of three proteins called *Ragulator*. This complex also anchors an arm of the pathway of MAPKs (MEK-ERK) to lysosomes. Understanding this can help us to elucidate complex diseases in which the mTOR pathway is upregulated. In this work, the *Ragulator* complex was obtained through the expression of p18 protein in inclusion bodies and their refolding by adding their partners MP1/p14 to dialysis. Biophysical studies were conducted with the intention of characterizing the complex, however its high degree of dissociation resulted in some difficulty to characterize it completely. In this work we characterized the aggregates formed by p18 and managed reduce its formation by dialysis containing reducing agent and its partner proteins. The p18 renaturation with MP1/p14 improve its yield, indicating interaction among these proteins, however the *Ragulator* complex wasn't stabilized.

The study of bacterial division is centrally dependent on FtsZ, a prokaryotic homologue of tubulin. FtsZ triggers the division to form the "Z ring", a supramolecular structure consisting in FtsZ polymers that surrounds the cell and acts as a frame of the division apparatus. The formation of the Z ring is regulated by modulators, proteins that affect both negatively and positively the ability of FtsZ to polymerize. The MinC protein is an inhibitor of FtsZ polymerization, recruits MinD to the inner surface of the plasma membrane, where the complex MinCD exerts its function. MinCD is an inhibitor of site-specific polymerization of FtsZ, preventing the formation of the Z ring at the poles of the cells but allowing this to happen in the central region. The elucidation of this process would be invaluable for the rational design of bacterial division inhibitors. In this work, we confirmed the interaction between MinC and FtsZ by Nuclear Magnetic Resonance. These proteins were not in their monomeric form and the high molecular weight of the complex prevented the identification of the amino acids involved in this interaction, due to limitations of the <sup>15</sup>NHSQC technology. At present, the MinC protein is being expressed in the presence of deuterium, which significantly increases the limit of this technique <sup>15</sup>NHSQC. Biophysical studies were also performed with the aim of characterizing the interaction.

# Capítulo 1

## Justificativa e Objetivo

## Capítulo 1

### 1. Capítulo 1

Neste trabalho serão abordados dois tópicos, o primeiro envolvendo proteínas relacionadas à via da quinase mTOR e o segundo abordando proteínas envolvidas na divisão celular bacteriana, sendo ambos os tópicos separados por capítulos.

#### 1.1. JUSTIFICATIVA

##### *Estudo do complexo Ragulator*

A via da mTOR coordena importantes mecanismos celulares tais como, crescimento, proliferação, autofagia e outros. Embora esta via tenha sido muito estudada, sua regulação por aminoácidos é pouco conhecida. O Complexo *Ragulator* parece ter um papel central nesta regulação. Como sugerido em levedura, este complexo ao se ancorar aos lisossomos e possivelmente se ligar às GTPases Rag, promove uma aproximação da mTOR com seu ativador, Rheb. O conhecimento da estrutura do complexo *Ragulator* e do mecanismo de interação das proteínas envolvidas é de grande importância para o entendimento da sinalização por aminoácidos. Este trabalho tem o objetivo de obter as primeiras informações estruturais da p18 e de sua interação com MP1/p14.

##### *FtsZ e seus moduladores*

A divisão bacteriana é um processo fundamental que só agora começa a ser compreendido em seus detalhes moleculares. A elucidação da estrutura de um modulador e a caracterização dos sítios de interação entre FtsZ e seus moduladores representaria informação inédita e relevante para este objetivo. O papel central de FtsZ na divisão bacteriana e a sua universalidade nestes organismos tornou esta proteína um promissor alvo para o desenvolvimento de novos antibióticos. Um antibiótico que atue inibindo o funcionamento do divisomo representaria um novo modo de ação e, portanto, teria um importante papel no combate à resistência bacteriana. Existem diversos relatos de pequenas moléculas inibidoras da função de FtsZ que foram identificadas a partir de triagens “cegas”

## Capítulo 1

do tipo “*high-throughput*”. Esta abordagem é poderosa e já produziu um candidato a antibiótico promissor [1]. Em contrapartida, existe muito espaço ainda para o desenho racional no desenvolvimento de inibidores de FtsZ. A compreensão a nível estrutural e com resolução atômica, de como inibidores naturais de FtsZ afetam o seu funcionamento, pode produzir uma base de conhecimentos capazes de impulsionar o desenho racional de inibidores da divisão bacteriana.

### 1.2. OBJETIVO

Este trabalho visa atingir os seguintes objetivos:

- a) Clonagem do cDNA e padronização da expressão e purificação da proteína p18;
- b) Mapear a interação entre a p18 e os outros componentes do complexo *Ragulator* (MP1 e p14);
- c) Padronização da expressão e purificação das proteínas MinC e FtsZ em meio mínimo (M9) para estudos de interação por RMN;
- d) Caracterizar/mapear da interação entre MinC e FtsZ de *Bacillus subtilis* por RMN.

# Capítulo 2

## O Complexo *Regulator*

## Capítulo 2

### 2. Capítulo 2 – O complexo *Ragulator*

#### 2.1 INTRODUÇÃO

##### *A quinase mTOR e o controle do crescimento celular*

A quinase TOR (*Target of Rapamycin*), desempenha um papel essencial na regulação do crescimento celular integrando diferentes sinais, tais como níveis de energia, nutrientes e fatores de crescimento, para coordenar o crescimento e proliferação celular. A regulação da mTOR está envolvida nos processos de angiogênese, adipogênese e ativação de linfócitos T [2] e falhas nessa regulação em células humanas estão associadas a patologias como diabetes e câncer [3];[4].

A via de sinalização mediada pela TOR é seletivamente inibida por Rapamicina, uma droga lipofílica isolada de uma cepa de *Streptomyces hygroscopicus*. A Rapamicina e seus análogos (*rapalogues*) apresentam interesse clínico para o tratamento de tumores sólidos, transplante de órgãos (onde atua como imunossupressor), reestenose coronária e artrite reumatóide [2]. O complexo formado entre a Rapamicina e seu receptor intracelular, FKBP12, interage especificamente com a quinase TOR por meio de um resíduo de Serina conservado [5].

Essa quinase é conservada em todos os eucariotos e, em mamíferos, é conhecida como mTOR, FRAP ou RAFT1 [6];[7]. mTOR é encontrada em dois complexos, mTORC1 e mTORC2, que conservam em grande parte a composição de proteínas, a sensibilidade a Rapamicina e os processos celulares regulados pelos complexos TORC1 e TORC2 em levedura. As proteínas Kog1 e Avo3/TSC11, encontradas nos complexos TORC1 e TORC2, respectivamente, apresentam ortólogas em mamíferos denominadas Raptor (*Regulatory-associated protein of mTOR*) e Rictor (*Rapamycin-insensitive companion of mTOR*) [8];[9]. Raptor parece regular o complexo formado por mTOR e recrutar substratos para o mesmo [10];[8]. Lst8, presente nos dois complexos, também é conservada e em mamíferos, sendo conhecida como mLst8/GβL; sua deleção não afeta a atividade de mTORC1 *in vivo* e desestabiliza o complexo mTORC2, tornando-a essencial para atividade

## Capítulo 2

deste complexo [11].

A quinase mTOR age como um integrador de diversos sinais celulares na regulação do crescimento. Além de responder à presença de nutrientes (aminoácidos) mTOR responde também aos níveis de energia, por meio da quinase AMPK; a diversos tipos de estresse (por exemplo, hipóxia, estresse genotóxico, ácido fosfatídico e inflamação); e a fatores de crescimento. Na resposta a fatores de crescimento, mTORC1 é um efetor da via de sinalização mediada pela fosfatidilinositol 3-quinase (PI3K).

Estudos recentes apóiam um modelo no qual receptores ativados de fatores de crescimento (por exemplo, a insulina) ativam a PI3K, que então ativa a quinase Akt (PKB) por meio da quinase dependente de 3-fosfo-inositol (PDK1). A quinase Akt fosforila e inativa o heterodímero TSC1/TSC2 (*Tuberous Sclerosis Complex*), permitindo que a GTPase Rheb ative a quinase TOR [3]. Outro modo de ativação da TOR por fatores de crescimento e mitógenos depende do ácido fosfatídico, que é gerado pela fosfolipase D e interage diretamente com a quinase TOR [12]. Recentemente mTOR foi apontada como reguladora da síntese lipídica. mTORC1 regula positivamente a atividade de SREBP1 (*sterol regulatory element binding protein*) [13] e PPAR $\gamma$  (*peroxisome proliferator – activated receptor  $\gamma$* ) [14], dois fatores de transcrição que controlam a expressão de genes que codificam as proteínas envolvidas na homeostase de lipídeos e colesterol. mTORC1 fosforila a Lipina 1, que então se localiza no núcleo, possivelmente sendo a responsável pelo fenótipo de excentricidade nuclear na inibição de mTORC1 por Torin 1. A Lipina 1 regula a homeostase de lipídeos e sua distribuição núcleo/citoplasma provavelmente está envolvida na regulação dos alvos de expressão gênica de SREBP1 por mTORC1 [15]. O complexo mTORC1 ainda regula o metabolismo e a biogênese mitocondrial [2].

Os principais alvos celulares de mTORC1 são a quinase S6K e a proteína 4E-BP1, envolvidas no controle da tradução. S6K é fosforilada e ativada por mTORC1 e fosforila a proteína ribossomal S6, promovendo o início da tradução. Já 4E-BP1 é fosforilada e inativada por mTORC1, e dessa forma libera da inibição o fator de início da tradução eIF4E, responsável pelo reconhecimento de mRNAs com *cap*.

O complexo mTORC2 é considerado insensível a Rapamicina, entretanto ele apresenta uma sensibilidade reduzida que se torna aparente em alguns tipos celulares e após longas exposições à droga. Seu principal alvo na célula é o citoesqueleto de actina,

## Capítulo 2

controlado por meio das GTPases Rho, Rac e da quinase PKC. Outro importante substrato de mTORC2 é a quinase Akt/PKB, que é fosforilada e ativada por esse complexo na serina 473, localizada no motivo C-terminal hidrofóbico [16]. Para a ativação completa de Akt/PKB, é necessária também a fosforilação por PDK1 na treonina 308 na alça de ativação. Apesar de Akt/PKB ser um ativador de mTORC1, não se observa a ativação de mTORC1 em resposta à ativação de mTORC2, sugerindo que estas diferentes quinases que atuam como substrato de mTORC2 e ativador de mTORC1 sejam diferentes. Ao contrário de mTORC2, mTORC1 funciona como um regulador negativo de Akt/PKB por meio de um *feedback* negativo envolvendo a inibição da expressão do substrato do receptor de insulina (IRS) pela quinase S6K.

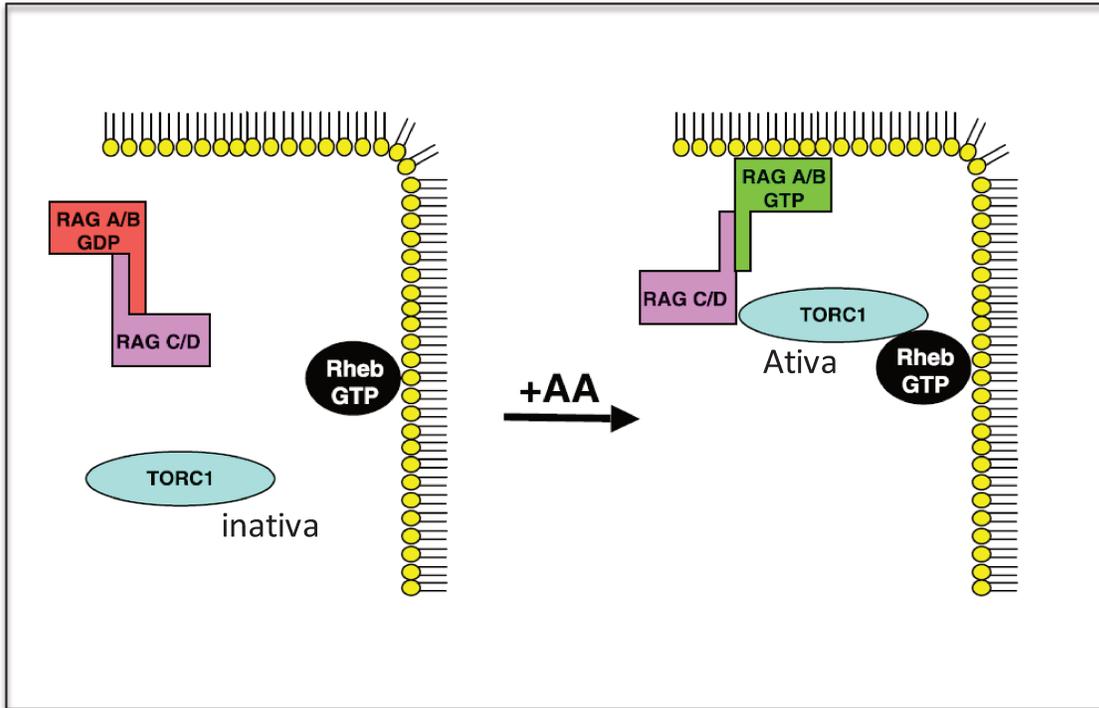
### *Regulação de mTORC1 por aminoácidos*

Uma das questões ainda não resolvidas com respeito à regulação de mTOR é a sua ativação por nutrientes, especialmente os aminoácidos de cadeia ramificada. O complexo mTORC1 é ativado pela presença de aminoácidos, permitindo que a célula coordene seu crescimento em função da disponibilidade de nutrientes. Essa é uma função primordial dessa via, presente desde organismos unicelulares, que foi integrada ao longo da evolução com sistemas de regulação por fatores de crescimento, o que é essencial a organismos multicelulares. Porém, enquanto a regulação de mTOR por fatores de crescimento envolvendo a via PI3K/Akt tem sido alvo de intensos estudos, o mecanismo de regulação por aminoácidos permaneceu pouco compreendido até recentemente [17].

A identificação das GTPases Rag como os principais mediadores da ativação de mTORC1 por aminoácidos, foi feita por dois grupos de pesquisa de forma independente [18]; [19]. Essa família de pequenas GTPases, pertencente à superfamília do protooncogene Ras, apresenta quatro membros que se associam na forma de heterodímeros RagA/B:RagC/D e alteram seu estado de ativação em função dos níveis de aminoácidos. Esses dímeros interagem diretamente com Raptor, componente de mTORC1, e ativam a quinase mTOR apenas quando RagA/B está no estado ativado (ligada a GTP), o que ocorre na presença de aminoácidos (**Figura 1**). A expressão de mutantes constitutivamente ativos dessas GTPases promove a ativação de mTORC1 mesmo na ausência de aminoácidos. A

## Capítulo 2

ativação de mTORC1 mediada por sua interação com as Rags depende de sua redistribuição para compartimentos vesiculares que contêm Rab7, e provavelmente o mecanismo de ativação se dá pela proximidade física de mTORC1 com Rheb, que é seu principal ativador [17].



**Figura 1 – Ativação do complexo mTORC1 por aminoácidos.** O Heterodímero Rag A/B, em presença de aminoácidos troca GDP por GTP e se liga a Raptor, um dos componentes do mTORC1, levando este seu complexo até os lisossomos, próximo a seu ativador, Rheb. Modificado de [20].

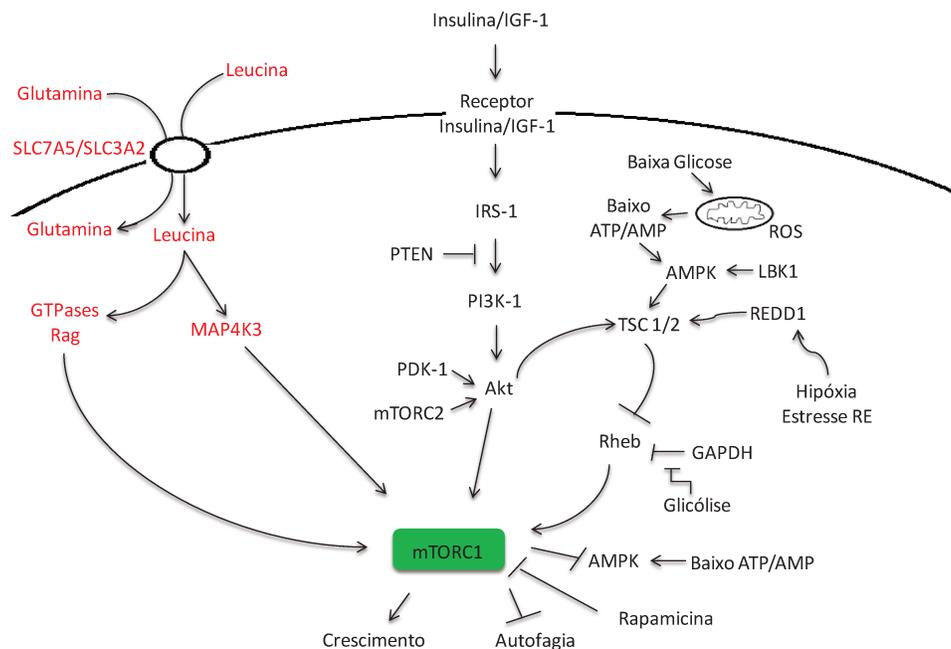
Estudos posteriores demonstraram que o compartimento subcelular para o qual mTORC1 se dirige após a ativação mediada por aminoácidos é o lisossomo, uma organela envolvida no processo de autofagia, que por sua vez é inibido pela ativação de mTOR. Nessa organela, foi identificado um novo complexo formado por três proteínas (MP1, p14 e p18) que ancora os heterodímeros de GTPases Rag por meio de uma interação direta. Devido a sua função como regulador das Rags, esse complexo foi denominado *Ragulator* [17]. Recentemente foi demonstrado que p18 não apenas interage com o lisossomo, como também tem um papel essencial na sua dinâmica. A perda p18 causa uma anormal

## Capítulo 2

distribuição e tamanho celular das organelas endossomo tardio/lisossomo e o acúmulo de componentes específicos dos endossomos tardios, tais como Rab7, RagC e LAMP1, sugerindo o acúmulo desta organela; e este fenótipo é repetido pela inibição da mTOR por Rapamicina, sugerindo uma interligação entre a atividade da p18 e via da mTOR. A perda da p18 inibe a integração entre as o endossomo tardio e o lisossomo, evidenciando sua da influência na dinâmica destas organelas [21].

As GTPases Rag também possuem um papel importante na atividade das biguanidinas metformina e de seu análogo mais hidrofóbico, a fenformina. Estes fármacos, que são amplamente utilizados no tratamento de diabetes tipo dois, não possuíam um mecanismo muito bem elucidado. Recentemente, [22], foi demonstrado que estes fármacos inibem a mTOR, diferentemente do que se acreditava, de forma independente de TSC1/2 e AMPK. Na realidade esta inibição ocorre através da inibição da função das Rags.

A maneira exata pela qual os aminoácidos transmitem o sinal de ativação às Rags ainda não é bem compreendida. O sensor inicial parece ser um trocador de Glutamina por Leucina presente na membrana celular, denominado SLC7A5/SLC3A2. Esse transportador promove a troca de Glutamina intracelular por Leucina extracelular, e a perda de sua função inibe o crescimento celular e ativa a autofagia [23] (**Figura 2**).



**Figura 2 – Vias de Regulação do Complexo mTORC1.** Representação esquemática das vias de regulação do complexo mTORC1, destacando a via que responde à presença de aminoácidos. Modificado de [24].

## Capítulo 2

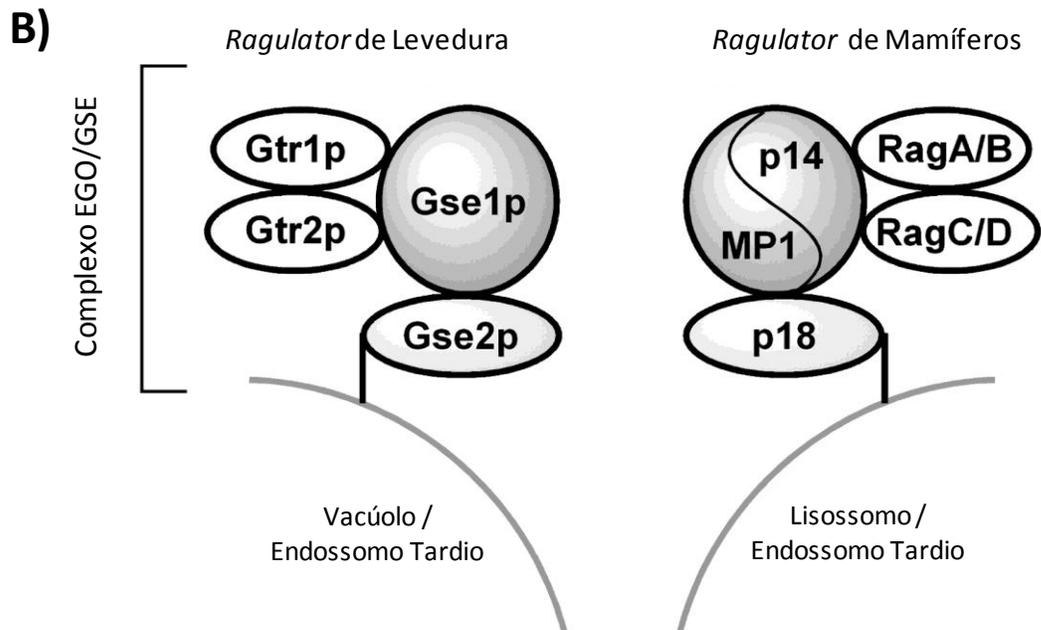
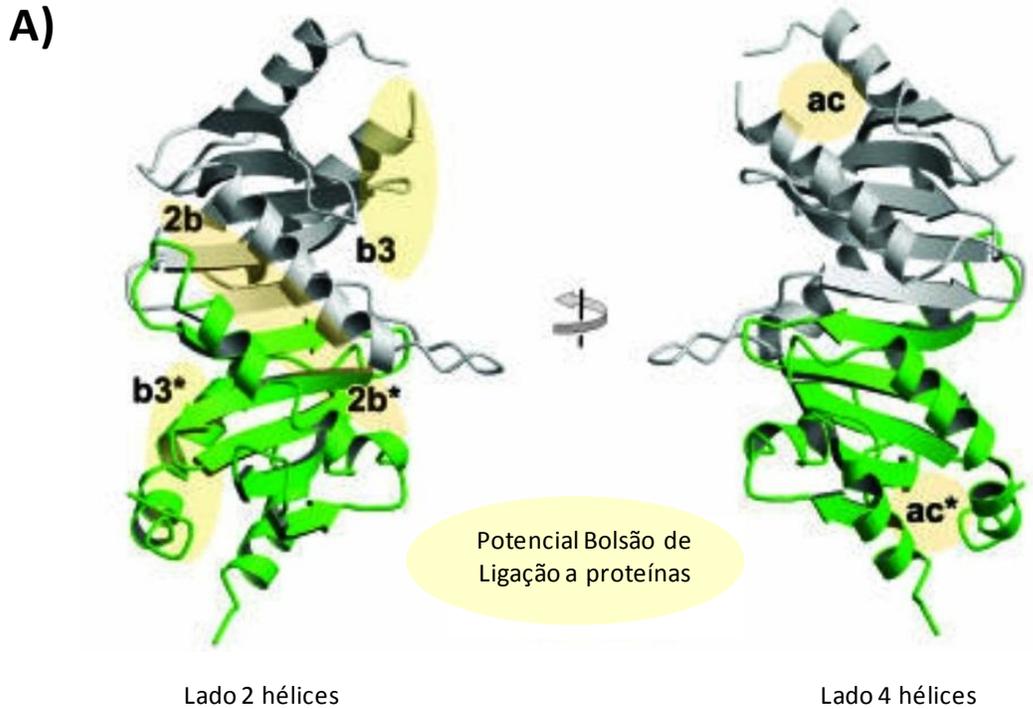
Uma possível conexão entre sistemas de transporte de aminoácidos e a ativação de TOR mediada por GTPases Rag é sugerida por estudos em levedura. Gtr1p e Gtr2p, ortólogas de RagA/B e RagC/D, respectivamente, ativam a quinase TOR na presença de aminoácidos, de maneira semelhante ao que foi descrito em células humanas [25]. Essas GTPases são encontradas na membrana do vacúolo, que tem função semelhante ao lisossomo em mamíferos, onde são ancoradas pela interação com um complexo de três proteínas formando um complexo maior denominado EGO. O complexo EGO-GSE de *S. cerevisiae*, formado por Gse1p, Gse2p, Gtr1p e Gtr2p, está relacionado ao influxo de aminoácidos e a ativação da via TOR por aminoácidos. As proteínas Gse1p e Gse2p não possuem homólogos conhecidos em mamíferos, quando comparados por sua estrutura primária. Gse2p é associada à membrana via N-miristilação e palmitoilação e, Gse1p pode se ligar a inositol fosfato, o que sugere afinidade e especificidade por membranas. Gse1p e Gse2p são, portanto, mediadores da ligação do resto do complexo EGO-GSE às membranas de levedura [26]. A estrutura cristalográfica da proteína Gse1p [26], revelou uma notável semelhança com a estrutura do heterodímero MP1/p14, apesar do complexo *Ragulator* não apresentar similaridade de sequência com os componentes do complexo EGO de levedura [17].

A proteína Vam6 de levedura, já conhecida por sua participação na formação do vacúolo [27], foi identificada como o fator de troca de nucleotídeos de guanina (GEF – *Guanine Exchange Factor*) das proteínas Gtr1p e Gtr2p [25]. Estudos anteriores haviam demonstrado um papel do complexo EGO na regulação do transportador de aminoácidos GAP1 e de TORC1 em resposta a Glutamina na superfície do vacúolo [28]. A proteína Vam6 apresenta dois homólogos em mamíferos ainda pouco caracterizados, HsVam6 e TRAP-1. A superexpressão de HsVam6 promove a agregação e fusão de lisossomos em estruturas justanucleares [29], o que sugere que essa proteína tem função semelhante à da sua ortóloga em levedura.

## Capítulo 2

### *Estrutura das proteínas do complexo Ragulator*

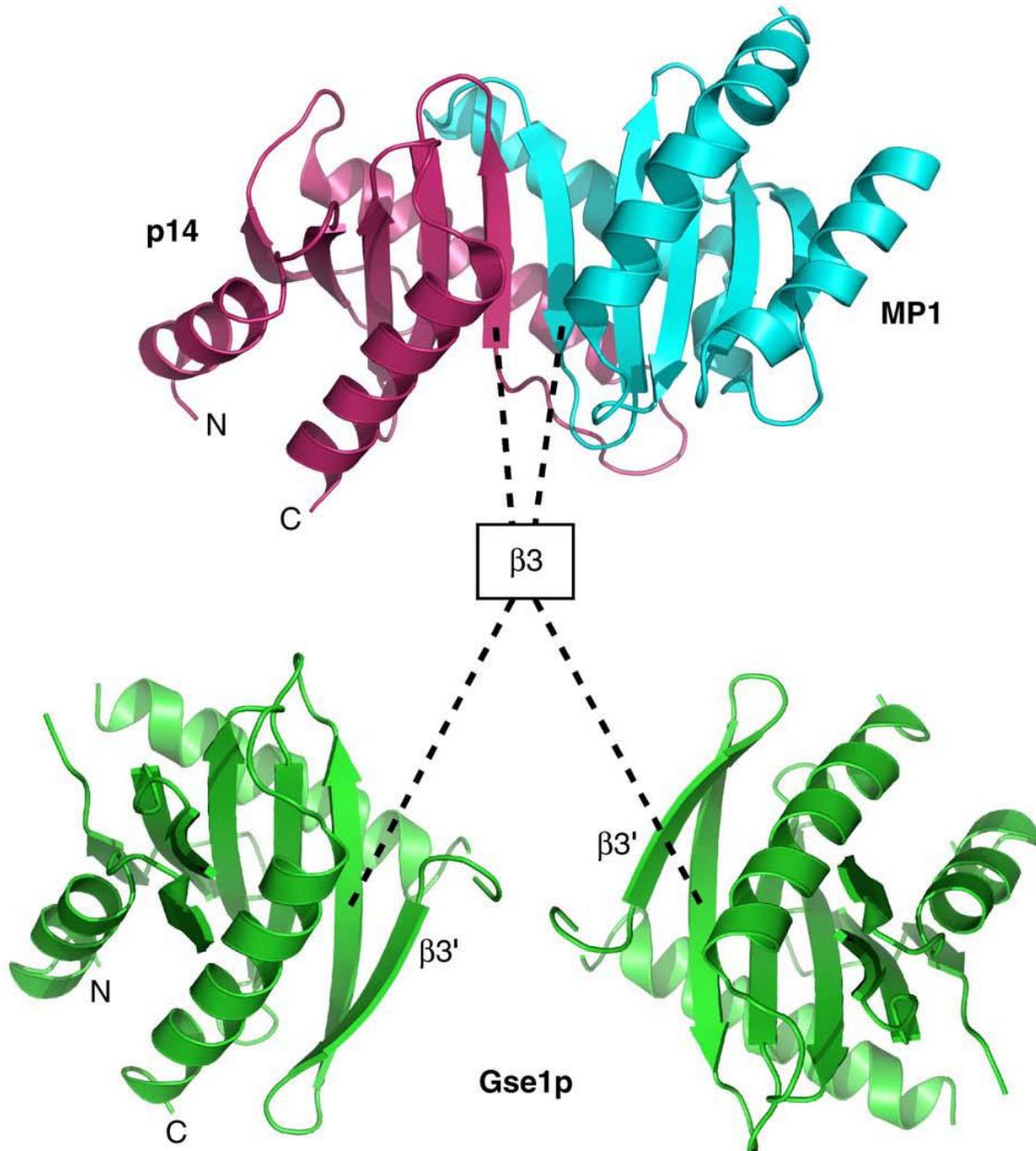
O complexo *Ragulator* é formado pelas proteínas p18, MP1 e p14, das quais apenas p18 não tem estrutura resolvida. MP1 e p14 formam um heterodímero estável, descrito anteriormente como um adaptador e ativador para a sinalização mediada por MEK1, localizado na superfície dos endossomos tardios [30];[31];[32]. Apesar de haver pouca similaridade de sequência entre essas duas proteínas, elas apresentam um enovelamento muito semelhante, sendo formadas por uma folha- $\beta$  composta de cinco fitas- $\beta$  e cercada por três  $\alpha$ -hélices, sendo uma de um lado da folha- $\beta$  e duas do outro lado. Na estrutura do dímero, as faces com duas alfa hélices se localizam do mesmo lado, e as folhas- $\beta$  se unem formando uma estrutura central composta de dez fitas [30];[31];[33] (**Figura 3**).



**Figura 3 – Estrutura do heterodímero MP1/p14.** **A.** Representação da estrutura do heterodímero de MP1 (verde) com p14 (cinza) – (PDB 1VEU 1VET, respectivamente), mostrando potenciais sítios de interação com outras proteínas (em amarelo claro). As imagens estão rotacionadas 180° ao longo do eixo vertical. Modificado de [30]; **B.** Esquema do complexo EGO-GSE e seu correspondente em mamíferos, *Ragulator* [26].

## Capítulo 2

Embora não haja similaridade alguma de estrutura primária entre Gse1p, MP1 e p14, a definição da estrutura da proteína de levedura Gse1p apresentou uma grande semelhança com o heterodímero humano MP1/ p14 (**Figura 4**).



**Figura 4 – Comparação entre as estruturas de Gse1p e o heterodímero MP1/p14.** Modelo Gse1p avaliado por *MolProbity* [34] e sobreposto a MP1/p14 (PDB ISKO). A fita  $\beta 3'$  pode interferir na dimerização de maneira semelhante ao observado no heterodímero MP1/p14. Em magenta e azul está representado o dímero MP1/p14 mostrando a folha  $\beta$  entre as subunidades do heterodímero. A fita  $\beta 3$  de cada subunidade participa da

## Capítulo 2

interface do dímero. Em Gse1p (verde), a fita  $\beta 3$  parece bloquear a folha  $\beta$  impedindo a formação de uma grande estrutura pseudocontínua oligomérica. Modificado de [26].

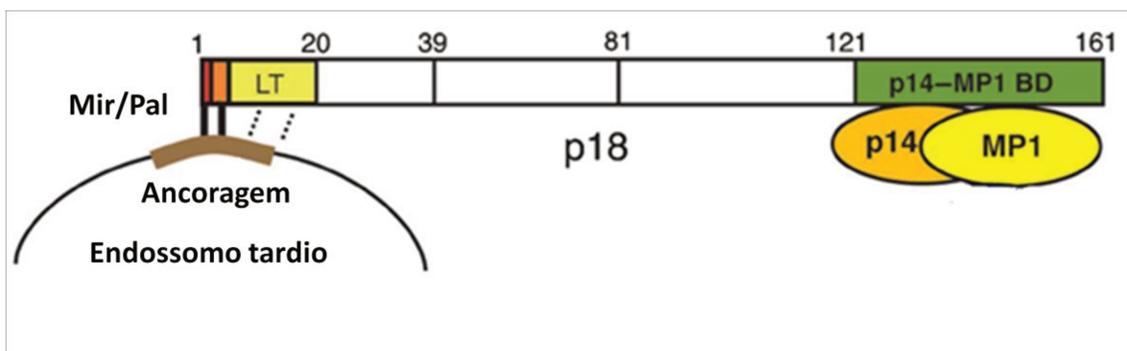
Estudos com mutantes demonstraram que o *loop* b3 de p14 é o motivo de ancoragem aos endossomos, e o *loop* correspondente na proteína MP1, denominado b3\*, define a interação de MP1 com p14 [30]. Posteriormente, foi demonstrado que a interação com p18 é essencial para a ancoragem do heterodímero MP1/p14 aos endossomos, esta interação ocorre possivelmente no *loop*. A proteína p18 possivelmente interage com o *loop* b3 de p14 por esta região ter sido previamente identificada como essencial para que o heterodímero MP1/p14 se ancore à membrana [32].

A proteína p18, assim denominada por possuir 18 kDa, foi identificada através de uma triagem de proteínas que interagem com membranas endossomais [32]. Esta proteína estava associada a um complexo, que por espectrometria de Massas, foi identificado como o heterodímero MP1/p14, e em outro trabalho [17] observou-se que p18 interage também com a membrana lisossomal. A associação de p18 com a membrana lisossomal ocorre por meio de modificações pós-traducionais (miristilação e palmitoilação) localizadas na sua extremidade N-terminal, e sua região C-terminal (resíduos 121 a 161, **Figura 5**) é responsável pela interação com MP1-p14 [32]. Em vista desses dados, é possível que o *loop* b3 de p14 seja responsável pela interação com p18.

Estudos recentes mostraram que a proteína p18 heteróloga contendo os aminoácidos de 47 a 104 se ligou às suas proteínas parceiras MP1/p14 [35]. Esta informação, somada ao fato de que estudos onde mutantes de p18 com deleções nos 40 resíduos de aminoácidos em sua região C não se ligaram a MP1/p14 [21], pode ser útil para identificação de uma área ainda mais restrita da região de interação de p18 com MP1/p14.

Pdro, outra nomenclatura utilizada para a proteína p18, está envolvida na homeostase do colesterol [36]. A depleção desta proteína por siRNA aumenta o teor de colesterol livre nas células, que é acompanhado pelo aumento do efluxo de colesterol das células. Por outro lado, células estáveis, ao superexpressar p18 mostram reduzido teor de colesterol livre celular. P18 diminui o excesso de colesterol, pelo menos em parte, por estimular a recaptação da lipoproteína de baixa densidade (LDL) e aumentar o egresso de colesterol dos endossomos tardios/lisossomos [36].

## Capítulo 2



**Figura 5 – Esquema de interação da proteína p18 com seus ligantes.** P18 interage na sua região N terminal com a membrana lisossomal através dos ácidos graxos miristoil e palmitoil. Ela também interage, em sua região C terminal, com o MP1/p14. Modificada de [32].

A proteína de levedura Gse2p (20 kDa) possui tamanho molecular semelhante ao da proteína p18 (18 kDa) e também sofre miristilação e palmitoilação na região N terminal. Ela apresenta um motivo na região N terminal,  $MGX_{0-4}CX_{0-2}C$ , com uma significativa identidade de sequência com a proteína p18. Outra importante semelhança entre a proteína p18 e Gse2p é que ambas possuem predição de estrutura secundária predominantemente em  $\alpha$  hélice e uma região extensa que poderia sugerir uma forma desenovelada quando na forma nativa, e possivelmente a formação do complexo estabilizaria suas estruturas (**Figura 6**). Estudos estruturais com estas duas proteínas devem ser realizados para verificar se são estrutural e funcionalmente homólogas [26].

## Capítulo 2



**Figura 6 – Comparação entre a Gse2p (código de acesso ao UniProt Q02205) e a proteína humana p18 (código de acesso ao UniProt Q61AA8).** Embora essas proteínas não possuam similaridade em sua sequência primária, ambas possuem predição de estrutura secundária rica em  $\alpha$  hélice e possuem grandes regiões que podem ser desenoveladas em sua forma nativa, de acordo com IUPRED [37] (não mostrado). Gse2p e p18 possuem tamanhos moleculares similares (aproximadamente 20 kDa), ambas contem resíduos de glicina na porção N terminal que são potenciais sítios de miristilação (setas negras), e ambas contêm potenciais sítios de palmitoilação (setas amarelas). Além disso, estas duas proteínas, assim como suas homólogas, possuem um motivo de seis aminoácidos (sublinhado) no início da sequência, que contém um ou mais resíduos ácidos seguidos por duas Leucinas conservadas, o que sugere uma sequência de localização lisossomal. Modificado de [26].

Neste trabalho, expressamos o complexo *Ragulator* com a finalidade de realizar estudos estruturais por Ressonância Magnética Nuclear e, a partir destes estudos aprofundarmos o conhecimento sobre a ativação da via da mTOR por aminoácidos. A análise estrutural pode nos auxiliar a compreender como ocorre a transdução de sinal da via de mTOR em resposta a aminoácidos.

## Capítulo 2

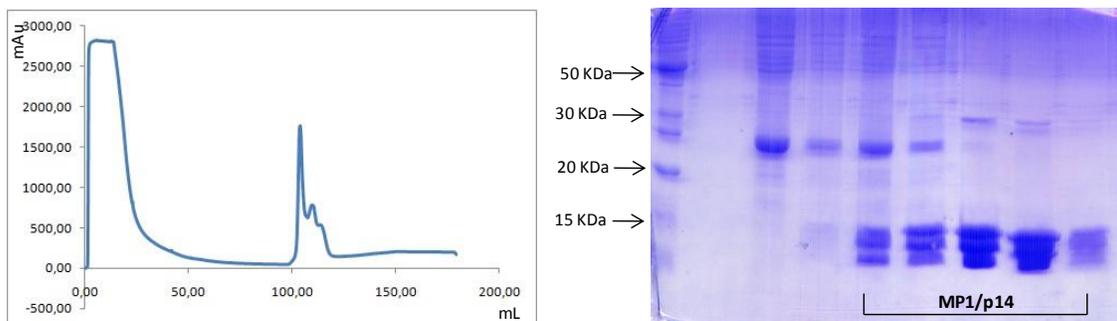
### 2.2. RESULTADOS OBTIDOS NA CARACTERIZAÇÃO DA PROTEÍNA P18

Apesar da interação entre a p18 e o heterodímero MP1/p14 já ter sido confirmada [32], os resultados a seguir foram norteados com a finalidade de mapearmos a interação, bem como a dinâmica envolvida entre as proteínas do complexo *Ragulator*.

#### 2.2.1. Expressão e purificação dos clones de MP1/p14

O clone de escolha para expressão foi o 688 [30]. Este clone foi enviado pelo grupo do pesquisador Dr. Tim Clausen (*Reserch Institute of Molecular Pathology, IMP, Viena, Áustria*) para a co-expressão das proteínas MP1/p14. Ele possui ambas as proteínas contidas no vetor pET28a (Novagen), e cauda de Histidina ligada a proteína MP1. Este foi um critério de escolha do clone, pois se a cauda de Histidina fosse na proteína p14 poderia atrapalhar a interação com p18, uma vez que imaginamos que a ligação entre p18 e MP1/p14 ocorre no *loop*  $\beta$ 3 da proteína p14. O heterodímero foi expresso em LB e utilizado para adição em diversas diálises da p18, auxiliando em seu processo de renaturação.

Na cromatografia por afinidade a Níquel, cujos detalhes estão citados na seção Metodologia, obtivemos o heterodímero MP1/p14 quase puro, como mostra a **Figura 7**.

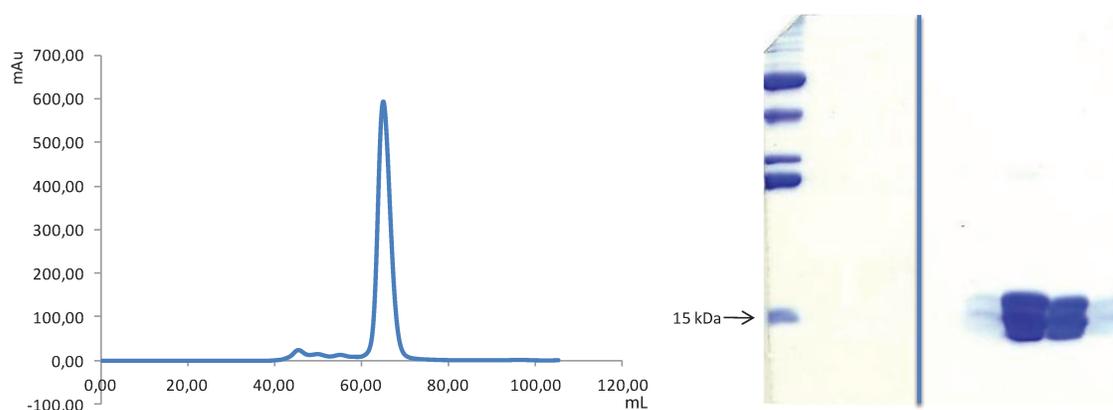


**Figura 7 - Cromatograma de Afinidade do heterodímero MP1/p14.** O heterodímero protéico foi submetido a uma cromatografia por afinidade a Níquel. O gel SDS-PAGE contém as frações referentes ao pico do cromatograma, onde as proteínas MP1 e p14 saem separadas (pois se trata de um gel desnaturante) na região do marcador molecular da *BenchMark™* (primeira amostra) correspondente a banda de 15 kDa. Coluna: Histrap FF 5 mL. Gradiente linear: 0 a 100% de B em 10 volumes de coluna. Tampão A: 20 mM

## Capítulo 2

Tris-HCl pH 6,5, 100 mM NaCl. Tampão B: Tampão A acrescido de 1 M de Imidazol. Peso Molecular das proteínas MP1 e p14: aproximadamente 15 kDa cada.

O heterodímero ainda foi submetido a uma cromatografia de exclusão por tamanho molecular (coluna Superdex HiLoad 200 16/60 – GE Healthcare), utilizando o mesmo tampão de carregamento de amostra da afinidade. Esta etapa permitiu a obtenção de MP1/p14 com alto grau de pureza. Apesar de não se tratar de uma filtração em gel analítica, o volume de eluição de MP1/p14 (60 mL) pode indicar que o heterodímero esteja com um tamanho molecular maior que o esperado (**Figura 8**).



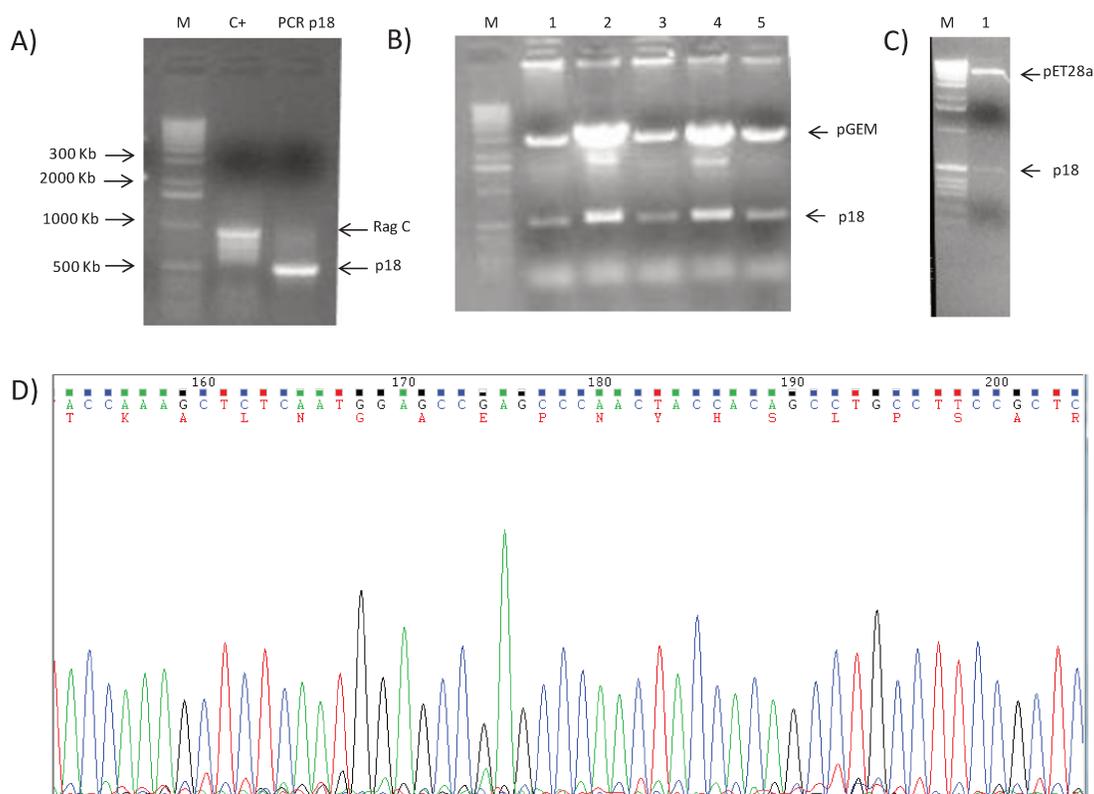
**Figura 8 - Cromatograma de exclusão por tamanho molecular do heterodímero MP1/p14.** A amostra de MP1/p14 purificada por afinidade foi concentrada e submetida a uma filtração em gel onde pudemos obter MP1/p14 com alto grau de pureza. O gel desnaturante SDS-PAGE mostra a banda de ambas as proteínas com aproximadamente 15 kDa, presentes nas frações correspondentes ao pico. Coluna: Superdex 200 16/60 (GE Healthcare®). Corrida: 1 volume de coluna. Tampão: 20 mM de tampão Tris-HCl pH 6,5 e 100 mM NaCl. Peso molecular do heterodímero: aproximadamente 30 kDa.

### 2.2.2. Estratégias de clonagem da p18

A amplificação do gene p18 foi realizada por PCR, sendo confirmada por eletroforese em gel de agarose e conferida pelo peso molecular (**Figura 9A**). A eletroforese mostrou que a amplificação da p18 ocorreu especificamente, observando-se a presença de uma única banda, que possibilitou a purificação do inserto direto do produto da PCR com o *Kit Qiaquick PCR purification*, da empresa *Qiagen*.

## Capítulo 2

A ligação do inserto com o vetor pGEM foi verificada clivando-se o DNA plasmidial de algumas colônias de bactérias com as enzimas *NdeI* e *XhoI*, que clivaram a região onde estava o inserto, que pode ser visualizada por eletroforese em gel de agarose (**Figura 9B**). Em seguida, esses DNAs plasmidiais foram enviados para o sequenciamento (**Figura 9D**) e depois foram analisadas com a ferramenta BLAST ([www.ncbi.nlm.nih.gov/blast/Blast.cgi](http://www.ncbi.nlm.nih.gov/blast/Blast.cgi)).

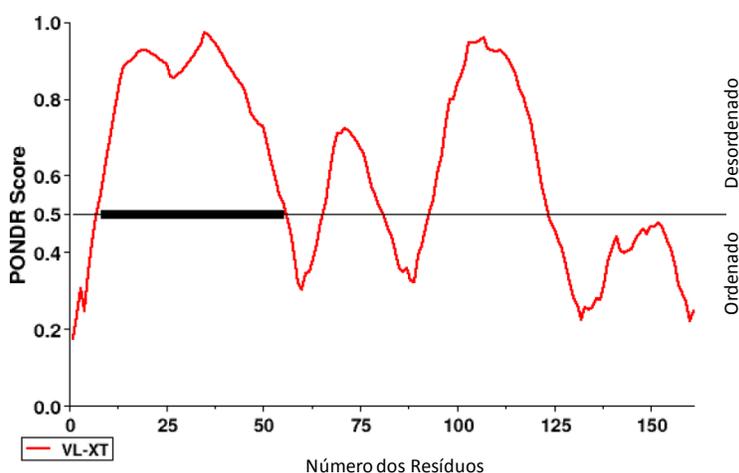


**Figura 9 - Clonagem da proteína p18.** **A)** Confirmação da reação de PCR da p18. Marcador Molecular 1 kb da Invitrogen® (M), Controle positivo RagC (C+) e amostra da reação de PCR da p18. **B)** A miniprep realizada com cinco clones de p18 no vetor pGEM, foi digerida com as enzimas de restrição *NdeI* e *XhoI* para checagem da clonagem **C)** Digestão de checagem da subclonagem de p18 em pET28a. Marcador Molecular 1 kb da Invitrogen® e digestão da miniprep do clone de p18. **D)** Fragmento do eletroferograma do sequenciamento da p18. Os picos de leitura altos e uniformes demonstram que a qualidade da amostra está boa e a leitura é confiável.

## Capítulo 2

### 2.2.3. Predição de Desordem da p18

O programa PONDR [38-40] foi utilizado para realizarmos a predição de desordem da proteína p18 (**Figura 10**). Segundo esta predição, a proteína p18 apresenta 3 regiões de desordem, sendo que uma destas regiões é formada por 48 aminoácidos. Dos 161 resíduos de aminoácidos constituintes desta proteína, 94 se encontram em regiões desordenadas, representando um percentual de desordem de aproximadamente 59%. Esta predição de desordem associada a presença de 2 resíduos de Cisteínas na p18 pode explicar a instabilidade e a facilidade com que esta proteína forme agregados.



**Figura 10 – Predição de desordem da proteína p18.** A predição foi realizada segundo algoritmo do PONDR [38-40]. Segundo esta predição, p18 apresenta 3 regiões de desordem representando 59% dos resíduos de aminoácidos.

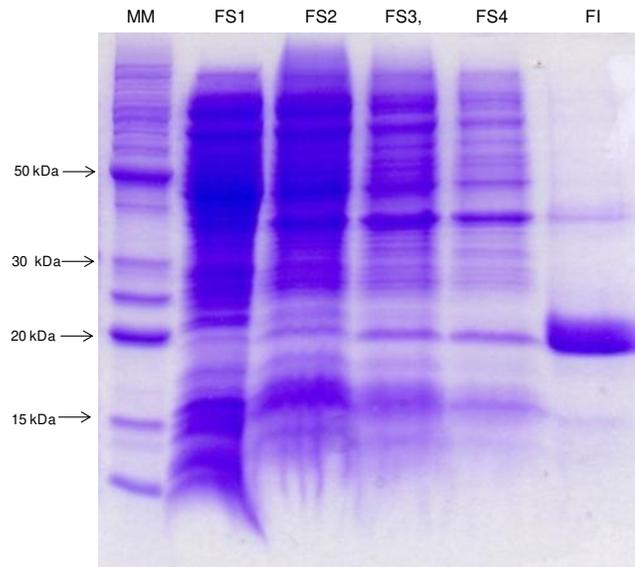
### 2.2.4. Expressão da proteína de mamífero p18 em bactéria

Foi realizado um teste para verificar a melhor forma de expressão da proteína de mamífero p18 (**Tabela 4**). Neste teste, foram variadas a cepa de expressão, a temperatura e o tempo de indução, entretanto o vetor (pET28a), a velocidade de incubação do inóculo (200 rpm) e a concentração de IPTG foram mantidos constantes.

Em todas as condições testadas, a proteína foi expressa na fração insolúvel, em corpos de inclusão. Devido ao alto grau de rendimento e pureza desta proteína, escolhemos

## Capítulo 2

a condição de expressão em que *E. coli* (DE3)  $\Delta$ SlyD foi expressa a 37°C a 200 rpm por 16h. A **Figura 11** demonstra a presença da proteína exclusivamente na fração insolúvel.



**Figura 11. Fração insolúvel da p18.** A proteína p18, com aproximadamente 20 kDa, foi induzida com 0,5 mM de IPTG a 37°C, 200 rpm por 16 h. No gel SDS-PAGE corado com *Comassie Blue*, MM corresponde ao marcador molecular *BenchMark™*, e FS correspondem às frações solúveis, pois para chegar a fração insolúvel (FI) são necessários 4 ciclos de sonicação.

### 2.2.5. Padronização das condições de diálise da p18

A ressolubilização de uma proteína por diálise muitas vezes resulta em agregados muito recalcitrantes. Na tentativa de obter p18 monomérica a partir da diálise, resolvemos estudar melhor a composição do tampão, a fim de obter uma proteína estável, solúvel e que não forme agregados.

Proteínas ricas em cisteínas, como a p18 (que possui 2 Cisteínas), tendem a formar agregados por pontes dissulfeto, o que induz a polidispersidade na amostra. Agentes redutores devem ser usados preventivamente, pois a formação de pontes dissulfeto dificilmente é revertida sem causar danos à proteína. Os agentes redutores mais utilizados são o  $\beta$ -ME e o DTT, sendo este último uma melhor escolha por apresentar apenas um pico no espectro de RMN.

## Capítulo 2

As proteínas apresentam uma solubilidade em função da concentração de sais. Isso porque, muitas proteínas insolúveis em água tornam-se cada vez mais solúveis conforme se aumenta a concentração de sal a certas concentrações limitantes, que dependem da proteína e do tipo de sal escolhido (*salting in*). Os íons adicionados (positivos e negativos), presentes em solução, interagem com os grupos carregados das moléculas de proteínas, atenuando a interação entre elas. Quando a concentração de sal atinge valores muito elevados, a solubilidade desta proteína diminui, até sua precipitação (*salting out*) [41].

Os agregados protéicos ainda podem ser formados por interações hidrofóbicas entre duas ou mais proteínas. O glicerol e alguns detergentes, como Triton-X 100 e Tween 20, podem ser utilizados para evitar que isso aconteça.

Com base nas interações existentes entre os diferentes tipos de aminoácidos constituintes em uma proteína, elaboramos algumas condições de tampão que poderiam evitar a formação de agregados na renaturação da p18. Nas condições de 1 a 6 citadas na **Tabela 1**, a proteína p18 se manteve agregada após a ressolubilização. Estas proteínas foram submetidas a cromatografias de afinidade e exclusão por tamanho molecular, além de estudos biofísicos para sua análise (serão descritos abaixo). Após estas análises, pudemos elaborar uma nova condição (7 da **Tabela 1**) onde p18 se ligou à coluna de afinidade, podendo estar em sua forma monomérica. Para isso devem ser realizados experimentos para confirmar seu estado monomérico.

**Tabela 1** – Condições de tampão variadas na diálise da p18 a 4°C.

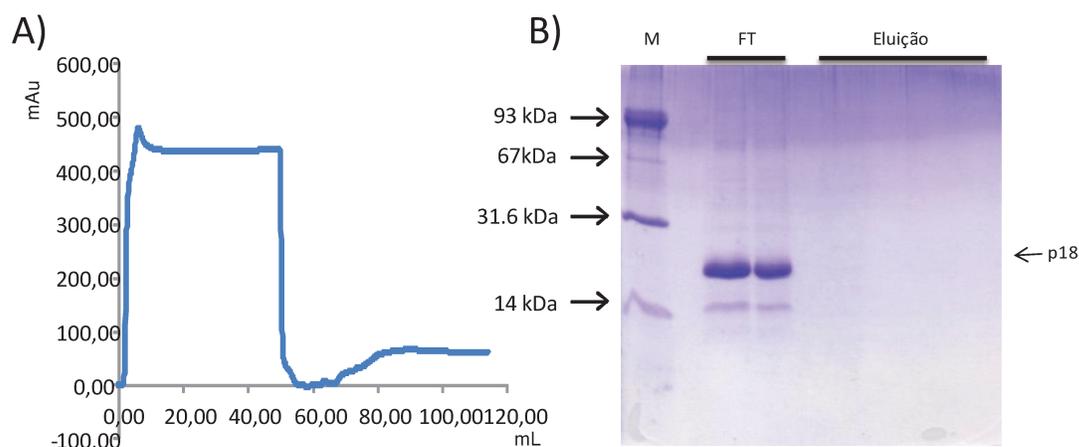
<i>Condição</i>	<i>Tampão utilizado na diálise primeira etapa da diálise</i>	<i>Condição da p18</i>
1	Uréia 6M, 50mM e Tris-HCl pH 8,8	Agregado Solúvel
2	Uréia 6M, 50mM Tris-HCl pH 8,8 e 50mM NaCl	Agregado Solúvel
3	Uréia 6M, 50mM Tris-HCl pH 8,8 e 150mM NaCl	Agregado Solúvel
4	Uréia 6M, 50mM Tris-HCl pH 8,8 e 0,1M $\beta$ ME	Agregado Solúvel
5	Uréia 6M, 50mM Tris-HCl pH 8,8, 150mM NaCl, 5mM $\beta$ ME, 0,1% Tween 20 e 5% glicerol	Agregado Solúvel
6	Uréia 6M, 50mM Tris-HCl pH 8,8, 50mM Arg/Glu e 150mM NaCl	Agregado Solúvel
7	Uréia 6M, 50mM Tris-HCl pH 8,8, 50mM NaCl e 0,5 mM TCEP**	p18 complexada com MP1/p14

\* Condição à qual foi adicionado MP1/p14 na última etapa de diálise, sem uréia. \*\* Condição à qual foi adicionado MP1/p14 com 1 M de uréia.

## Capítulo 2

### 2.2.6. Padronização da purificação da proteína p18 renaturada com uréia

A proteína p18, dialisada nas condições de 1 a 6 (Tabela 1), foi submetida a cromatografia por afinidade a Ni (Figura 12). Esta proteína, embora possua cauda de Histidina, não se ligou a coluna, provavelmente por apresentar um estado de agregação de modo a formar uma conformação com um núcleo hidrofóbico, onde a cauda de Histidina não estaria exposta.

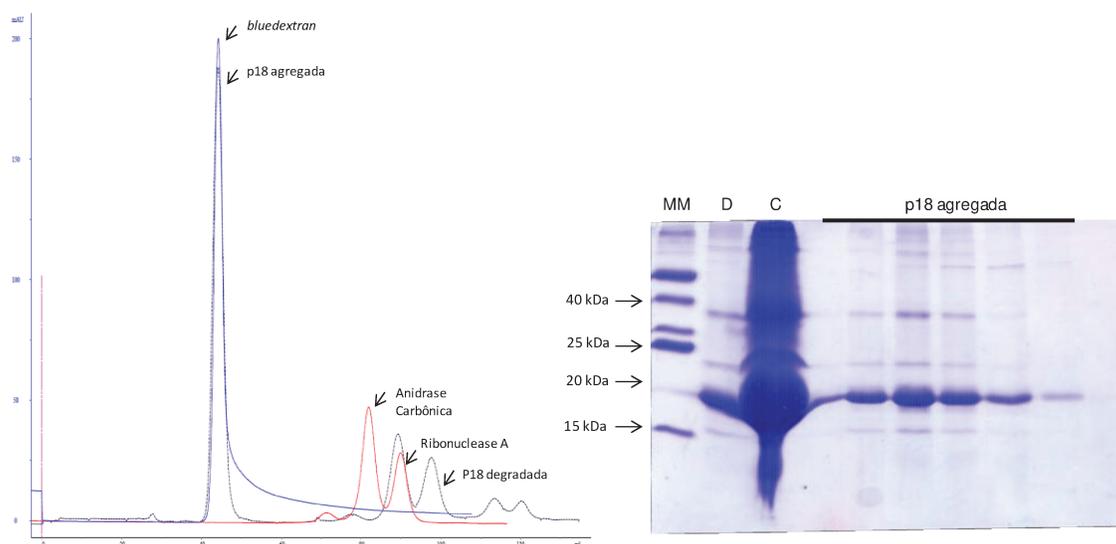


**Figura 12 – Cromatograma de afinidade da proteína p18 dialisada nas condições de 1 a 6 da Tabela 1.** A) Cromatograma de afinidade da proteína p18 dialisada onde a proteína não se ligou na coluna. B) Gel SDS-PAGE corado com *Comassie Blue*, referente à cromatografia corado A, contendo um marcador molecular, a proteína p18 que passou direto na coluna (*flowthrough*) e as amostras eluidas que não contém a proteína. Coluna: Histrap FF 5 mL. Gradiente linear: de 0 a 100% de B em 10 volumes de coluna. Fluxo: 1 mL/min.

A proteína recuperada da afinidade foi submetida a uma cromatografia de exclusão por tamanho molecular na coluna Superdex 200 16/60 (GE-Healthcare®). A fim de se investigar o tamanho do agregado em que p18 se encontrava, foram acrescentados ainda alguns marcadores de tamanho molecular, como o *blue dextran* (padrão que passa direto pela coluna, possui um tamanho maior do que a proteína é capaz de separar), Ribonuclease A (padrão para proteínas com 13 kDa) e Anidrase Carbônica (padrão para proteínas 29 kDa). P18 apresentou um tamanho bem maior que o esperado, uma vez que esta saiu junto com o *blue dextran*, sugerindo a formação de grandes agregados, além de pequenos picos

## Capítulo 2

de degradação indicando relativa instabilidade. O cromatograma da p18 também apresentou alguns picos próximos às proteínas padrões que possuem pesos moleculares próximos a 20 kDa (**Figura 13A**), porém através de um gel SDS-PAGE constatamos que se trata de bandas de degradação, não mostradas no gel da **Figura 13B**.



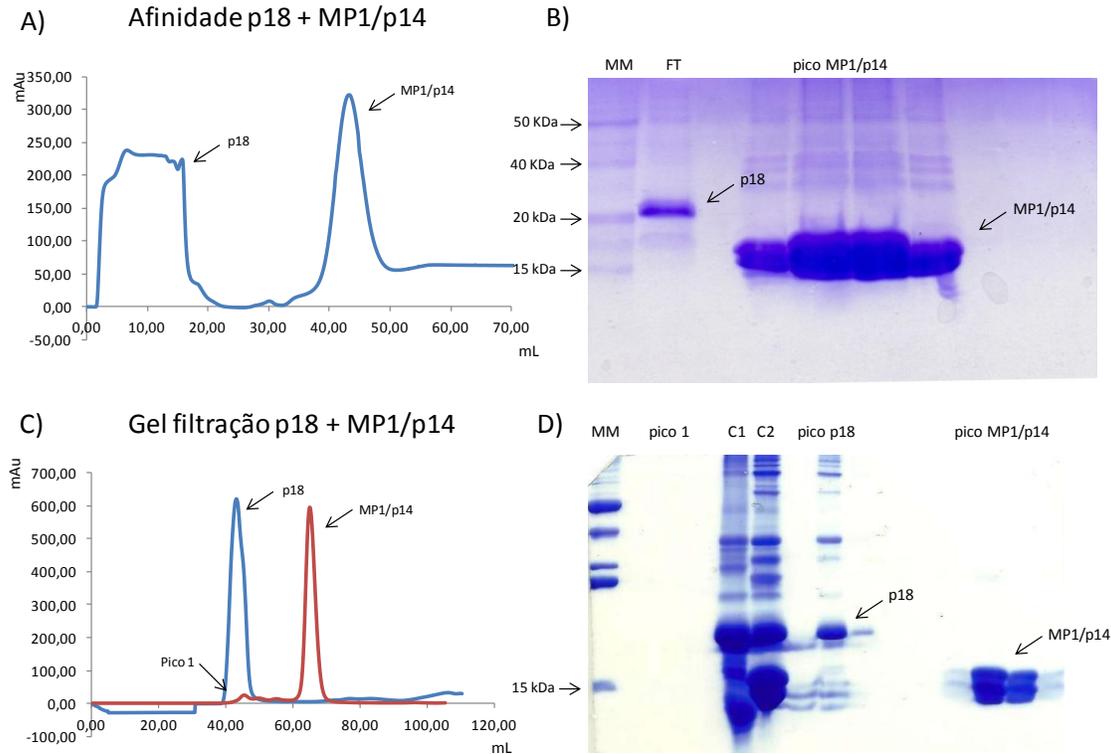
**Figura 13. Cromatograma da exclusão por tamanho molecular da p18. A)** Sobreposição do cromatograma da cromatografia de exclusão por tamanho molecular da proteína p18 com o cromatograma dos padrões *blue dextran*, Anidrase Carbônica e Ribonuclease A. Podem-se observar picos da p18 em regiões de volume morto da coluna e em regiões próximas aos do tamanho esperado (aproximadamente 20 kDa). **B)** Gel SDS-PAGE corado com *Comassie Blue*, onde MM corresponde ao marcador molecular *BenchMark<sup>TM</sup>*, D a proteína pós diálise, C a proteína concentrada para filtração em gel e as frações sinalizadas como p18 agregada correspondem ao pico de p18 que saem com volume equivalente ao *Comassie Blue*. Coluna: Superdex 200 16/60 (GE-Healthcare<sup>®</sup>). Corrida: 1.5 volume de coluna. Tampão: 20 mM Tris-HCl pH 6,5, 100 mM NaCl.

Com o intuito de formar o complexo *Ragulator*, as proteínas MP1/p14 foram adicionadas à última etapa da diálise da p18, quando praticamente não havia mais uréia. O fato da p18 se encontrar agregada possivelmente a impossibilitou de ligar às suas parceiras, MP1/p14.

Em uma nova tentativa de unir estas proteínas, a p18 foi incubada com o heterodímero MP1/14 e esta amostra foi submetida à cromatografia de exclusão por tamanho molecular (Superdex 200 16/60 – GE Healthcare<sup>®</sup>). As frações referentes a esta filtração em gel foram avaliadas em um gel SDS PAGE e, em nenhum pico foram

## Capítulo 2

observadas ambas as proteínas, o que sugeriria a formação do complexo (**Figura 14**). Mais uma vez, p18 não se ligou a seus parceiros por estar agregada (seu pico saiu em aproximadamente 40 mL, indicando um valor muito maior que o tamanho molecular teórico) e, portanto se faz necessário encontrar uma condição específica em que possamos obter p18 em sua forma monomérica e só então tentarmos associá-la a MP1/p14.



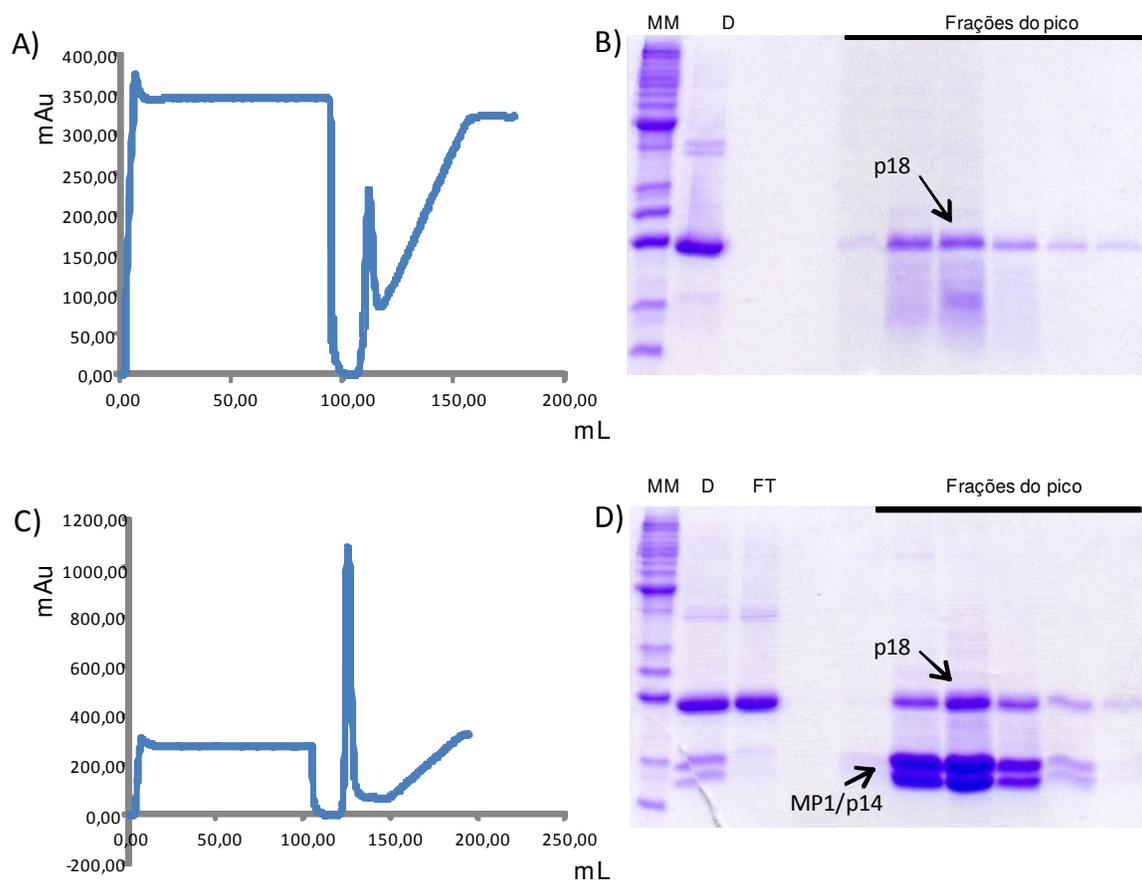
**Figura 14. Teste de interação do complexo *Ragulator*.** **A)** Cromatograma de afinidade por Níquel das proteínas P18+MP1/p14, onde apenas o heterodímero MP1/p14 se ligou à coluna. **B)** Gel SDS-PAGE corado com *Comassie Blue* contendo o marcador molecular *BenchMarker*<sup>®</sup> a proteína p18 que não se ligou à coluna (FT, passou no *flowthrough*) e as frações do pico contendo MP1/p14. **C)** Cromatograma de filtração em gel das proteínas p18 (~20 kDa) e MP1/p14 (~30 kDa) saíram em picos separados. **D)** Gel SDS-PAGE corado com *Comassie Blue* contendo marcador molecular, frações do primeiro pico, C1 e C2 são controles contendo p18 e p18 + MP1/14.

Uma última diálise foi testada com o tampão 7 da **Tabela 1**. A esta diálise, foram adicionadas as proteínas MP1/p14 quando ainda havia 1 M de uréia. A uréia foi totalmente retirada e a proteína foi submetida às cromatografias de afinidade e filtração em gel. Em um mesmo béquer, foi realizada uma diálise controle, sob as mesmas condições, porém as proteínas MP1/p14 não foram adicionadas. Sob estas condições, foi possível obter a proteína p18 capaz de se associar a coluna de Níquel, supostamente em sua forma

## Capítulo 2

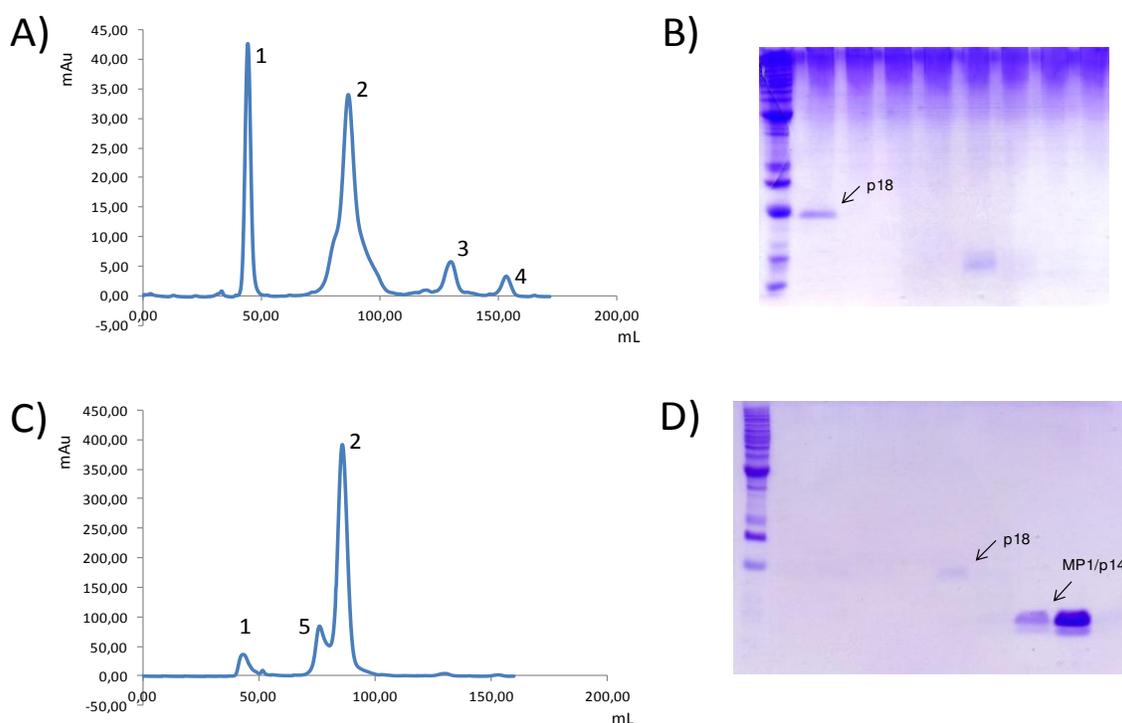
monomérica. Tanto a p18 renaturada na diálise controle, quanto a renaturada com adição das parceiras MP1/p14, se ligou à coluna de afinidade a Níquel. Embora em ambos os casos tenha se obtido p18 em sua suposta forma monomérica, o rendimento foi muito maior quando adicionado MP1/p14. Foi obtida uma grande quantidade das proteínas p18 e MP1/p14 na cromatografia de afinidade. Como tanto p18 quanto o heterodímero MP1/p14 possuem caudas de Histidina, as proteínas podem ou não estar complexadas na forma de *Ragulator*. (**Figura 15**). As frações correspondentes ao pico da afinidade, que continham p18 e MP1/p14 foram submetidas à cromatografia de exclusão por tamanho molecular. Nesta cromatografia, p18 e MP1/p14 saíram em picos separados, sugerindo dissociação do complexo, provavelmente por possuir alto coeficiente de dissociação [35]. É importante observar que p18 é eluída em menor volume que o esperado, sugerindo agregação (**Figura 16**). Enquanto realizávamos estes experimentos, um artigo foi publicado com a renaturação do complexo, também através de adição de MP1/p14 em uma etapa da diálise da p18, porém com outro tampão [35].

## Capítulo 2



**Figura 15 – Afinidade da diálise da proteína p18 sob a condição de tampão 7 da Tabela 1, com e sem adição de MP1/p14.** **A)** Cromatograma da Afinidade a Ní em Coluna His Trap FF 5 mL, da proteína p18 dialisada sem adição de MP1/p14. **B)** Gel SDS-PAGE corado com *Comassie Blue* contendo marcador molecular *BenchMark™* (MM), a proteína pós diálise (D) e as frações do pico contendo a p18 que se ligou a coluna. **C)** Cromatograma da afinidade nas mesmas condições acima citadas, porém com a amostra de diálise em que foi adicionado MP1/p14. **D)** Gel SDS-PAGE corado com *Comassie Blue* contendo MM, proteína pós diálise (D), proteína que não se ligou a coluna (FT), provavelmente por MP1/p14 possuir concentração limitante e as frações do pico correspondem a p18 provavelmente associada a MP1/14.

## Capítulo 2



**Figura 16 – Cromatografia de exclusão por tamanho molecular, em coluna Superdex 200 16/60<sup>®</sup>, da diálise da proteína p18 sob a condição de tampão 7 da Tabela 1, com e sem adição de MP1/p14. A) Cromatograma da filtração em gel da proteína p18 sem adição de MP1/p14. B) Gel SDS-PAGE contendo marcador molecular *BenchMarker*<sup>®</sup>, a fração do pico 1, em que p18 aparece em sua forma agregada e os demais picos correspondem a proteína degradada. C) Cromatograma da filtração em gel da proteína p18 com adição de MP1/p14. D) Uma menor quantidade de proteína foi obtida na forma de agregados, imperceptível no gel SDS-PAGE (pico 1), onde podemos visualizar p18 apenas no segundo pico e MP1/p14 no terceiro, sugerindo dissociação do complexo *Ragulator*.**

Uma repetição da diálise sob o tampão 7 (**Tabela 1**) com adição de MP1/p14 foi realizada. Nesta repetição, as proteínas contidas nas frações do pico da afinidade foram selecionadas para realização de *cross-linking* e não foi realizada filtração em gel.

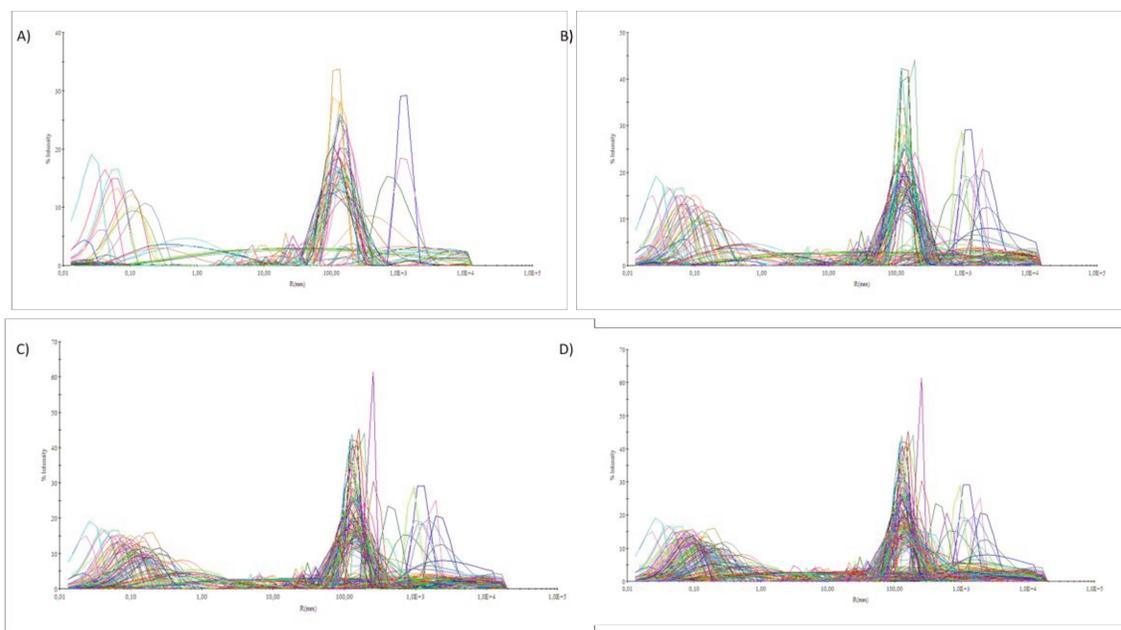
## Capítulo 2

### *2.2.7. Espalhamento de luz dinâmico – DLS*

O espalhamento de Luz Dinâmico (DLS) é uma técnica amplamente utilizada no estudo de partículas em solução, como polímeros, proteínas, vesículas lipídicas e várias outras moléculas de interesse biológico.

A proteína p18, dialisada sob o tampão 5 (**Tabela 1**) foi submetida à análise de DLS, com a finalidade de investigar as características do agregado formado. A triagem foi realizada com 96 condições diferentes (variando-se tampões e aditivos, citados na seção Metodologia) nas temperaturas 20°C, 25°C, 37°C, repetindo a medida de 25°C ao final para avaliar se houve alguma alteração na amostra após a medida a 37°C. Em todas as medidas, obtivemos um  $R$  (índice de dispersão fornecido pelo equipamento) próximo de 1, ou seja, um valor muito alto indicando uma polidispersividade muito grande. O raio hidrodinâmico foi de aproximadamente 100 nm, indicando a presença de agregados protéicos desta ordem e, portanto nenhuma das condições foi capaz de manter p18 em sua forma monomérica (**Figura 17**). Mesmo de posse destes dados de DLS, ainda se faz necessária a elucidação do que estaria ocorrendo com a p18 e a suspeita de grandes agregados nos fez optar pelo recurso da microscopia eletrônica.

## Capítulo 2



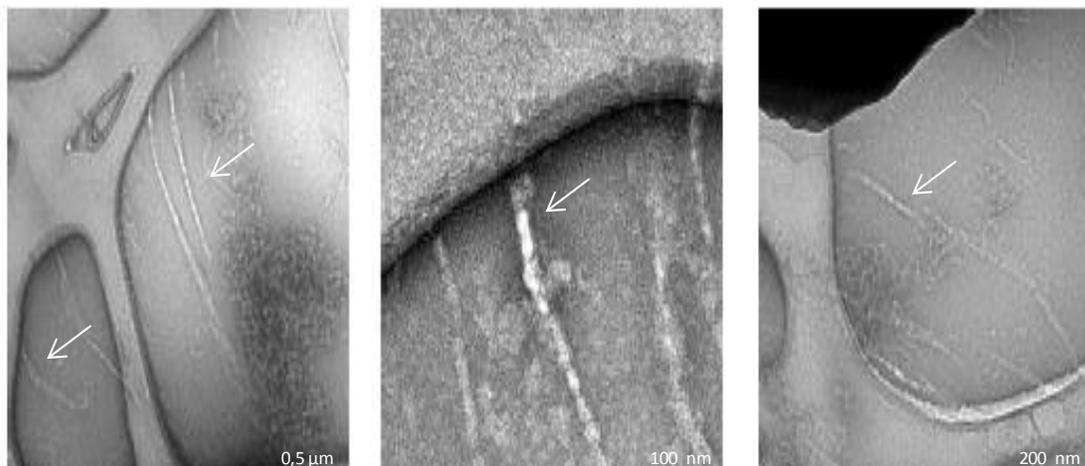
**Figura 17. Gráficos de regularização de espalhamento Dinâmico de Luz (DLS) da p18.** As figuras A, B, C e D correspondem, respectivamente as análises da p18 a 20°C, 25°C, 37°C e 25°C (após medição a 37°C). Cada linha no espectro representa uma das condições de tampão citadas no texto, em que se variou o pH de 5,5 a 9,5 e 12 diferentes de aditivos. Em todas as condições visualizamos grandes agregados da ordem de 100 nm, sugerindo a formação de fibras, que foram confirmadas posteriormente por microscopia eletrônica.

### 2.2.8. Microscopia eletrônica

A microscopia eletrônica tem sido uma ferramenta largamente utilizada para entendimento de biomoléculas e, através dela é possível determinar estruturas e analisar interações entre proteínas. Com os resultados obtidos no Espalhamento Dinâmico de Luz, resolvemos investigar os agregados protéicos formados pela p18.

Na primeira amostra utilizada na microscopia foram observados agregados fibrosos, como mostrados na **Figura 18**. A amostra foi dialisada sob o tampão 5 (**Tabela 1**) se encontrava com 50 mM Tris-HCl pH 8,8, 150 mM NaCl, 5 mM  $\beta$  ME, 0,1% Tween 20 e 5% glicerol.

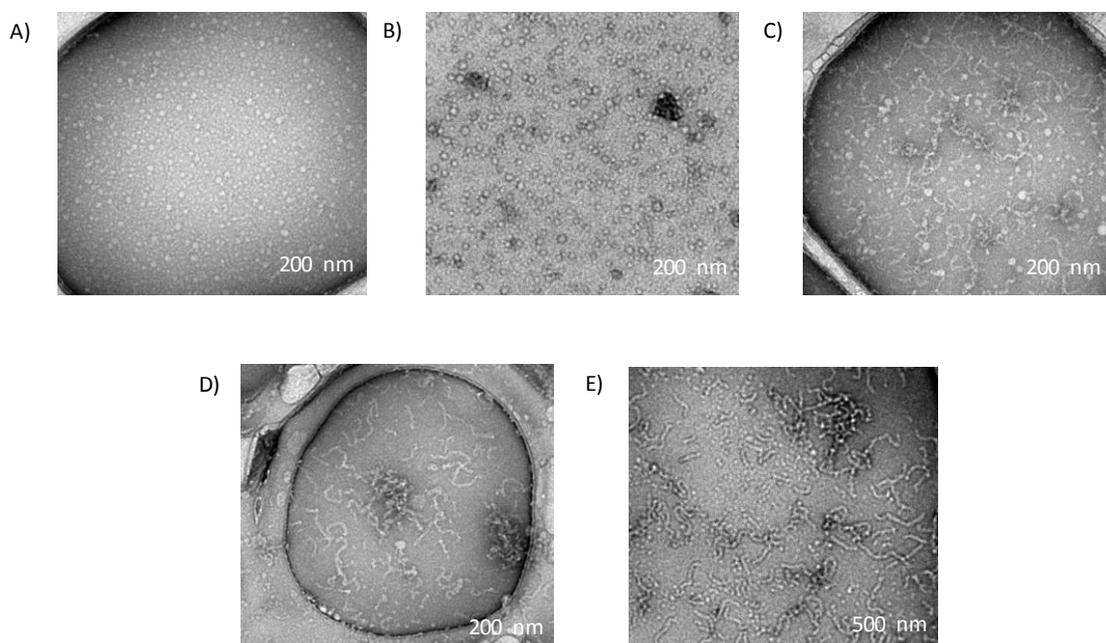
## Capítulo 2



**Figura 18. Microscopia eletrônica da proteína p18 renaturada.** A proteína p18 foi dialisada sob o tampão 5 (**Tabela 1**) e observada em um microscópio eletrônico utilizando acetato de uranila como contraste. Esta análise confirmou a presença de grandes agregados protéicos (setas brancas), principalmente na forma de longas fibras.

Foi realizada uma nova análise de microscopia eletrônica da p18 renaturada (**Figura 19**), desta vez com o tampão 2 (**Tabela 1**). Diferentemente do que esperávamos, este tampão sem aditivos, contendo apenas 50 mM Tris-HCl pH 8,8 e 50 mM NaCl (além da uréia, que foi retirada gradativamente nas etapas de diálise), foi capaz de reduzir a formação destas fibras, porém a proteína ainda se encontrava agregada. Nesta nova etapa, foram retiradas alíquotas em todas as etapas da diálise. Apenas as alíquotas T2, T3, T4, T6 e T8, que continham respectivamente 3 M – 1,5 M – 0,75 M – 0,19 M – 0,04 M de uréia, foram analisadas por microscopia eletrônica. Esta análise nos auxiliou a encontrar um ponto chave onde poderemos adicionar o heterodímero MP1/p14 à diálise, ou seja, em uma concentração entre 0,75 M e 1,5 M de uréia poderíamos tentar realizar a formação do complexo *Ragulator*, condição esta em que havia menor quantidade de agregados fibrosos.

## Capítulo 2



**Figura 19 – Microscopia eletrônica dos pontos de diálise da p18.** As figuras A, B, C, D e E correspondem à microscopia eletrônica da proteína p18 em 50 mM NaCl e 50 mM de Tris-HCl pH 8,8, contendo respectivamente 3 M – 1,5 M – 0,75 M – 0,19 M – 0,04 M de uréia. A condição C foi escolhida como ponto de diálise no qual serão adicionadas as proteínas MP1/p14, na tentativa de formar o complexo *Ragulator*.

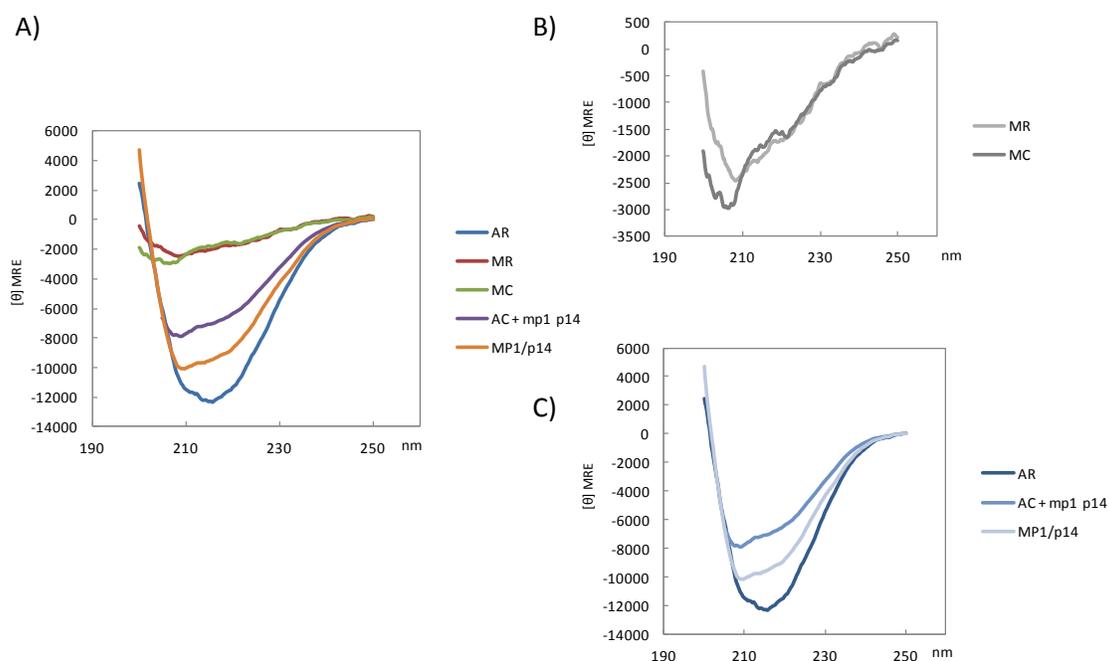
### 2.2.9. Dicroísmo circular – CD

Após a diálise em que foi obtido o complexo *Ragulator*, separamos uma única fração de cada pico (exceto picos 3 e 4, por não ter quantidade de proteína suficiente para observação em gel SDS-PAGE corado com *Coomassie*) presente na cromatograma de exclusão por tamanho molecular da p18 renaturada sozinha (controle) e com MP1/p14 (ver cromatograma **Figura 16**). As frações correspondentes ao primeiro pico de cada cromatograma, denominados p18 agregada GF p18 (para proteína dialisada sozinha - **AC**) e p18 agregada GF *Ragulator* (para proteína dialisada com MP1/p14 - **AR**). À fração do pico 2 do cromatograma de p18 renaturada sozinha, denominamos p18 monomérica GF p18 (**MC**), por supormos que esta apresenta-se em sua forma monomérica. Pelo mesmo motivo, denominamos a fração correspondente ao pico 5 do cromatograma referente a p18 renaturada com MP1/p14 de p18 monomérica GF *Ragulator* (**MR**). A fração fração 2 deste cromatograma continha **MP1/p14**.

## Capítulo 2

As frações continham quantidades muito pequenas de proteínas e foi preciso concentrá-las para analisá-las no CD. Além destas frações, ainda analisamos o espectro de p18 (**Figura 20**).

Os resultados obtidos por CD, sugerem uma variação no perfil de estrutura secundária entre a proteína p18 em sua forma “monomérica” e em sua forma agregada. A p18 “monomérica” renaturada com e sem MP1/p14 apresentam espectros de CD semelhantes. Espectros de CD da p18 agregada sugerem a formação de agregados amilóides, que devem ser confirmados por outra técnica. O espectro de CD contendo p18 agregada, somada a MP1/p14 não sugere interação entre estas proteínas.

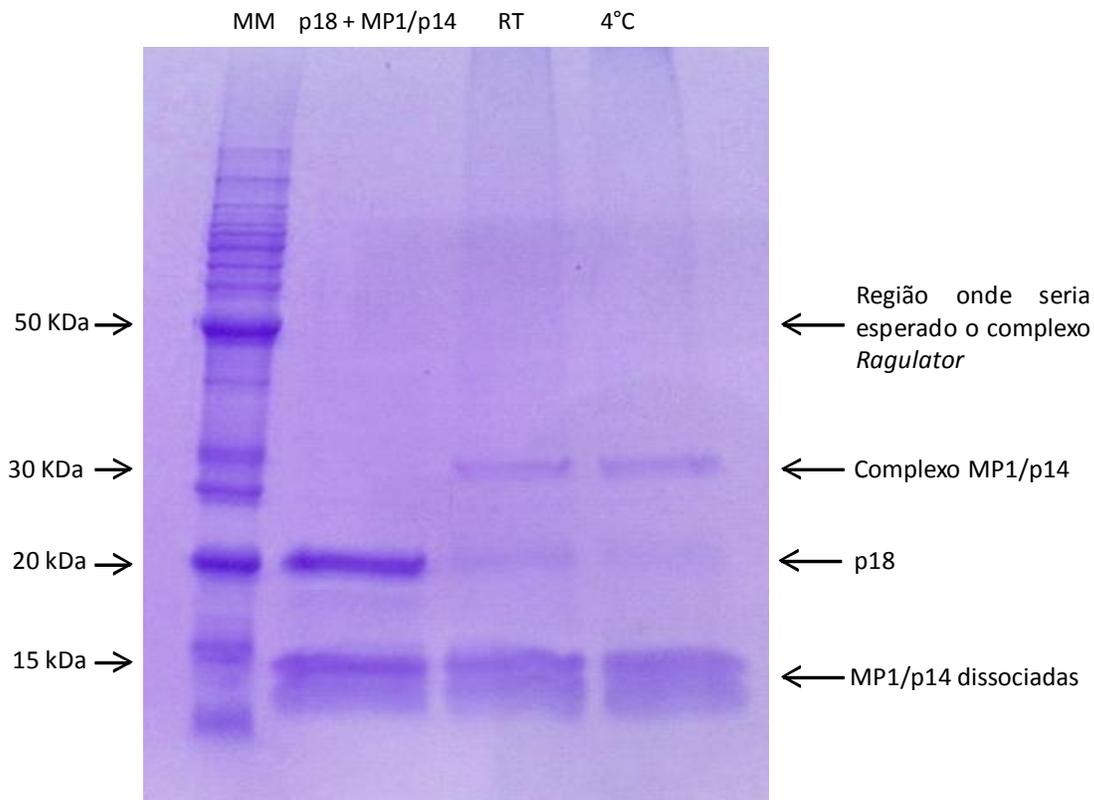


**Figura 20 – Espectro de CD das proteínas do Complexo *Ragulator*.** Espectros realizados a 10°C, com 1 mm de caminho óptico e a uma faixa de 205 a 250 nm. As abreviações contidas na figura estão explicitadas no texto. A) Comparação dos espectros de CD das proteínas p18 agregadas e “monoméricas”, com a proteína MP1/p14 e desta última associada a AC minutos antes da obtenção do espectro. B) Comparação dos espectros da proteína p18 “monomérica” renaturada com (MR) e sem (MC) MP1/p14. C) Comparação dos espectros de MP1/p14, da proteína p18 agregada (AR) e da junção de ambas (AC + MP1/p14).

## Capítulo 2

### 2.2.10. *Cross-linking*

A proteína p18 renaturada em presença de MP1/p14 (condição 7, **Tabela 1**) e submetida à cromatografia de afinidade a Ni foi submetida a uma reação de *cross-linking* a fim de estabilizarmos a interação entre as proteínas constituintes do *Ragulator*. A reação ocorreu por 4 h, nas temperaturas 4°C e ambiente. A reação de *cross-linking* promoveu ligações covalentes apenas entre as proteínas MP1/p14, e não entre p18, MP1 e p14, como era o esperado (**Figura 21**). Provavelmente a banda do complexo ternário não aparece devido à baixa quantidade de amostra. Outras tentativas de *cross-linking* estão sendo realizadas e, em caso de sucesso na ligação entre as três proteínas do *Ragulator*, a amostra será analisada por espectrometria de Massas.



**Figura 21 – Gel SDS-PAGE do *Cross-linking* do *Ragulator*.** MM corresponde ao marcador molecular *BenchMarker*<sup>®</sup>, a amostra anterior ao *cross-linking* e as amostras em que foi realizado o *cross-linking* a temperatura ambiente (RT) e a 4°C, respectivamente. A banda de aproximadamente 30 kDa corresponde ao heterodímero MP1/p14, nenhuma banda foi observado próximo ao 50 kDa, onde estaria o complexo *Ragulator*.

## Capítulo 2

Enquanto realizávamos estes experimentos, um artigo foi publicado onde os autores conseguiram uma co-expressão contendo MP1/p14 e p18 [35]. Para isso, eles fizeram deleções nas regiões N e C terminal da p18 e pela primeira vez provaram sua interação com MP1/p14 *in vitro*. Entretanto a real região de interação entre estas proteínas ainda é desconhecida. O clone utilizado nesta co-expressão foi solicitado aos autores do artigo e no momento estamos aguardando sua chegada. Com esta co-expressão pretendemos mapear, através de cross-linking e Espectrometria de Massas, os resíduos envolvidos nesta interação.

# Capítulo 3

## A proteína FtsZ e sua moduladora MinC

### 3. Capítulo 3 – FtsZ e sua moduladora MinC

#### 3.1. INTRODUÇÃO

A organização subcelular de proteínas em bactérias possui notável similaridade com a de células eucarióticas, de modo que as três maiores famílias de proteínas de citoesqueleto eucarióticos (tubulina, actina e filamentos intermediários) são homólogas de proteínas bacterianas, incluindo FtsZ, MreB e crescentina. Apesar da similaridade estrutural, proteínas do citoesqueleto eucariótico e bacteriano estão envolvidas em diferentes funções fisiológicas, e a função *in vivo* destas proteínas bacterianas ainda não está completamente compreendido [42]. Seu estudo pode auxiliar o entendimento de processos patológicos bem com o desenvolvimento de novos fármacos bactericidas.

#### *FtsZ e a divisão celular*

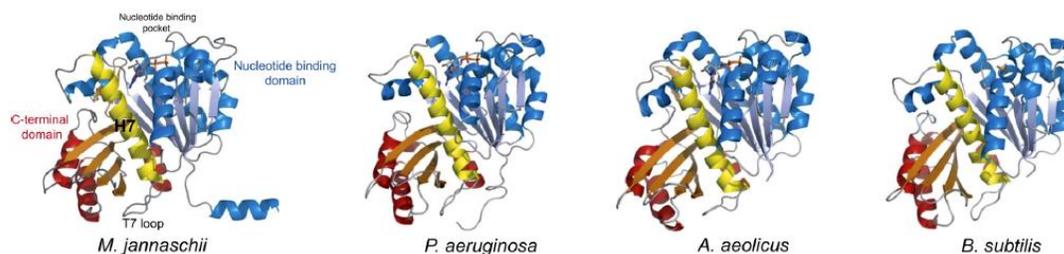
A proteína FtsZ é uma homóloga bacteriana da tubulina, essencial no processo de divisão celular. Está presente virtualmente em todas as bactérias, na maioria das Archaeas e em cloroplastos e mitocôndrias de eucariotos. É a proteína mais conservada entre as proteínas relacionadas no processo de divisão celular [43].

Esta proteína polimeriza em presença de GTP e forma protofilamentos que se associam formando a base do anel Z. Este anel consiste em um complexo protéico que se estende ao redor da circunferência celular e define o plano de divisão durante o processo de divisão celular. Este septo ocorre no centro da célula, onde FtsZ recruta inúmeras proteínas da divisão celular, incluindo as responsáveis pela degradação e pela construção da parede celular [42]. O anel Z é ancorado a superfície interna da membrana plasmática pela proteína FtsA [44] e provavelmente também por ZipA [45].

O septossomo é um multicomplexo citocinético formado pela reunião de pelo menos 10 proteínas da divisão celular associadas ao anel Z, e esta associação é dependente da presença do anel Z no sítio de divisão. Por isso, acredita-se que o anel Z é uma plataforma inicial para associação de todos, ou pelo menos da maioria dos componentes essenciais da maquinaria de divisão celular [43]. FtsZ é responsável pelo formato do

invaginamento do septo, uma vez que mutantes de FtsZ que formam espirais ao invés da estrutura de anel possuem septos com a mesma diferença no formato [46]. Não se sabe se FtsZ é apenas uma proteína suporte durante a reunião das proteínas da divisão bacteriana, se ela apresenta alguma função regulatória ou se participa ativamente no encolhimento do novo septo [43].

A proteína FtsZ possui estrutura conhecida de espécies diferentes, entre elas *M. jannaschii*, *P. aeruginosa*, *A. aeolicus* e *B. subtilis* (**Figura 22**). FtsZ e a tubulina, sua homóloga, compartilham um enovelamento arquitetado em 2 domínios: a região N terminal, que contém o sítio de ligação ao nucleotídeo, é ligada ao domínio C terminal através de uma hélice central (H7), que é envolvido formando um protofilamento. O sítio de atividade GTPásica é formado na interface entre os monômeros por inserção dos resíduos ácidos do *loop* T7 do domínio C terminal no bolso de ligação a nucleotídeos do monômero anterior do protofilamento [47]. FtsZ e tubulina possuem propriedades de estrutura e polimerização similares, entretanto, estas proteínas não são sensíveis as mesmas moléculas inibitórias [42], e esta especificidade faz de FtsZ um bom alvo farmacêutico no combate de diversas bacteremias.



**Figura 22 – Estruturas Cristalográficas da proteína FtsZ de 5 espécies diferentes.** Representação da estrutura de FtsZ em *M. jannaschii*, *P. aeruginosa* (resíduos 2 -6), *A. aeolicus* e *B. subtilis* (resíduos de 12 – 315). Em azul observa-se o domínio de ligação a nucleotídeos, a hélice H7 em amarelo e o domínio C terminal em vermelho/laranja [47].

Para que a divisão celular seja eficaz, ela possui precisão espacial e deve ser perfeitamente coordenada com a replicação e segregação do DNA. O septo deve ser formado no centro da célula, dando origem a duas células filhas idênticas [43]. Se um erro espacial ou cronológico ocorrer neste processo, será formada uma célula filha menor que o ideal e sem nucléolo e uma célula mãe remanescente aumentada e com dois nucléolos. A

regulação do processo de divisão celular necessita ser finamente regulada para a manutenção da vida. Em *B. subtilis* e *E. coli* foram identificados diversos moduladores do anel Z, dentre os quais se destacam o Sistema Min e o Sistema Noc/Slma [48]

#### ***O sistema MinCDE***

O processo de regulação do septo da divisão celular começou a ser desvendado há quatro décadas, quando Adler e colaboradores obtiveram mutantes de *E. coli* em que um grande número de pequenas células anucleadas, incapazes de se dividirem, eram formadas durante a fase logarítmica da divisão celular. Estas células, em forma de bastonetes, foram isoladas através de centrifugação por gradiente de densidade e utilizadas em diversos estudos biológicos [49]. Estes mutantes formavam septos em ambientes diferenciado, como nos pólos celulares. A formação do septo na região polar da célula dá origem a uma célula filha esférica com ausência de DNA cromossomal (mini células) e uma célula mãe residual com tamanho maior que o normal e dois nucleóides. Interessantemente o DNA plasmidial pôde ser dividido entre as células [43].

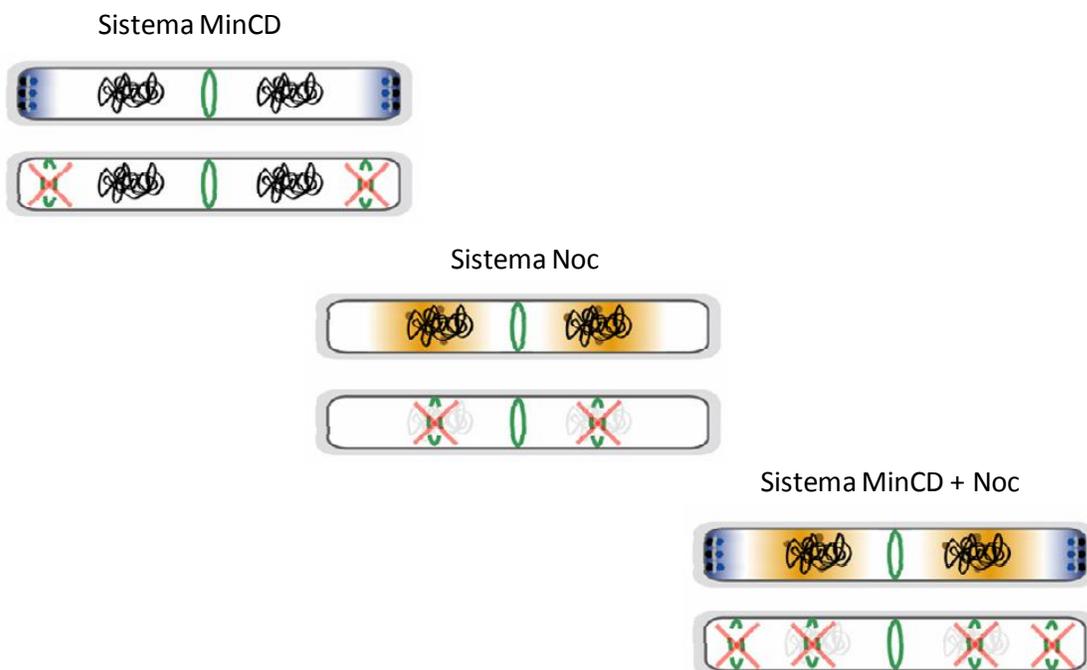
Em mini células de *E. coli* foi identificada uma mutação em um operon de 3 genes, o locus genético *minCDE* [50]. Os produtos de expressão destes genes evitam a formação do septo e divisão nos pólos celulares.

MinC é o inibidor da divisão deste sistema, interagindo com FtsZ e prevenindo a formação do anel Z estável [43]. MinC forma um heterocomplexo associado a membrana capaz de impedir a divisão nos pólos celulares [51] e a deleção destas proteínas resulta em uma divisão assimétrica causada pela formação de septos nos pólos celulares [52, 53]. A proteína MinE é um fator de especificidade topológica, responsável por dar especificidade a MinC, limitando sua atuação aos pólos celulares, para que a célula possa se dividir em sua região central [43]. MinD é a responsável pela associação de MinC e MinE à membrana plasmática [54, 55].

A proteína MinD é membro de uma superfamília de ATPases, onde estão inclusas ParA (envolvida em segregação dos cromossomos) e proteínas relacionadas, e se associa à membrana através de uma hélice anfipática localizada em sua extremidade N terminal [48, 56, 57]. O domínio C terminal da proteína MinC é o responsável por sua interação com

MinD [58], entretanto ambos os domínios C e N terminal são necessários para que MinC exerça sua função inibitória [58, 59].

A localização do sistema MinCDE nos pólos das células em *B. subtilis* se dá através da interação com a proteína DivIVA [43], delimitando, portanto o local de atuação deste sistema. Outro importante sistema regulador de FtsZ é o Sistema Noc/SlmA. O Sistema Noc em *B. subtilis* e SlmA em *E. coli* interagem com DNA e inibem a polimerização do anel Z [60, 61], impedindo que a divisão ocorra na região cromossomal.



**Figura 23 – Esquema de inibição da divisão bacteriana em *B. subtilis*.** O sistema MinCD impede que a divisão celular ocorra nos pólos, o sistema *Noc* impede que a mesma ocorra na região onde contém material genético, a soma dos dois sistemas faz com que a única região permitida para a divisão celular seja o centro. Modificado de [48]

## **3.2 RESULTADOS DA CARACTERIZAÇÃO DA INTERAÇÃO DE FtsZ COM MinC**

Os experimento a seguir foram realizados com o intuito de confirmar a interação de FtsZ e MinC de *Bacillus subtilis*, bem como identificar os resíduos nela envolvidos. A interação entre a região N terminal da MinC com FtsZ de *E. coli* foi confirmada recentemente [62], entretanto a interação entre as proteínas de *Bacillus subtilis* ainda precisa ser confirmada. Outros experimentos já foram realizados pelo grupo do pesquisador Dr. Frederico G. Filho, sem sucesso, provavelmente por se tratar de uma interação fraca. Dessa forma, esta colaboração visa utilizar uma ferramenta extremamente sensível como a RMN, capaz de detectar interações fracas.

### ***3.2.1. Expressão das proteínas MinC e FtsZ***

Os clones contendo as sequências gênicas referentes às proteínas FtsZ e MinC foram obtidos através de uma colaboração com o pesquisador Dr. Frederico Gueiros Filho (IQ-USP SP) (Projeto SMOLBNET 2.0, Processo FAPESP 2010/51866-0). Os protocolos de indução já estavam estabelecidos por este mesmo colaborador, porém foi preciso aperfeiçoá-los para que estas proteínas pudessem ser expressas em M9[63] e utilizadas no espectrômetro de Ressonância Magnética Nuclear. O protocolo pode ser visto na seção Metodologia.

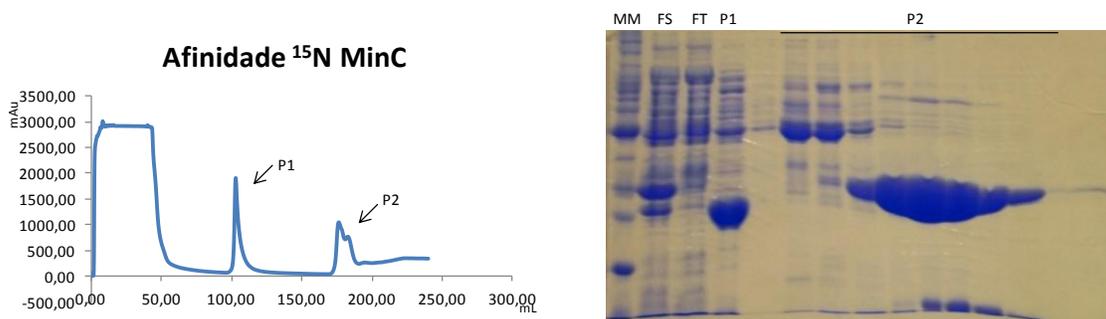
### ***3.2.2. Expressão e purificação das proteínas MinC e FtsZ***

A purificação das proteínas expressas em meio LB e em meio mínimo M9 [63] para marcação com isótopos estáveis para estudos de RMN, foi realizada de forma similar, exceto pelos tampões que tiveram de ser adaptadas as exigências das técnicas de RMN. A padronização prévia foi feita no tampão HEPES 10 mM, 50 mM de Arginina, 50 mM de Ácido Glutâmico, 100 mM KCl, 1 mM de EDTA e 5% de glicerol. Para estudos por Ressonância Magnética Nuclear, é preferível utilizar o tampão fosfato a outros tampões, há

uma limitação na concentração de sal (máximo 100 mM, para utilização da sonda criogênica) e aditivos como Arginina e Ácido Glutâmico utilizados para aumentar a solubilidade, apresentam sinais no espectro e, por se encontrarem em uma concentração muito superior ao da proteína, saturam o receptor de sinal, inviabilizando a visualização de sinais da proteína.

### 3.2.2.1 Purificação da MinC

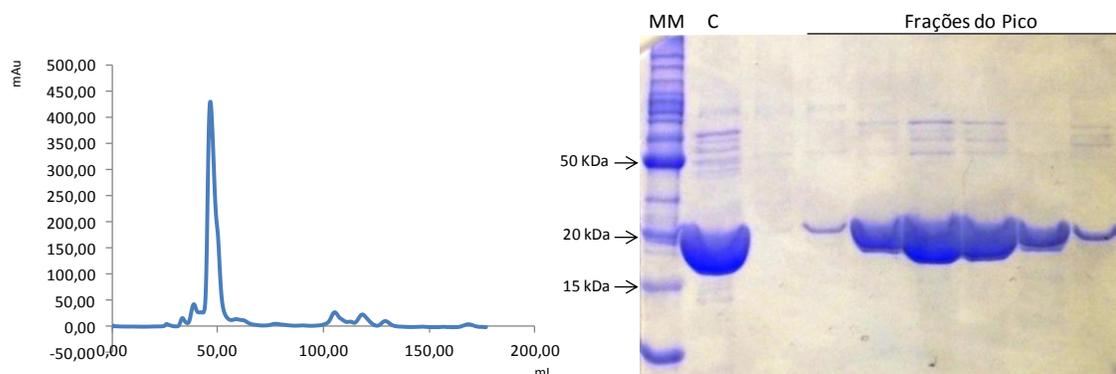
O lisado de células foi purificado em coluna de afinidade (His Trap Chelating FF-5 mL). A afinidade se mostrou um excelente método de captura para MinC por separar bem MinC das demais proteínas celulares. A purificação foi confirmada em gel de poliacrilamida (SDS-PAGE corado com Coomassie blue) a 12% (**Figura 24**)



**Figura 24 – Cromatograma da proteína MinC marcada com  $^{15}\text{N}$  e purificada por afinidade.** À direita, gel SDS-PAGE onde: MM corresponde ao marcador molecular *BenchMark*<sup>®</sup>, FS a fração solúvel aplicada à coluna de afinidade, FT a fração que passou pelo *flowthrough*, P1 a fração do primeiro pico, as próximas frações correspondem ao segundo pico, onde observamos MinC em alto rendimento.

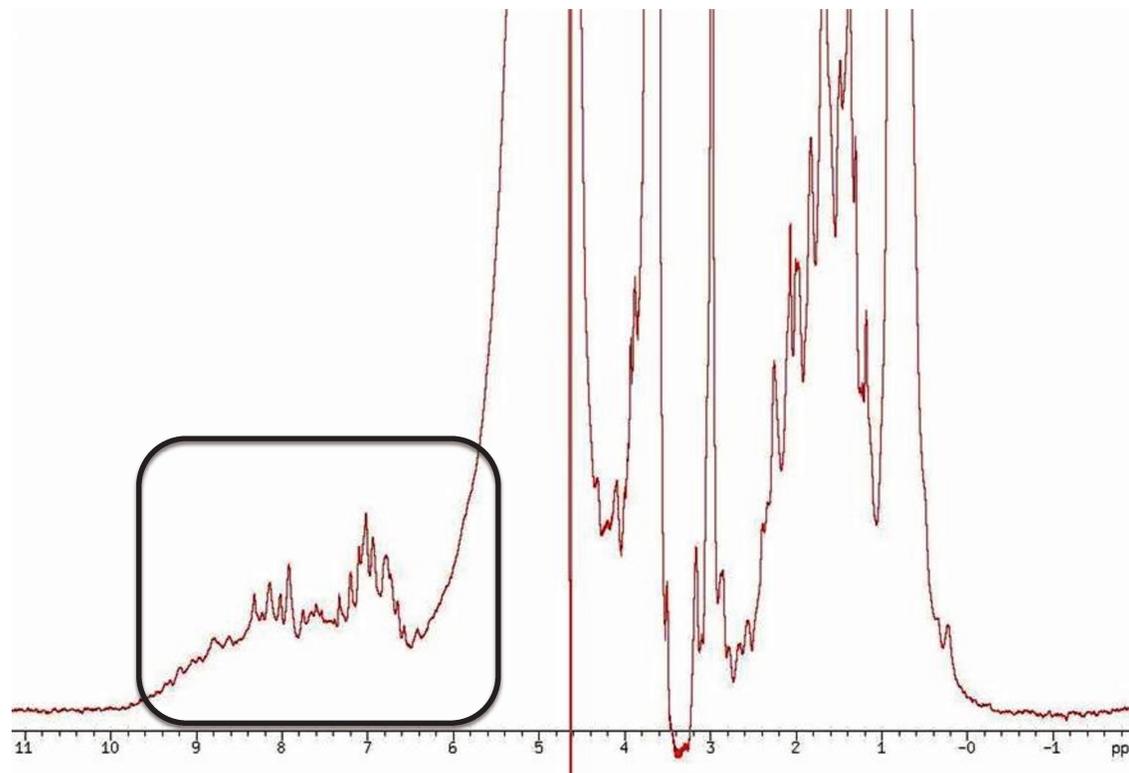
Embora a purificação por afinidade tenha fornecido resultados satisfatórios quanto a pureza e concentração de MinC, foi necessário um método de refinamento para evitar que alguns contaminantes interferissem no espectro de RMN. O método de escolha foi a cromatografia de exclusão por tamanho molecular, ou filtração em gel (coluna Superdex HiLoad 200 16/60 – GE Healthcare<sup>®</sup>). Esta cromatografia é capaz de separar eventuais contaminantes e possíveis dímeros ou até mesmo oligômeros que dependendo das características da proteína podem ser formados. Pode-se observar através desta cromatografia, apesar de não se tratar de uma filtração em gel analítica, que a proteína

MinC (~27,7 KDa) é eluída com 50 mL, sugerindo que esta não se encontra em sua forma monomérica (**Figura 25**).



**Figura 25.** Cromatografia da proteína MinC purificada por filtração em gel na coluna Superdex 200. As frações do gel SDS-PAGE correspondem respectivamente ao marcador molecular *BenchMark<sup>TM</sup>* (MM), proteína MinC concentrada pós afinidade (C) e as frações do pico contendo MinC.

A proteína pura e concentrada foi levada a espectrômetro de RMN, onde foi obtido um espectro de prótons em 1D, para confirmação de que a proteína se encontra estruturada e na concentração necessária para obtenção de sinais. No espectro pode-se observar que a proteína se encontra enovelada por apresentar dispersão na região do NH, próximo a 8 ppm. Proteínas desenoveladas apresentam pouca dispersão na região no NH, pois todos os aminoácidos experimentam o mesmo ambiente químico. A concentração alcançada pela purificação foi suficiente para obtermos um bom nível de sinal no espectro. O próximo passo foi obter espectros 2D com a proteína marcada com nitrogênio 15, analisar sua estabilidade e posteriormente sua interação com FtsZ não marcada (**Figura 26**).



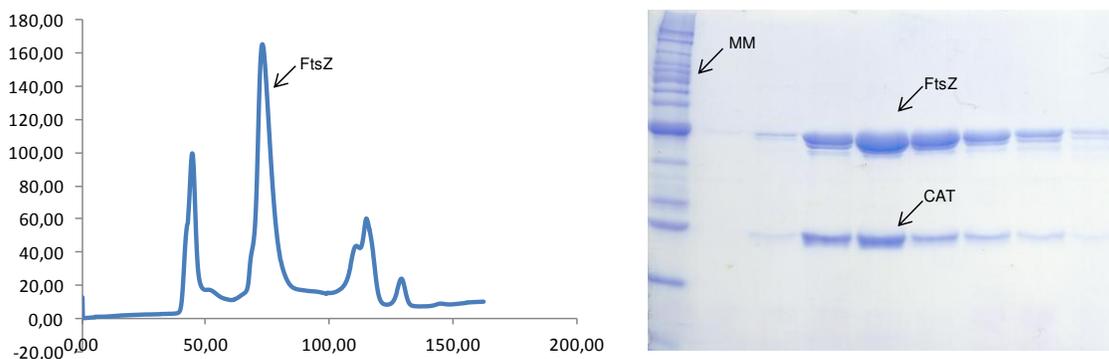
**Figura 26. Espectro unidimensional de Hidrogênio da proteína MinC.** O espectro foi realizado em um espectrômetro Varian Inova 600 MHz. Em destaque, observa-se a região do espectro onde são analisados os NHs das cadeias laterais dos aminoácidos.

Após confirmação, foi padronizada a expressão de MinC em meio mínimo M9, contendo uma única fonte de nitrogênio, no caso o isótopo  $^{15}\text{N}$ . Foram repetidos os procedimentos de expressão e purificação, e a proteína marcada foi submetida à espectroscopia de RMN para visualização de espectros 2D  $^{15}\text{N}$ -HSQC (**Figuras de 32 a 38**).

#### 3.2.2.2. Purificação da FtsZ

Para FtsZ, a cromatografia por afinidade em coluna de Níquel não se mostrou uma boa opção de purificação. A proteína purificada sob estas condições se mostrava incapaz de se polimerizar na presença de GTP e Cloreto de Cálcio e, portanto inativa. Por este motivo sua purificação foi feita precipitando-a com  $(\text{NH}_4)_2\text{SO}_4$ , como citado na seção Metodologia. Outra purificação foi necessária para garantir maior pureza de FtsZ. Para isto, a amostra obtida da precipitação com  $(\text{NH}_4)_2\text{SO}_4$  foi submetida a filtração em gel na coluna Superdex HiLoad 200 16/60<sup>®</sup>. As frações foram posteriormente analisadas por gel de

poliacrilamida (SDS-PAGE colorido com *coomassie blue*) para confirmação. Conseguimos uma alta pureza de FtsZ, porém esta se apresentava constantemente com um contaminante de menor tamanho molecular que, por espectrometria de Massas, realizada no LNBio, descobrimos ser a cloranfenicol acetil transferase (CAT), provavelmente proveniente do plasmídeo pRIL que utilizamos na expressão e possui marca de seleção para cloranfenicol. (Figura 27).



**Figura 27. Cromatografia de exclusão por tamanho molecular da proteína FtsZ purificada por precipitação em sulfato de amônio.** A cromatografia foi realizada na coluna Superdex 200 16/60<sup>®</sup>. A proteína foi encontrada no segundo pico, sendo o terceiro pico uma provável contaminação por DNA, uma vez que este foi drasticamente diminuído com o tratamento da amostra com a enzima DNase. O gel SDS-PAGE contém um marcador molecular *BenchMarker*<sup>®</sup> e as frações correspondentes ao segundo pico da filtração, onde é possível visualizar as proteínas FtsZ e CAT (contaminante).

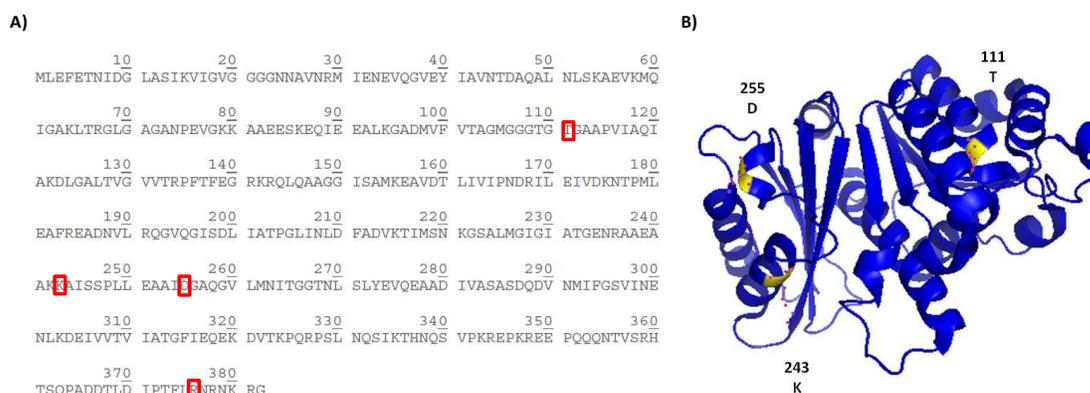
A proteína pura e concentrada foi utilizada para ensaios de interação com MinC por Ressonância Magnética Nuclear. Não foi necessária sua marcação com isótopos estáveis, uma vez que estaremos visualizando quais resíduos de MinC são envolvidos na interação com FtsZ.

Para retirada da CAT, contaminante de FtsZ, esta proteína foi expressa na cepa (DE3) C43 *E. coli*, cepa esta indicada para expressão de proteínas tóxicas. A ausência do plasmídeo pRIL impediu a expressão de CAT e a proteína FtsZ pode ser purificada sem seu contaminante, levando a um alto grau de pureza de FtsZ. A purificação de FtsZ sem CAT pode ser observada na purificação dos mutantes, no próximo item.

### 3.2.2.3. Purificação dos mutantes de FtsZ

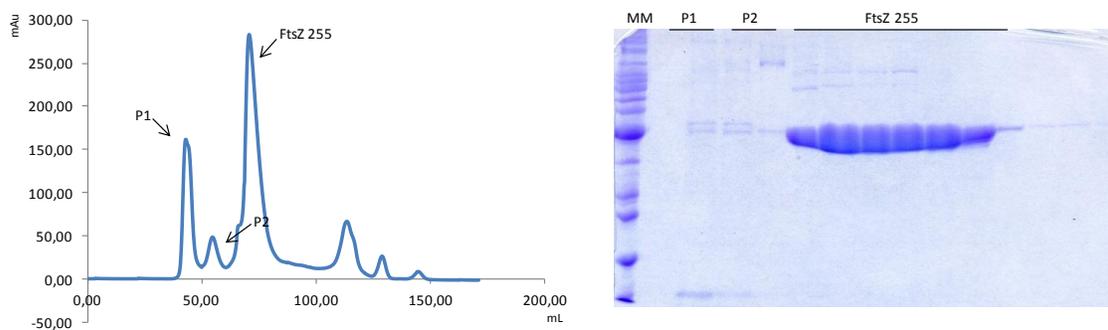
Além das proteínas selvagens, foram expressos também quatro mutantes pontuais de FtsZ de *B. subtilis* para titulação com MinC no espectrômetro de RMN. Estes mutantes provêm de uma biblioteca de mutantes ao acaso de FtsZ e desta biblioteca foram isolados os mutantes que continham algum tipo de resistência à MinC, seja por impedir a interação FtsZ/MinC, seja por tornar FtsZ mais estável. [48]. O grupo do pesquisador Frederico G. Filho nos cedeu alguns destes para que comparássemos suas interação com MinC em relação à com FtsZ selvagem.

Estes mutantes possuem alteração de um único resíduo por uma Alanina e foram nomeados de acordo com o resíduo mutado. Os resíduos utilizados foram 111, 243, 255, e 376 (**Figura 28**).



**Figura 28 – Mutantes de FtsZ.** A) Sequência de aminoácidos da FtsZ (*ProtParam*), destacando os aminoácidos que foram mutados por alanina. Número de aminoácidos total: 382. Peso Molecular 40394.8. Ponto Isoelétrico 5,01. B) FtsZ (PDB 2AVM) destacando os resíduos de aminoácidos de FtsZ. FtsZ 111 possui mutação no aminoácido 111, FtsZ 243 no 243, FtsZ255 no 255 e o FtsZ 287 possui mutação no resíduo 287, que não pode ser representado por não haver densidade eletrônica suficiente nesta região.

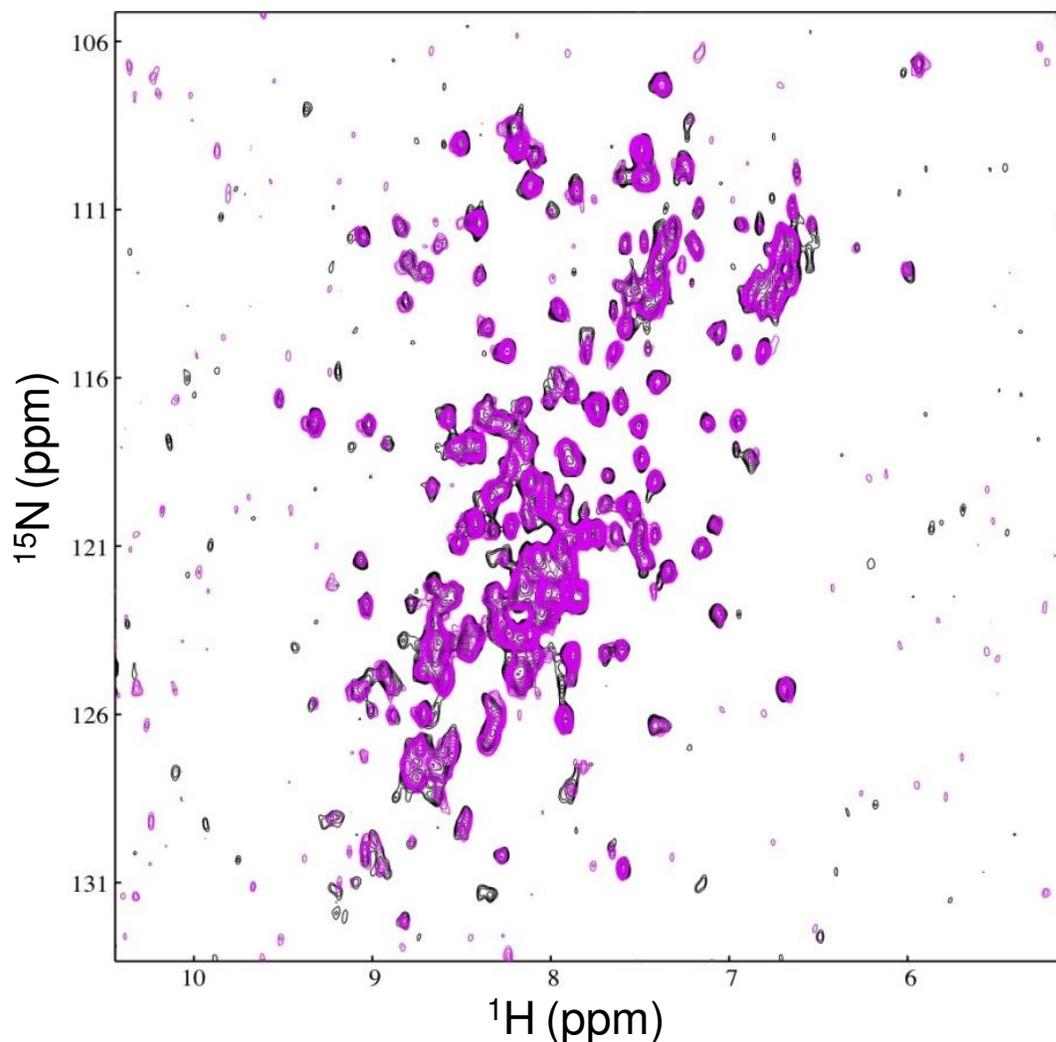
A expressão e purificação dos mutantes se deu da mesma forma que a FtsZ selvagem. Os mutantes apresentaram um perfil de filtração em gel muito parecido com o do selvagem e por utilizarmos a cepa de *E. coli* (DE3) C43, não apresentaram contaminação com a CAT (**Figura 29**).



**Figura 29 – Cromatograma da filtração em gel do mutante 255 sem a CAT.** O cromatograma do mutante é semelhante a de FtsZ selvagem. O gel SDS-PAGE ao lado contém o marcador molecular *BenchMarker*<sup>®</sup> (MM), amostras referentes aos picos 1 e 2 (P1 e P2) e amostras referentes ao terceiro pico contendo o mutante de FtsZ 255 sem a contaminação da CAT.

### 3.2.3. *Teste de Estabilidade de MinC a 40°C*

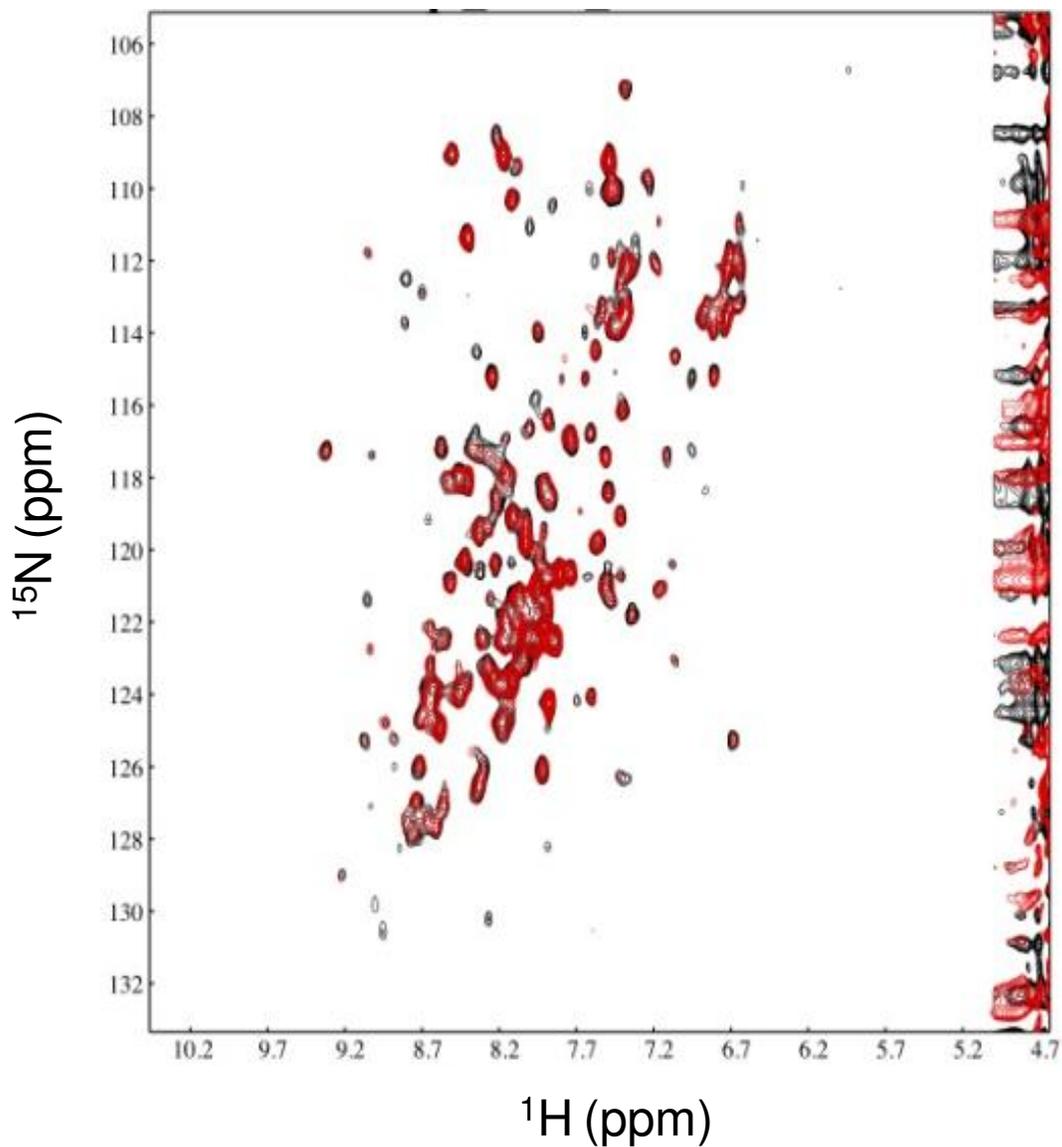
Antes de realizarmos a titulação realizamos um teste de estabilidade com a proteína MinC, a fim de descartar a possibilidade de que as mudanças no espectro estariam ocorrendo graças a instabilidade da amostra e não por causa da interação com FtsZ. Foram coletados nove espectros 2D-<sup>15</sup>NHSQC, cada um durando 23 minutos, e os espectros foram sobrepostos. Foram obtidos repetidamente espectros da proteína MinC a 40°C por 207 minutos a fim de acompanhar sua estabilidade. Estes espectros foram sobrepostos e não houve variação entre os picos, sugerindo que a proteína se encontra estável e que após a titulação com FtsZ a variação dos picos será exclusivamente devido à interação e não por instabilidade da MinC (**Figura 30**). Nestes espectros não observamos variações nas posições dos picos, mostrando que a amostra está estável e que a titulação pode ser feita.



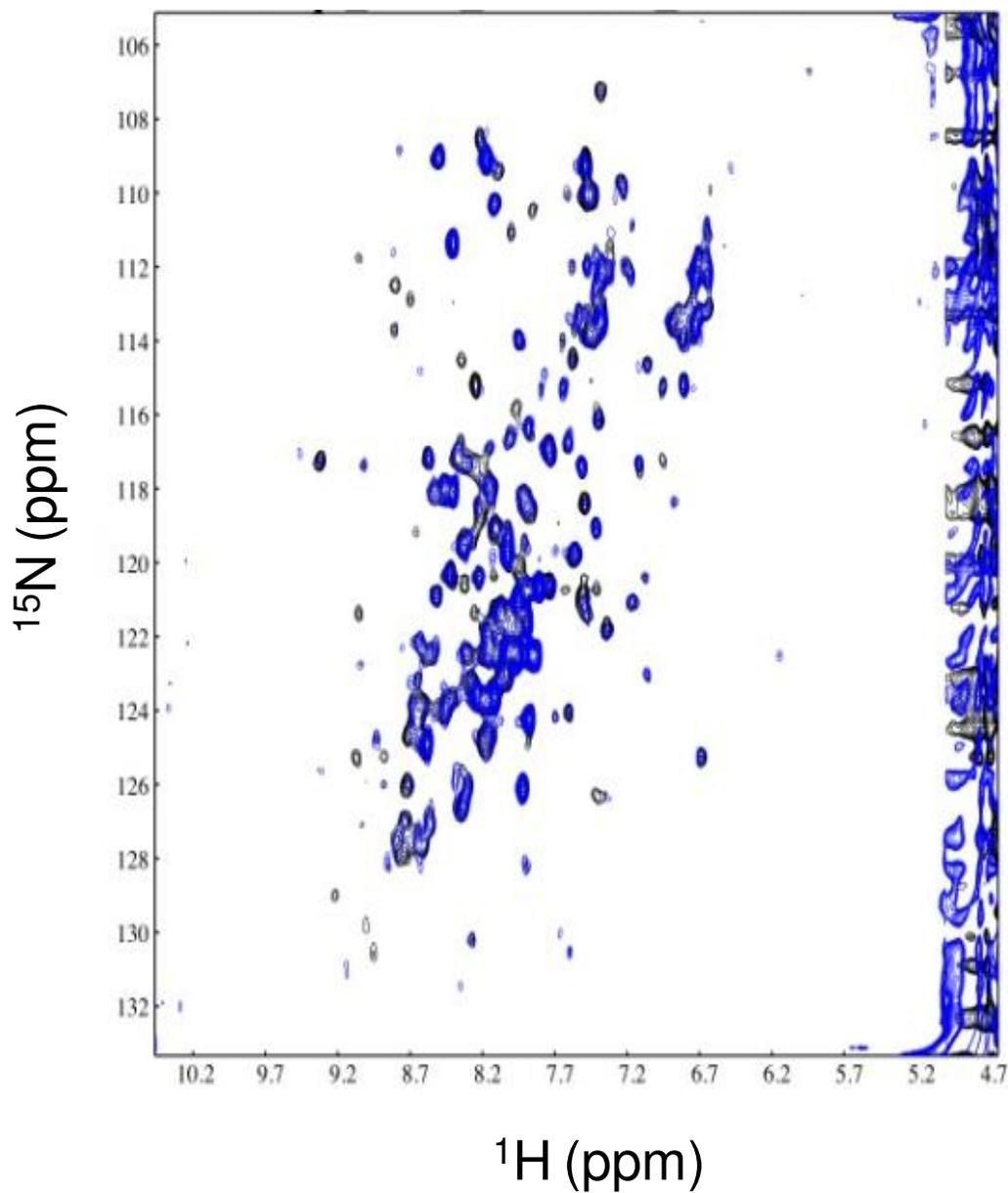
**Figura 30 – Espectro  $^{15}\text{N}$ -HSQC do teste de estabilidade da MinC.** Em preto pode-se observar o primeiro espectro, onde a proteína permaneceu 23 minutos a  $40^{\circ}\text{C}$ , em roxo, sobreposto, temos o espectro do último  $^{15}\text{N}$ -HSQC, em que a proteína já havia permanecido a  $40^{\circ}\text{C}$  por 207 minutos.

#### 3.2.4. Titulação de MinC com FtsZ

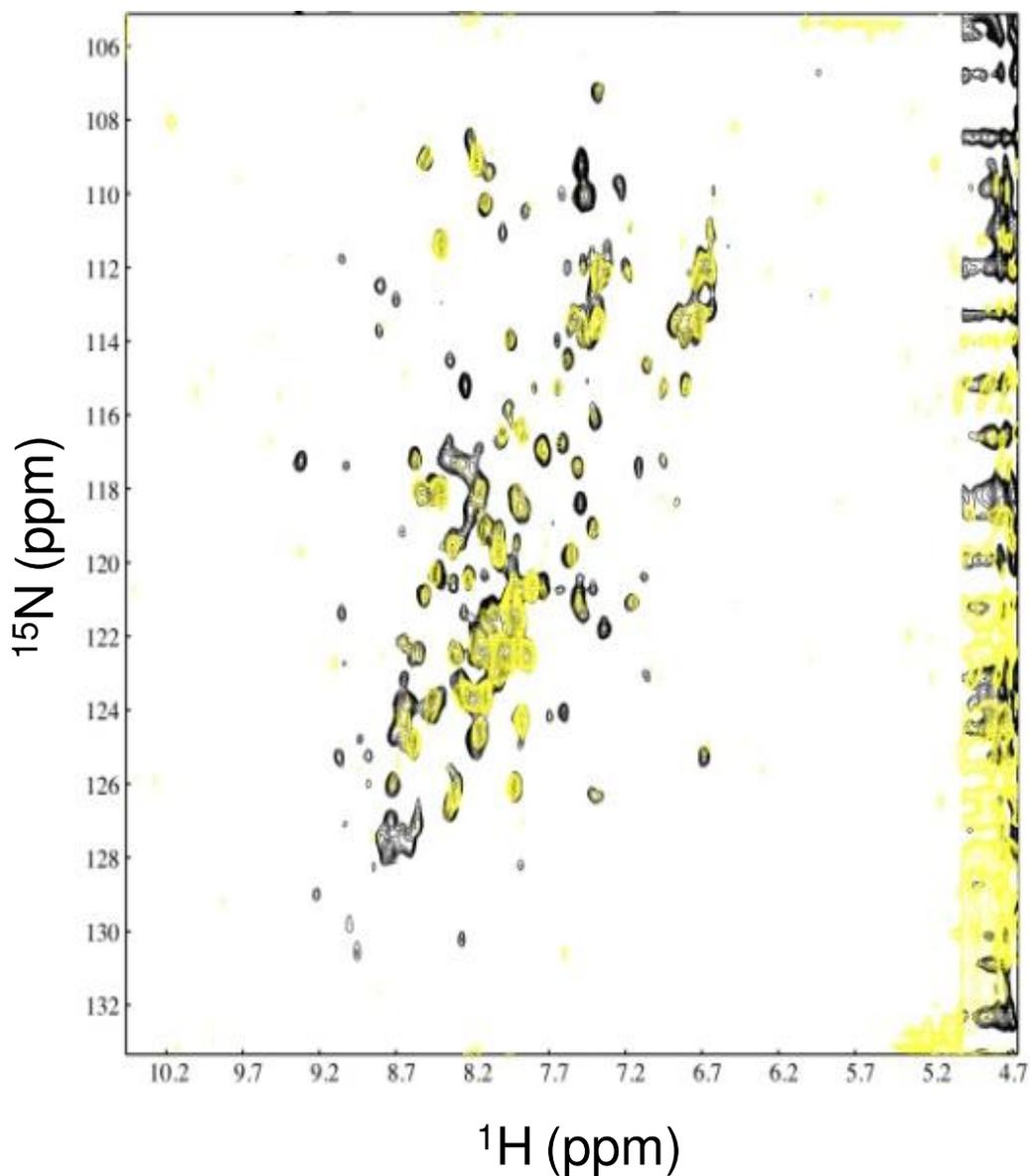
A proteína MinC foi titulada com FtsZ nas proporções 50:1, 10:1, 5:1, 2,5:1 e finalmente, 1:1. Em cada uma destas condições, obtivemos um espectro  $^{15}\text{N}$ -HSQC. Os espectros foram sobrepostos e neles pudemos observar o deslocamento de alguns picos sugerindo interação entre estas duas proteínas (**Figura de 31 a 36**). O deslocamento químico de um pico depende do ambiente em que o átomo se encontra e o fato de um átomo, interagindo com outro, modificar este ambiente químico.



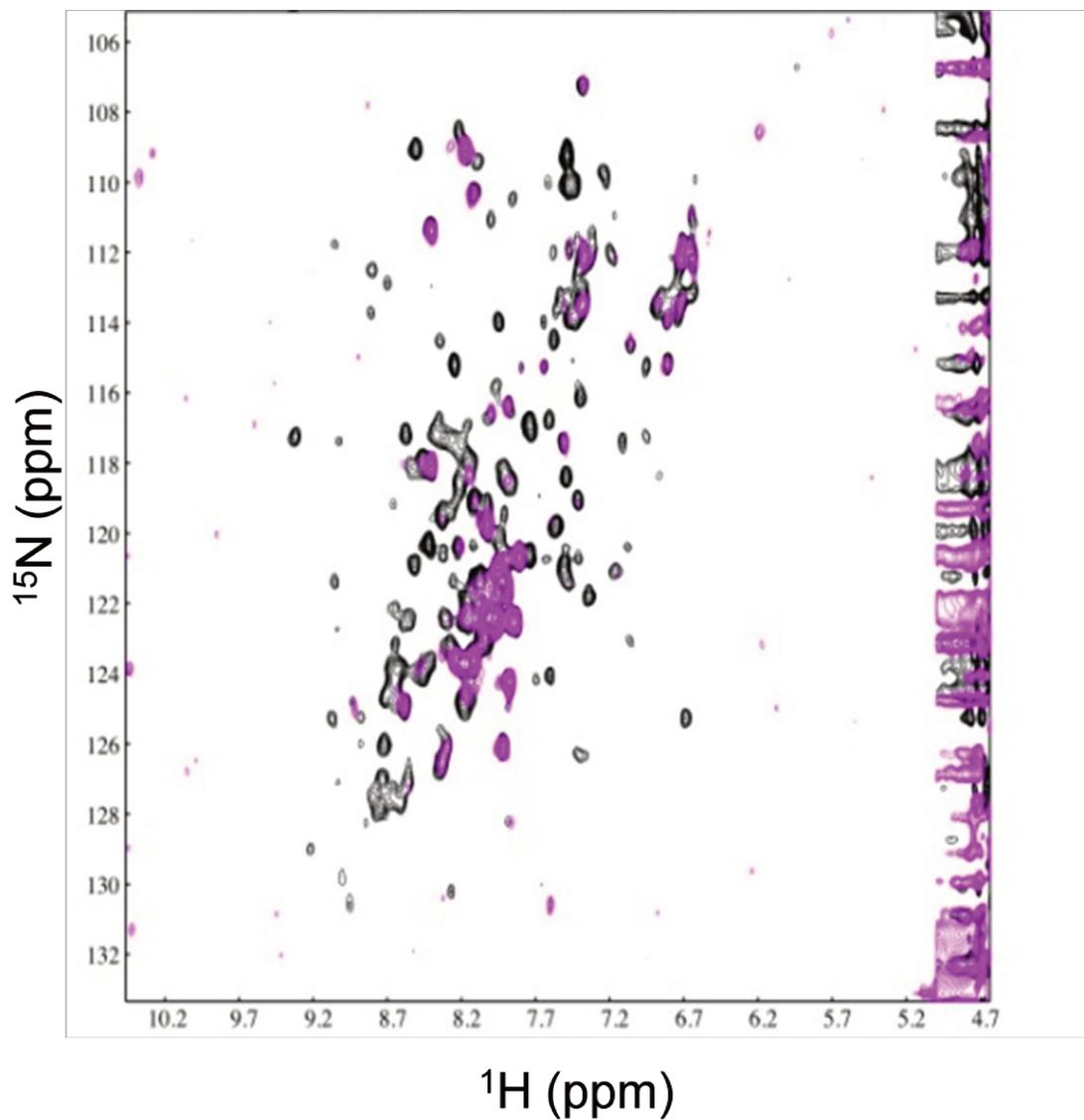
**Figura 31 – Espectro  $^{15}\text{N}$ -HSQC da titulação de MinC com FtsZ na proporção 50:1.** Em cor preto observamos os picos correspondentes a proteína MinC marcada com  $^{15}\text{N}$  sem FtsZ. Sobreposto em vermelho observamos espectro MinC titulada com na proporção 50:1.



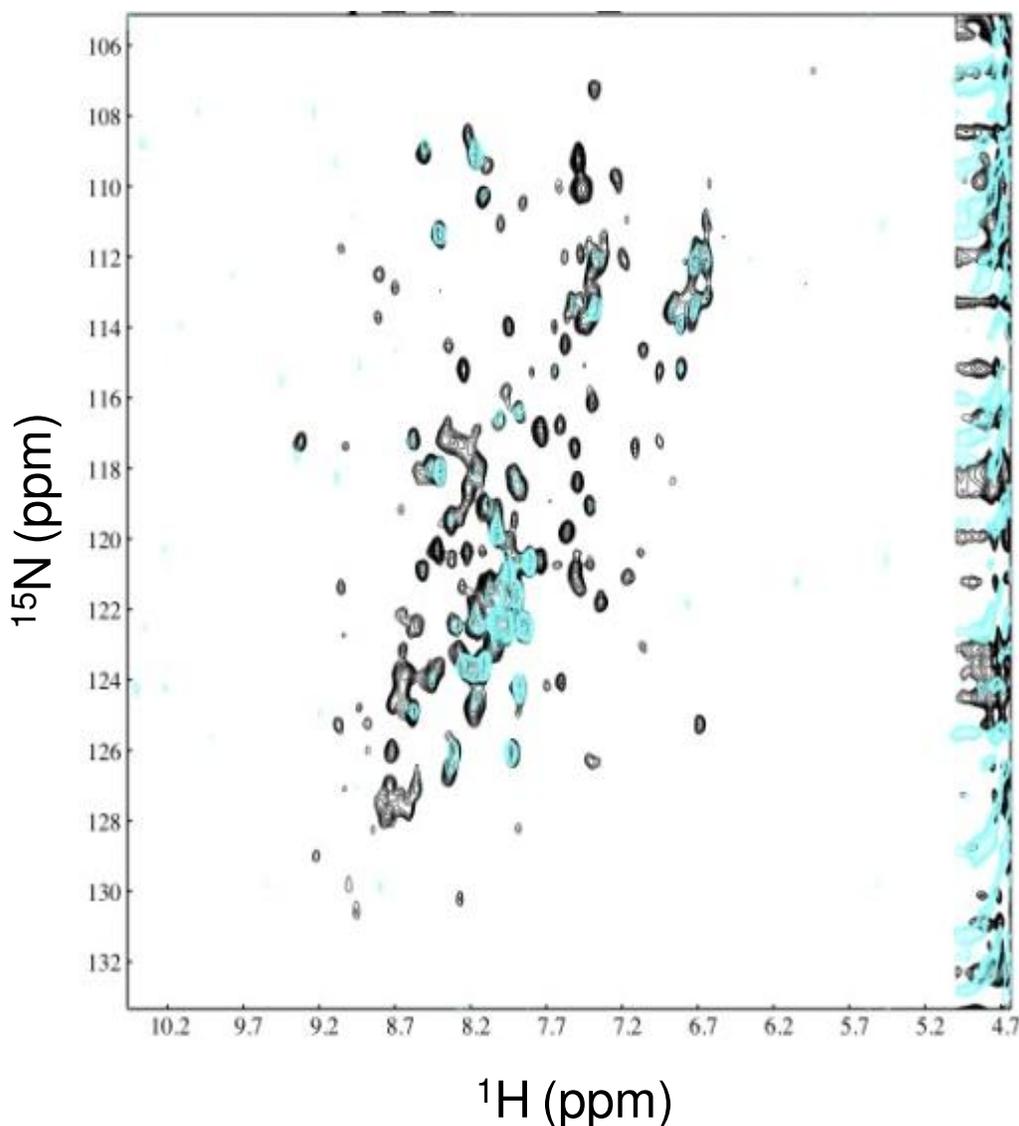
**Figura 32 – Espectro  $^{15}\text{N}$ -HSQC da titulação de MinC com FtsZ na proporção 10:1.** Em cor preto observamos os picos correspondentes a proteína MinC marcada com  $^{15}\text{N}$  sem FtsZ. Sobreposto em azul observamos espectro MinC titulada com na proporção 10:1.



**Figura 33 – Espectro  $^{15}\text{N}$ -HSQC da titulação de MinC com FtsZ na proporção 5:1.** Em cor preto observamos os picos correspondentes a proteína MinC marcada com  $^{15}\text{N}$  sem FtsZ. Sobreposto em amarelo observamos espectro MinC titulada com na proporção 5:1.



**Figura 34 – Espectro  $^{15}\text{N}$ -HSQC da titulação de MinC com FtsZ na proporção 2,5:1.** Em cor preto observamos os picos correspondentes a proteína MinC marcada com  $^{15}\text{N}$  sem FtsZ. Sobreposto em roxo observamos espectro MinC titulada com na proporção 2,5:1.

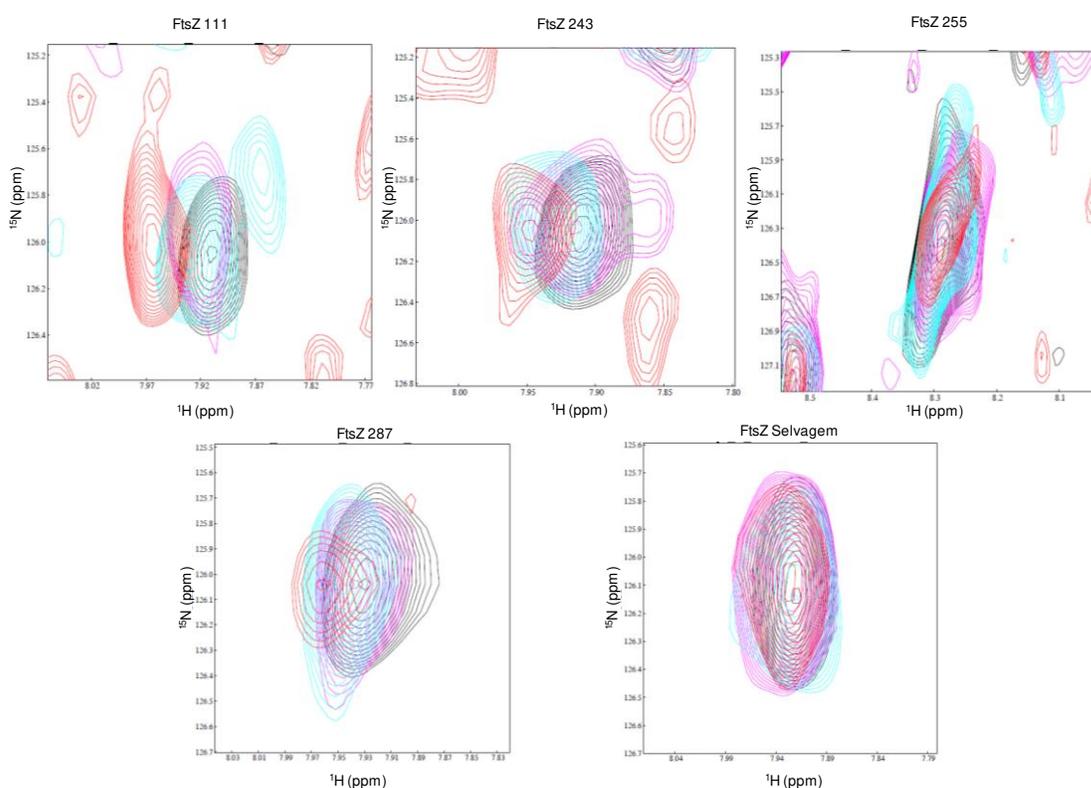


**Figura 35 – Espectro  $^{15}\text{N}$ -HSQC da titulação de MinC com FtsZ na proporção 1:1.** Em cor preto observamos os picos correspondentes a proteína MinC marcada com  $^{15}\text{N}$  sem FtsZ. Sobreposto em azul observamos espectro MinC titulada com na proporção 1:1.

As titulações apresentadas acima confirmam a ocorrência de uma interação fraca entre as proteínas FtsZ e MinC. O deslocamento químico de determinados picos (resíduos de aminoácidos) de MinC provocados pela titulação com FtsZ é indicativo desta interação. Pode-se identificar que o deslocamento do pico provém da interação da proteína a ser titulada simplesmente pelo fato de que o deslocamento é proporcional ao aumento da concentração proteína a ser titulada, ou seja, quanto mais adicionarmos da proteína titulada, mais o pico é deslocado, até um momento em que este deixa de mover. A interação é

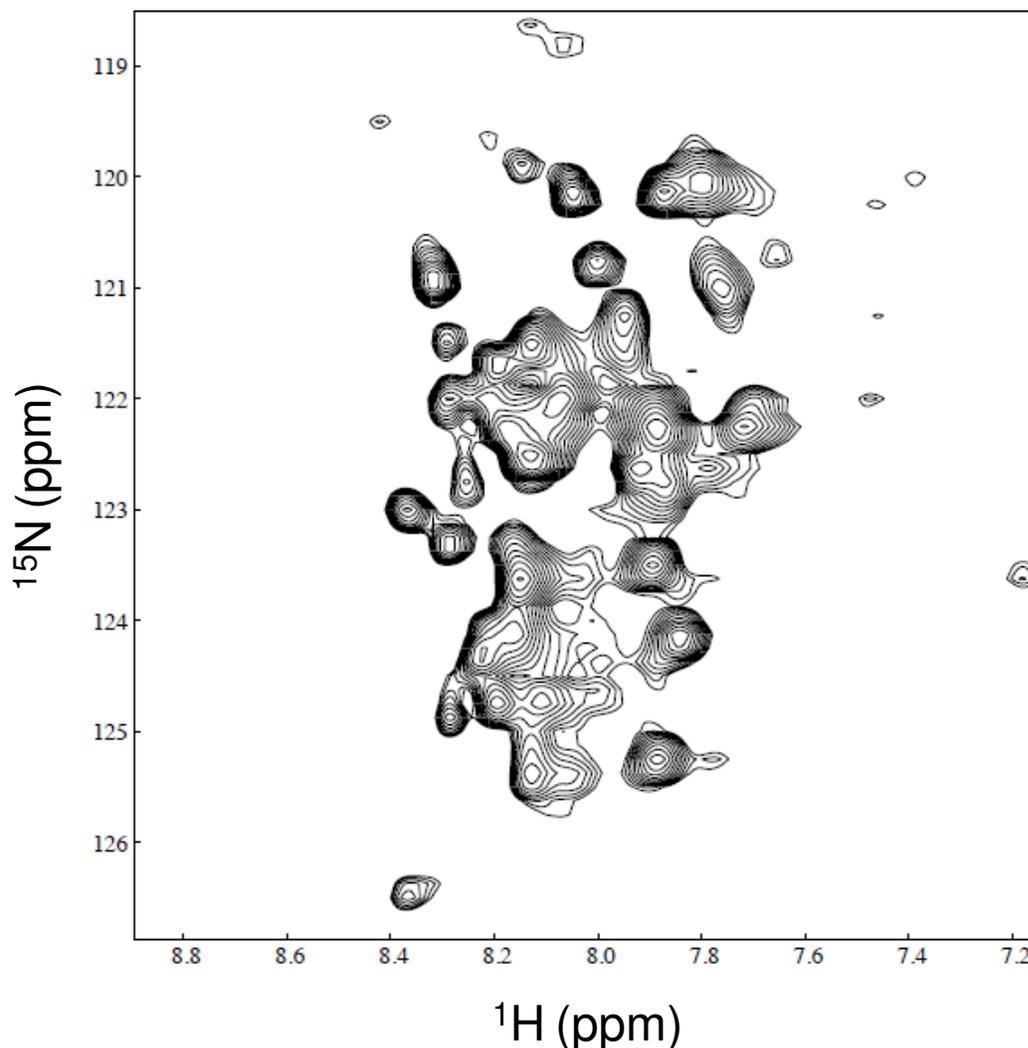
considerada fraca, pois nos picos em que pudemos observar interação, a variação de deslocamento químico provocada pela titulação é pequena.

Não foi possível identificar quais os resíduos envolvidos nesta interação, pois os espectros apresentaram sinais agregados na região central do espectro, impossibilitando a análise de cada pico separadamente. Entretanto, alguns picos isolados puderam ser analisados, onde conseguimos observar mudanças entre FtsZ mutante e selvagem, sugerindo que estes estão envolvidos na alteração entre MinC e as FtsZ e que as mutações pontuais podem estar modificando o perfil desta interação (**Figura 36**).



**Figura 36 – Sobreposição dos espectros da titulação de MinC (preto) na proporção 5:1 de FtsZ (magenta), 2,5:1 (azul) e 1:1 (vermelho).** Nesta sobreposição observamos apenas um pico localizado na região a 116.15 ppm no  $^{15}\text{N}$  e 7.9 ppm do  $^1\text{H}$ . Neste pico, particularmente, FtsZ selvagem e o mutante 255 não apresentam mudanças em seus deslocamentos químicos durante sua titulação à MinC, diferentemente dos demais mutantes, 111, 243 e 287, que apresentaram mudanças.

Para fins investigativos, foi realizada ainda a titulação inversa, em que a proteína FtsZ foi marcada com  $^{15}\text{N}$  e MinC expressa em LB. O espectro de  $^{15}\text{N}$  FtsZ apresentou menos sinais/picos que o número total de seus aminoácidos, indicando que a proteína FtsZ está com um tamanho maior que o limite para a técnica  $^{15}\text{N}$ -HSQC e ainda maior que o tamanho da MinC (Figura 37).



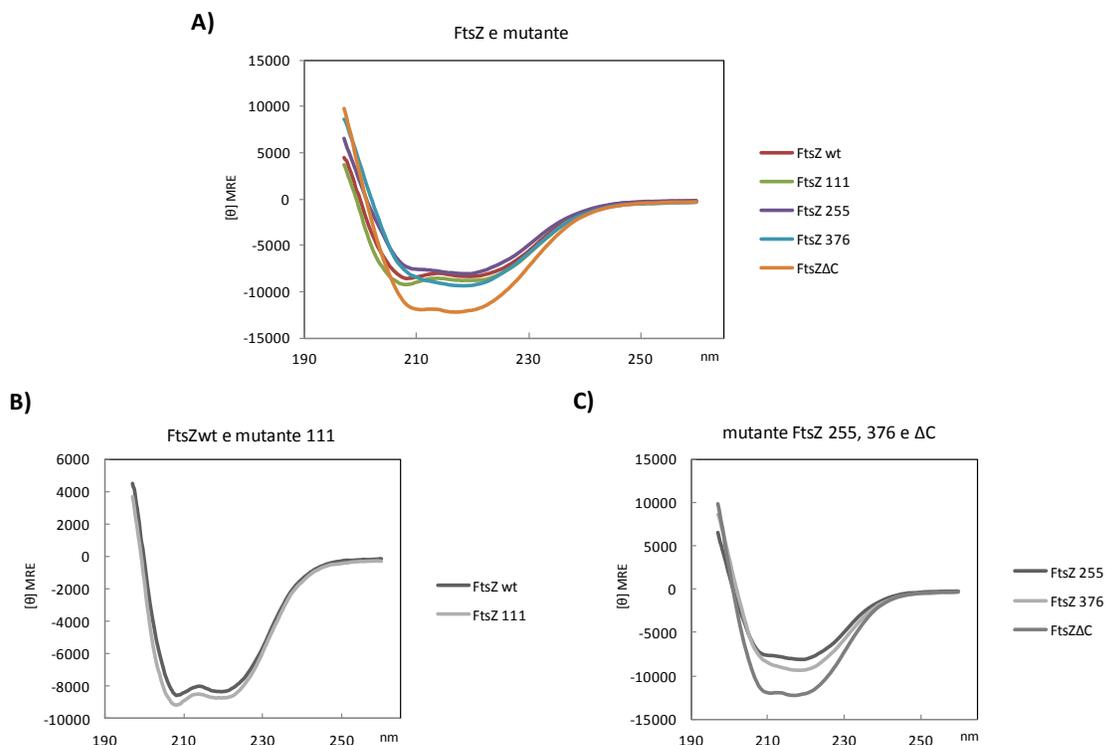
**Figura 37 – Espectro de  $^{15}\text{N}$ -HSQC de FtsZ.** A proteína FtsZ foi marcada com  $^{15}\text{N}$ , entretanto o agregado de picos na região central não viabilizou a identificação de picos de interação com MinC (titulação não mostrada).

Estas análises ainda não são conclusivas, e este perfil aglomerado dos espectros se dá provavelmente por as proteínas não se encontrarem em sua forma monomérica. Estudos de SAXS ainda não concluídos, realizados por nosso colaborador Dr. Wanius José Garcia

da Silva (UFABC), sugerem que MinC está em sua forma tetramérica e FtsZ em sua forma dimérica. O estado de agregação destas proteínas foram investigados também por DLS. Considerando que estas proteínas não estejam em sua forma monomérica, o tamanho esperado para o complexo será além do tamanho limite para a técnica de RMN  $^{15}\text{NHSQC}$ . Uma solução para este impasse seria expressar a proteína MinC em água deuterada, o que aumentaria significativamente o limite de tamanho da técnica. No momento, estamos expressando a proteína MinC em água deuterada a fim de testar novas titulações.

### *3.2.5. Espectros de Dicroísmo Circular de MinC, FtsZ e mutantes*

A fim de obtermos maiores informações sobre a interação entre FtsZ e MinC, realizamos análises de Dicroísmo circular onde os espectros de FtsZ e seus mutantes foram comparados entre si. FtsZ selvagem e o mutante 111 apresentaram um perfil muito semelhante, típico de proteínas com maior conteúdo de  $\alpha$  hélice. Os mutantes de FtsZ 255, 376 apresentam um maior conteúdo de folhas  $\beta$ , quando comparado com FtsZ selvagem e o mutante FtsZ $\Delta$ C apresenta uma conformação intermediária entre os perfis citados (**Figura 38**).



**Figura 38 – Espectro de CD da proteína FtsZ e mutantes.** Foram realizados espectros de CD da proteína FtsZ e seus mutantes no espectômetro J-810 (Jasco), com um caminho óptico de 1 mm, a 40°C e na concentração de 5 μM. Os espectros mostraram diferenças entre os mutantes de FtsZ. A) Espectro de CD de FtsZwt (selvagem) e dos mutantes 111, 255, 376 e ΔC (com deleção na região C terminal). B) Espectro de CD de FtsZwt e do mutante 111 e C) Espectro de CD dos mutantes de FtsZ 255, 376 e ΔC.

Foram realizadas ainda análises do desenovelamento térmico das proteínas MinC, FtsZ selvagem e mutantes 111, 255, 376 e FtsZΔC (em que foram deletados 39 aminoácidos da região C terminal); bem como de cada uma destas FtsZs associadas a MinC. Posteriormente cada FtsZ (selvagem e mutante) foi associado a MinC, e o desenovelamento térmico destas duas proteínas foi analisado por Dicroísmo Circular. Para confirmar a interação entre estas proteínas, foram observadas as TMs (temperatura média em que metade do complexo se encontra enovelado e a outra metade desenovelado). Com os dados obtidos, não foi possível confirmar a interação, pois a presença de MinC não teve efeito sobre a estabilidade térmica em nenhum dos casos (pois a variação entre as TMs foi menor que 5°C). FtsZ111 é menos estável termicamente em relação à selvagem, enquanto que FtsZΔC é termicamente mais estável (**Tabela 2**).

**Tabela 2** – Variação da temperatura média de desenovelamento térmico de FtsZ selvagem e mutantes com MinC

<b>Proteína</b>	<b>TM (°C)</b>	<b>TM na presença de MinC (°C)</b>	<b> ΔTM </b>
MinC	43,19973	-	-
FtsZ wt*	40,37729	42,34165	1,96
FtsZ 111	32,91962	34,24732	1,33
FtsZ 255	40,73696	43,43944	2,70
FtsZ 376	42,73721	45,08769	2,35
FtsZ ΔC	48,61548	45,16039	3,46

\* FtsZwt: Selvagem.

Nossos resultados mostraram a interação entre as proteínas MinC e FtsZ. Adicionalmente pudemos observar diferenças na estrutura secundária, através de espectros de dicroísmo circular, entre a proteína FtsZ selvagem e seus mutantes pontuais. Novos estudos ainda devem ser realizados até a identificação dos resíduos de aminoácidos responsáveis pela interação entre FtsZ e MinC. A proteína MinC será deuterada e a titulação com FtsZ será observada através do experimento HSQC com pequenas modificações na sequência de pulso. A proteína FtsZ será deuterada e titulada com os domínios C e N terminal de MinC, separadamente.

# Capítulo 4

## Metodologia

## 4. METODOLOGIA

### *4.1. Meios de cultura, cepas e clones de E. coli*

Para padronização da expressão das proteínas utilizadas neste trabalho, utilizamos o meio de cultura LB (Extrato de levedura 0,5% m/v; Triptona 1% m/v; NaCl 200 mM/L; pH 7,4 ). Para preparo de placas contendo LB sólido foi adicionado a esta preparação bacto-ágar 1,5% m/v [64]. O meio de escolha para a marcação isotópica da proteína com  $^{15}\text{N}$  e  $^{13}\text{C}$  foi o meio mínimo M9 [63]. Esse meio consiste na adição de uma única fonte de carbono e outra única fonte de nitrogênio (Glicose e Cloreto de Amônio, respectivamente). A bactéria terá de utilizar estas fontes de C e N na síntese de seus aminoácidos e, conseqüentemente as proteínas expressas apresentarão estes átomos que, quando marcadas com os isótopos de spin  $\frac{1}{2}$ , poderão ser visualizadas nos espectros de RMN. Pode-se ainda, quando necessário, acrescentar alguns aditivos/metais a esse meio de cultura para melhorar o rendimento da expressão protéica, porém neste trabalho não foi necessário.

Para amplificação dos vetores de expressão empregados, utilizamos a cepa de *E. coli* DH5 $\alpha$ . As cepas, plasmídeos e antibióticos utilizados estão descritos na **Tabela 3**.

**Tabela 3** – Cepas e antibióticos utilizados neste trabalho

Clones (plasmídeo + inserto)	Cepas de <i>E. coli</i>	Antibiótico ( $\mu\text{g/mL}$ )
pET28a + p18	<i>E. coli</i> BL21 (DE3) $\Delta\text{slyD}$	50 canamicina
	<i>E. coli</i> BL21 (DE3) $\Delta\text{slyD}$ Prare	50 canamicina
	<i>E. coli</i> (DE3) C41 .	50 canamicina
pETDuet + MP1/p14 <sup>*1</sup>	<i>E. coli</i> BL21 (DE3)	50 ampicilina
pET28a + His- MP1/p14 <sup>*2</sup>	<i>E. coli</i> BL21 (DE3)	50 canamicina
pET28a + MP1/p14-His <sup>*2</sup>	<i>E. coli</i> BL21 (DE3)	50 canamicina
pET24b + MinC <sup>*3</sup>	<i>E. coli</i> BL21 (DE3) pRIL	30 canamicina + 50 cloranfenicol
pET28a + FtsZ <sup>*3</sup>	<i>E. coli</i> BL21 (DE3) pRIL	30 canamicina + 50 cloranfenicol

Clones obtidos através de colaboração com: \*1 Dr. Miroslaw Cygler (*Biotechnology Research Institute, NRCC, Quebec, Canadá*); \*2 Dr. Tim Clasen (*Reserch Institute of Molecular Pathology, IMP, Viena, Áustria*) e \*3 Dr. Frederico Gueiros Filho (IQ-USP/SP).

Os plasmídeos foram inseridos na bactéria através de transformação por choque térmico ou eletroporação. Todos os clones foram estocados em glicerol 60% (v/v) a  $-80^{\circ}\text{C}$ .

#### 4.2. Preparo de células de *E. coli* competentes

##### 4.2.1. Preparo de *E. coli* competente para transformação por choque térmico

Para preparo de células competentes por choque térmico, utilizamos 500 mL de inóculo, incubado a  $37^{\circ}\text{C}$  sob agitação de 200 rpm até atingir  $\text{D.O.}_{600\text{nm}} = 0,3$ . Após atingir o número de células desejado, levamos o inóculo a um banho de gelo por 5 minutos e centrifugamos a  $1.600 \times g$  a  $4^{\circ}\text{C}$  por 10 minutos. Ressuspendemos o “*pellet*” de células em 100 mL de tampão  $\text{Ca}^{+2}$ /glicerol gelado e descartamos o sobrenadante. Foram realizadas etapas de centrifugação e lavagem do “*pellet*”. As células (“*pellet*”) foram ressuspensas em

12 mL de  $\text{Ca}^{+2}$ /glicerol gelado, distribuídas em tubos de micro centrífugas e armazenados em freezer biológico (-80°C). Modificado de [64].

#### **4.2.2. Preparo de *E. coli* eletrocompetente**

Para preparo de células de *E. coli* eletrocompetentes, foi utilizado um pré inóculo de 100 mL, incubado a 37°C por 16 h sob agitação de 200 rpm. Diluímos então este pré inóculo em 1 L de meio fresco pré-aquecido, que foi incubado novamente a 37°C, 200 rpm até atingir  $\text{D.O.}_{600\text{nm}} = 1,0$  a 1,2. As células foram coletadas através de uma centrifugação a 1360  $\chi\text{g}$  por 20 minutos a 4°C. Após etapas de lavagens, o sobrenadante foi descartado e ressuspendemos as células em 3 mL de glicerol 20% gelado e alíquotamos 40  $\mu\text{L}$  por tubo e micro centrífuga (0,6 mL) e congelamos em freezer biológico (-80°C). Modificado de [64].

### **4.3. Transformação de *E.coli***

#### **4.3.3.1. Transformação de *E.coli* por choque térmico**

Após descongelar uma alíquota de *E. coli* competente em banho de gelo, adicionamos aproximadamente 200 ng do DNA plasmidial. Após 30 minutos de banho de gelo, a amostra foi colocada em um banho térmico a 42°C por 2 minutos. Após o choque térmico, adicionamos 1 mL de meio LB e incubamos por 1 hora a 37°C, para que o plasmídeo entre na célula e a bactéria o replique. Foram plaqueados 100  $\mu\text{L}$  da suspensão de bactérias em uma placa de LB contendo antibiótico, para seleção clonal. A placa foi incubada a 37°C por 16 horas. Modificado de [64].

#### **4.3.2. Eletroporação de *E.coli***

Descongelamos as alíquotas de células competentes em gelo, adicionamos o plasmídeo e transferimos as células para uma cubeta de eletroporação (0,1 cm). Submetemos as células a um pulso elétrico nas seguintes condições: 1,7 kV, 200 W e 25  $\mu\text{F}$ . Imediatamente após o pulso, adicionamos 1 mL de meio LB na cubeta para remover as células e as transferimos para um tubo estéril. Após incubação a 37°C por 1 hora, sob agitação, plaqueamos as células em meio seletivo e incubamos a 37°C, 200 rpm por 16 h, para obtenção de colônias. Modificado de [64].

#### 4.4. Estratégias de clonagem gênica

##### 4.4.1. Clonagem da p18

O cDNA da p18 foi clonado no vetor pGEM T (Promega) e a partir deste clone foi realizada uma subclonagem no vetor de expressão pET28a (Novagen). A proteína p18 está localizada no cromossomo 11, ORF 59 de mamíferos. Para desenho dos oligonucleotídeos utilizados nesta clonagem, obtivemos a sequência da proteína no site do NCBI (National Center for Biotechnology Information – <http://blast.ncbi.nlm.nih.gov>) e acrescentamos às suas extremidades, sequências específicas para as enzimas de restrição *NdeI* e *XhoI* (**Tabela 4**). A reação de PCR foi realizada a partir da biblioteca de cérebro fetal humana clontech. Foi realizado um ciclo de desnaturação inicial de 5 minutos a 95°C seguido de trinta ciclos de desnaturação a 95°C por dois minutos, anelamento a 55°C por dois minutos e extensão a 72°C por um minuto. Ao final, foi feita uma extensão de 7 minutos a 72°C. Foi usada a enzima Taq Platinum (Invitrogen). O produto de PCR foi purificado e à sua extremidade A (inseridas pela enzima TAQ Platinum), foi ligado à extremidade T do vetor pGEM T (Promega) através de uma reação de ligação com a enzima T4 DNA ligase e seu tampão, com incubação em um banho a 16°C por 16 h. O plasmídeo resultante da ligação foi transformado em células de *E. coli* DH5 $\alpha$  e a clonagem foi confirmada por sequenciamento. O inserto foi retirado do pGEM (Promega) através da digestão com as enzimas *NdeI* e *XhoI* e subclonado no vetor pET28a (Novagen). A confirmação desta subclonagem foi realizada através da digestão do novo clone com as mesmas enzimas da clonagem, *NdeI* e *XhoI*.

**Tabela 4.** Desenho dos oligonucleotídeos utilizados na estratégia de clonagem do cDNA da proteína p18 no vetor pET28a.

<i>Oligonucleotídeo</i>	<i>Enzima de Restrição</i>	<i>Sequência utilizada</i>
<i>Forward</i>	<i>NdeI</i>	CccatATGGGGTGCTGCTACAGCAG
<i>Reverso</i>	<i>XhoI</i>	ccctcgagttaTGGGATCCCAAACGTACAACCAG

#### **4.4.2. Clones obtidos por colaboradores**

O grupo do pesquisador Dr. Tim Clausen (*Reserch Institute of Molecular Pathology, IMP, Viena, Áustria*) nos enviou clones para a co-expressão das proteínas MP1/p14. O clone 688 possui ambas as proteínas contidas no vetor pET28a (Novagen), sendo que apenas MP1 possui cauda de Histidina, enquanto que no clone 689, apenas p14 possui cauda de Histidina. O grupo do pesquisador Dr. Miroslawler (*Biotechnology Research Institute, NRCC, Quebec, Canada.*) nos forneceu um clone de MP1/p14 no vetor bicistrônico pETDuet-1 (Novagen), onde ambas as proteínas contêm caudas de Histidina.

Os clones contendo FtsZ e MinC foram obtidos através de uma colaboração com prof. Frederico J. Gueiros Filho, da Universidade de São Paulo – SP (Processo FAPESP 2010/51866-0 Smolbnet 2.0). A sequência referente à MinC foi inserida no vetor pET24a (Novagen) e a referente a FtsZ no vetor pET28a (Novagen), assim como seus mutantes.

#### **4.5. Eletroforese em gel de agarose**

Preparou-se um gel de agarose 0,8% com 40 µL de Brometo de Etídeo para cada 100 mL do tampão TAE. As amostras são aplicadas em conjunto com o tampão de amostra (*loading buffer*) e submetidas a uma voltagem de 80 volts pelo tempo necessário para que as bandas fiquem separadas. Modificado de [64].

#### **4.6. Expressão de proteínas recombinantes em *E.coli*.**

##### **4.6.1. Expressão do clone de p18 em pET28a**

Foi realizado um teste de expressão da proteína p18 em diversas cepas de *E. coli* (**Tabela 3**). O pré inóculo foi preparado em meio LB, acrescido de canamicina (50 µg/mL) e incubado a 37°C por 16h a 200 rpm. No dia seguinte, acrescentou-se 60 mL do pré inóculo a 2 L de meio LB e com as mesmas concentrações de canamicina. As células cresceram a 37°C e 200 rpm até uma D.O.<sub>600</sub> próxima de 1, onde foi adicionado 0,5 mM de IPTG e as diferentes células foram incubadas nas condições citadas na **Tabela 5**.

**Tabela 5** – Condições de expressão testadas para o clona da p18 (p18 + pET28a )

Cepa de <i>E. coli</i>	Temperatura de indução (°C)	Tempo de indução (h)
BL21 (DE3) ΔSlyD*	25, 30 e 37*	4 e 16*
BL21 (DE3) ΔSlyD pRARE	25, 30 e 37	4 e 16
BL21 (DE3) C41	18	12

\* Condição de escolha para todas as expressões seguintes.

As células foram coletadas a 4630  $\chi$ g por 20 minutos 4°C. O “*pellet*” de células foi ressuspenso em 10 mL de tampão A (50 mM de tampão fosfato de sódio pH 7,2, 100 mM de NaCl, 5% glicerol) acrescido de 0,1 ug/ml de lisozima e 1 mM PMSF, e a suspensão de células foi incubada em gelo por 30 minutos, para ação da lisozima. Após incubação, a amostra foi submetida à sonicação (sonicador *Sonics*<sup>®</sup>) em seis ciclos de pulsos de 15 segundos com 30% de amplitude. Em seguida, a amostra foi submetida à centrifugação a 21700  $\chi$ g por 10 minutos 4°C. O sobrenadante foi reservado e a fração insolúvel (*pellet*) foi ressuspenso em tampão A acrescido de 2% do detergente Triton-X 100, sendo esta preparação novamente sonicada (em três ciclos de pulsos de 15 segundos com 30% de amplitude) e centrifugada. Este ciclo de sonicação/centrifugação se repetiu por mais duas vezes até a obtenção da última fração insolúvel (fração correspondente aos corpos de inclusão).

Os procedimentos realizados neste trabalho utilizaram a proteína p18 preparada a partir da expressão na cepa de *E. coli* (DE3) ΔSlyD, sob uma indução com 0,5 mM de IPTG a 37°C, sob agitação de 200 rpm por 16h.

#### 4.6.2. Diálise e purificação da p18

Os corpos de inclusão referentes a 500 mL de indução de p18 em LB foram ressuspenso em 100 mL de tampão contendo 8 M de uréia, 50 mM de Tris-HCl pH 8,8, 50 mM de NaCl e outros aditivos, como mostrado na **Tabela 6**. Esta solução foi colocada em uma membrana de diálise SPECTRA/POR 3<sup>®</sup> de poros de 3,5 kDa. A membrana foi adicionada a um béquer de 5 L que continha 2 L de tampão (Tris-HCl pH 8,8, 6M de uréia e aditivos). Os tampões foram testados separadamente, variando a concentração do tampão e dos aditivos, como mostrado na **Tabela 6**. A diálise ocorreu a 4°C (em uma câmara fria), sob agitação constante e, após intervalos que variaram entre 4 e 16 h, 1 L de tampão era

trocado por um novo tampão, desta vez contendo todos os componentes do tampão inicial exceto a uréia. Desta forma a uréia foi gradativamente diluída, e a proteína reenovelada. Neste protocolo, toda a uréia foi retirada.

**Tabela 6** – Condições de tampão variadas na diálise da p18 a 4°C

<i>Condição</i>	<i>Tampão de ressuspensão do “pellet”</i>	<i>Tampão utilizado na primeira etapa da diálise</i>
1	Uréia 8 M, 50 mM e Tris-HCl pH 8,8	Uréia 6 M, 50 mM e Tris-HCl pH 8,8
2	Uréia 8M, 50 mM Tris-HCl pH 8,8 e 50 mM NaCl	Uréia 6 M, 50 mM Tris-HCl pH 8,8 e 50 mM NaCl
3	Uréia 8 M, 50 mM Tris-HCl pH 8,8 e 150 mM NaCl	Uréia 6 M, 50 mM Tris-HCl pH 8,8 e 150 mM NaCl
4	Uréia 8 M, 50 mM Tris-HCl pH 8,8 e 0,1 M $\beta$ ME	Uréia 6 M, 50 mM Tris-HCl pH 8,8 e 0,1 M $\beta$ ME
5	Uréia 8 M, 50mM Tris-HCl pH 8,8, 150 mM NaCl, 5 mM $\beta$ ME, 0,1% Tween 20 e 5% glycerol	Uréia 6 M, 50 mM Tris-HCl pH 8,8, 150 mM NaCl, 5 mM $\beta$ ME, 0,1% Tween 20 e 5% glycerol
6	Uréia 8 M, 50 mM Tris-HCl pH 8,8, 50 mM Arg/Glu e 150 mM NaCl	Uréia 6 M, 50 mM Tris-HCl pH 8,8, 50 mM Arg/Glu e 150 mM NaCl
7*	Uréia 8 M, 50 mM Tris-HCl pH 8,8, 50 mM NaCl e 0,5 mM TCEP	Uréia 6 M, 50 mM Tris-HCl pH 8,8, 50 mM NaCl e 0,5 mM TCEP

\* Condição em que foi adicionado o complexo MP1/p14 na etapa da diálise onde o tampão estava com 1 M de uréia. Foram realizadas 11 trocas de metade do volume do tampão de forma que o tampão e todos os aditivos, exceto a uréia, foram mantidos. As concentrações de uréia diminuíram gradativamente e mais duas trocas foram realizadas onde todo o tampão (os 2 L contidos no bécquer) foi trocado.

A condição 7 da **Tabela 6** foi utilizada para a diálise de p18 com adição de MP1/p14. Como citado acima, após sucessivas trocas de tampão as concentrações de uréia foram gradativamente sendo diminuídas e, na etapa em que a concentração de uréia fosse 1 M, foi adicionado um excesso de MP1/p14 junto à membrana de diálise que continha a p18.

O produto de cada uma das diálises foi purificado por afinidade (coluna Histrap FF de 5 mL – GE Healthcare). Para isso utilizou-se um tampão de carregamento da amostra contendo 20 mM de Tris-HCl pH 6,5, 100 mM de NaCl e 0,5 mM de TCEP além de um tampão de eluição igual ao tampão de carregamento da amostra, acrescido de 1 mM de imidazol. A corrida foi realizada em um AKTA® FPLC™, em um fluxo de 1 mL/min, utilizando 10 volumes de coluna para uma eluição com um único passo, em um gradiente linear de zero a 100% de tampão de eluição.

As amostras contendo MP1/p14 e p18 foram concentradas em filtros ultra concentradores para centrífuga com poros de 10 kDa (Amicon-Millipore®). Estas amostras foram utilizadas para realização de *cross-linking* ou aplicadas à cromatografia de exclusão por tamanho molecular.

A cromatografia de exclusão por tamanho molecular também foi realizada em um AKTA® FPLC™, com um fluxo de 1 mL/min, utilizando de 1 a 1.5 volume de coluna para uma eluição sob o tampão 20 mM de Tris-HCl pH 6,5, 100 mM de NaCl. A coluna utilizada foi a Superdex 200 16/60 (GE Healthcare®).

O mesmo protocolo de purificação foi realizado para a p18 oriunda das diálises nos tampões de 1 a 6 da **Tabela 6**, onde não foram adicionadas MP1/p14.

#### **4.6.3. Expressão e purificação dos clones de MP1/p14**

O clone de MP1/14 + pETDuet-1 foi transformado em células de *Escherichia coli* BL21(DE3). Uma colônia foi acrescentada a 60 mL de LB contendo 50 µg/mL de canamicina. Estas células foram crescidas com a 37°C e 200 rpm por 16h. O pré inóculo foi adicionado a 2 L de LB e incubado a 37°C e 200 rpm até uma D.O.<sub>600</sub> 0,6, quando as células foram induzidas a expressar as proteínas com 1 mM IPTG, a 20°C e 200 rpm por 16h [31]. Em seguida, coletamos a indução centrifugando-a a 4°C, 4630  $\chi$ g por 20 minutos.

Os clones de MP1/p14 688 e 689 foram expressos em células de *E. coli* (DE3)  $\Delta$ SlyD. O pré inóculo foi preparado em meio LB, acrescido de antibiótico (**Tabela 7**) e incubado a 37°C e 200 rpm por 16h. No dia seguinte, acrescentou-se 60 mL do pré inóculo a 2 L de meio LB e com as mesmas concentrações de antibiótico. As células cresceram a 37°C e 200 rpm até uma D.O.<sub>600</sub> próxima de 1, quando foi adicionado 0,5 mM de IPTG e incubadas a 37°C a 200 rpm por 16h [30].

Os clones 688 e 689 foram expressos separadamente, enquanto que o clone MP1/p14 + pETDuet-1 foi utilizado na tentativa de co-expressão com p18.

Após indução, o inóculo foi centrifugado (4630  $\chi$ g, 20 min, 4°C) e as células foram resussuspensas em 20 mM de tampão fosfato pH 6,5, 20 mM de NaCl, 1  $\mu$ M de PMSF e lisozima em concentração final de 100  $\mu$ g/ml. Os extratos foram sonicados em sonicador Sonics® com seis ciclos de pulsos de 15 segundos com 30% de amplitude. Após sonicação, a amostra foi centrifugada a 21700  $\chi$ g por 20 minutos a 4°C. O sobrenadante obtido corresponde ao lisado de células que foi posteriormente purificado.

**Tabela 7 – Clones utilizados na expressão dos clones de MP1/p14**

Laboratório Responsável pela elaboração do Clone	Clone	Cauda de Histidina	Antibiótico utilizado (50 $\mu$ g/mL)
Mirosław Cygler	MP1/p14 + pETDuet	MP1 e p14	Ampicilina
Tim Clausen	His-MP1/p14 + pET28a (688)	MP1	Canamicina
	MP1/p14-His + pET28a (689)	P14	

Ambos os clones de MP1/p14 foram purificados da mesma forma que a proteína p18. O método de captura escolhido foi a afinidade (coluna His Trap FF de 5 mL – GE Healthcare®). O tampão de carregamento da amostra continha 20 mM de Tris-HCl pH 6,5 e 100 mM de NaCl, assim como o tampão de eluição, porém acrescido de 1 mM de imidazol. A corrida foi realizada em um AKTA® FPLC™, em um fluxo de 1 mL/min, utilizando 10 volumes de coluna para uma eluição com um único passo, em um gradiente linear de zero a 100% de tampão de eluição.

As amostras contendo MP1/p14 foram concentradas em filtros ultra concentradores para centrífuga com poros de 10 kDa (Amicon-Millipore®). Estas amostras foram aplicadas à cromatografia de exclusão molecular.

A cromatografia de exclusão por tamanho molecular também foi realizada em um AKTA® FPLC™, com um fluxo de 1 mL/min, utilizando de 1 a 1.5 volume de coluna para uma eluição sob o tampão 20 mM de Tris-HCl pH 6,5 e 100 mM de NaCl. A coluna utilizada foi a Superdex 200 16/60 (GE Healthcare®).

As amostras contendo MP1/p14 foram novamente concentradas em filtros ultra concentradores para centrífuga com poros de 10 kDa (Amicon-Millipore®) e utilizadas em ensaios biofísicos ou acrescidas na diálise da p18 para formação do complexo *Ragulator*.

#### 4.6.4. Expressão e purificação das proteínas *FtsZ* e *MinC*

*FtsZ* e *MinC* foram expressas em células de *E. coli* (DE3) pRIL. A expressão destas proteínas de *Bacillus subtilis* foi realizada em meio LB acrescido de canamicina em uma concentração final de 30 µg/mL e cloranfenicol (50 µg/mL). Foram utilizadas as construções plasmidiais referidas na **Tabela 8**, conferindo as estas proteínas a presença de uma cauda de Histidina. O pré inóculo foi incubado a 37°C, 200 rpm por 16h. O inóculo contendo 3% de pré inóculo e a mesma concentração de antibióticos citada acima foi incubado a 37°C, sob uma agitação se 200 rpm e induzidas em uma D.O.<sub>600</sub> próxima a 1 com 1 mM de IPTG. A indução foi incubada a 37°C, 200 rpm por 4h.

Os mutantes de *FtsZ* seguiram o mesmo protocolo de expressão e purificação que a proteína selvagem

**Tabela 8 – Clones utilizados na expressão de *FtsZ* e *MinC***

Clone	Vetor	Cepa de <i>E. coli</i> utilizada	Antibiótico utilizado (µg/mL)
MinC	pET24a	<i>E. coli</i> (DE3)-RIL	30 Kan + 50 Cl
FtsZ selvagem e mutantes	pET28a	<i>E. coli</i> (DE3)-RIL	30 Kan + 50 Cl
		<i>E. coli</i> (DE3)-C43	30 Kan

Kan = Canamicina e Cl = Cloranfenicol

Ao termino da indução, as células foram coletadas por centrifugação a 4630  $\chi$ g por 20 minutos a 4°C. O “pellet” de células foi ressuspenso em um tampão contendo 50 mM de tampão fosfato de sódio pH 6,5, 50 mM de KCl, 50 mM de Arginina, 50 mM de ácido glutâmico, 1 mM de EDTA, 5% de glicerol, 1 µM de PMSF e 100 µg/mL de lisozima. A amostra foi levada por 5 minutos a um banho térmico a 37°C, para melhor atuação enzimática. Em seguida, a amostra permaneceu por 15 minutos em um banho de gelo e posteriormente as células foram lisadas em um sonicador Sonics® em três ciclos de pulsos

de 30 segundos com 35% de amplitude. A amostra foi então submetida a uma centrifugação a 21700  $\chi$ g, 4°C por 45 minutos e a fração solúvel (contendo a proteína) foi submetida à purificação.

#### **4.6.4.1. Purificação da MinC**

O método de captura escolhido foi a afinidade (coluna Histrap FF de 5 mL – GE Healthcare®). O tampão de carregamento da amostra continha 50 mM de tampão fosfato de sódio pH 6,5, 50 mM de KCl, 50 mM de Arginina, 50 mM de ácido glutâmico, 1 mM de EDTA e 5% de glicerol, assim como o tampão de eluição, porém acrescido de 1 mM de imidazol. A corrida foi realizada em um AKTA® FPLC™, em um fluxo de 1 mL/min, utilizando 10 volumes de coluna para uma eluição com um único passo, em um gradiente linear de zero a 100% de tampão de eluição.

As amostras contendo MinC foram concentradas em filtros ultra concentradores para centrífuga com poros de 10 kDa (Amicon-Millipore®). Estas amostras foram aplicadas à cromatografia de exclusão molecular.

A cromatografia de exclusão por tamanho molecular também foi realizada em um AKTA® FPLC™, com um fluxo de 1 mL/min, utilizando de 1 a 1.5 volume de coluna para uma eluição sob o tampão 50 mM de tampão fosfato de sódio pH 6,5, 50 mM de KCl, 50 mM de Arginina, 50 mM de ácido glutâmico, 1 mM de EDTA e 5% de glicerol. A coluna utilizada foi a Superdex 200 16/60 (GE Healthcare®).

As amostras contendo MinC foram novamente concentradas em filtros ultra concentradores para centrífuga com poros de 10 kDa (Amicon-Millipore®) e utilizadas em ensaios biofísicos ou espectroscópicos.

#### **4.6.3.1. Purificação da FtsZ e mutantes**

FtsZ foi purificada por precipitação com sulfato de amônio e este foi retirado através de uma coluna do tipo PD-10 (Sephadex G25 – GE Healthcare®). À proteína eluída, foi adicionado  $\text{Ca}^{+2}$  20 mM e GTP 1mM, seguido de incubação a 37°C por 5 minutos, polimerizando-a e permitindo a seleção de proteínas ativas. A amostra foi centrifugada e o precipitado foi ressuspenso em 3mL de tampão fosfato 50 mM pH 6,5, 50 mM, KCl 50 mM de Arginina, 50 mM de ácido glutâmico, 1 mM de EDTA e 5% de glicerol e submetido a

filtração em gel. As frações foram posteriormente analisadas por gel de poliacrilamida (SDS-PAGE colorido com *coomassie blue*) para confirmação.

Em alguns experimentos incubamos a amostra com GDP e magnésio e posteriormente utilizamos EDTA, favorecendo assim a forma monomérica de FtsZ.

A cromatografia de exclusão por tamanho molecular, ou simplesmente filtração em gel, também foi realizada no AKTA<sup>®</sup> FPLC<sup>™</sup> – GE Healthcare<sup>®</sup>, e os dados foram analisados no UNICORE 5.2. A coluna utilizada foi a Superdex Hiload 200 16/60 (GE Healthcare). Em geral, as corridas ocorreram sob um fluxo de 1 mL/minuto por 180 mL (1,5 volume de coluna).

As amostras contendo FstZ foram concentradas em filtros ultra concentradores para centrífuga com poros de 10 kDa (Amicon-Millipore<sup>®</sup>) e utilizadas em ensaios biofísicos ou espectroscópicos.

#### **4.6.5. Expressão e purificação de proteínas FtsZ e MinC em meio mínimo M9**

A forma como estas proteínas foram expressas em meio mínimo M9 [63] possui algumas diferenças em relação à expressão em LB. O inóculo, mas não o pré inóculo, foi realizado em meio mínimo M9 [63] ao invés de LB. Para evitar contaminação do M9 [63] com LB, centrifugamos o pré inóculo em tubos estéreis, a 37°C, 4630  $\chi$ g por 20 minutos, descartamos o sobrenadante em ambiente estéril e ressuspendemos o “*pellet*” de células em meio mínimo.

As condições de indução também foram ligeiramente alteradas, as proteínas foram expressas a 25°C a 200 rpm por 16 h.

#### **4.7. Eletroforese em gel de poliacrilamida (SDS-PAGE).**

As amostras protéicas foram misturadas a um tampão de amostra 2X [64], fervidas a 96°C por 5 minutos e aplicadas a um gel de acrilamida SDS-PAGE. Realizamos a eletroforese entre 100 V e 200 V, o coramos com *coomassie blue* por 30 minutos à temperatura ambiente e em seguida descoramos com uma solução contendo 30% de metanol e 10% de ácido acético.

#### **4.8. Estudos Biofísicos**

Os ensaios biofísicos foram realizados no Laboratório de Espectrometria e Calorimetria do LNBio.

##### **4.8.1. Espalhamento de Luz Dinâmico – DLS**

O raio hidrodinâmico da proteína p18 renaturada na condição de tampão 6 (**Tabela 6**) foi medido utilizando o instrumento *DynaPro™ Plate Reader™*, da Wyatt sob um ângulo de 90°, a 824.2 nm com laser de 100% de intensidade, onde foram coletadas 10 aquisições de 5 segundos para cada amostra. Neste experimento, testamos as diferentes temperaturas 20°C, 25°C, 37°C, repetindo a medida de 25°C ao final para avaliar se houve alguma alteração na amostra após a medida a 37°C. As análises foram feitas no programa *DYNAMICSTM* - Versão 6.12.0.3.

A triagem foi realizada em 96 condições diferentes, onde se combinaram aditivos (5 mM  $\beta$ -ME, 1 mM DTT, 20% Glicerol, % Tween 20, 0,1% Igepal, 50 mM Arginina/ Ácido Glutâmico e NaCl nas concentrações 50 mM, 150 mM, 200 mM, 250 mM, 300 mM, 400 mM e 500 mM) e tampões a 100 mM com diferentes pHs ( 5,5 Tampão citrato; 6,0 Tampão Fosfato; 6,5 Tampão ADA; 7,0 Tampão Fosfato; 7,5 Tampão Hepes; 8,0 Tampão Imidazol; 8,5 Tampão Tris-HCl e 9,0 Tampão Ches).

##### **4.8.2. Dicroísmo Circular – CD**

Esta técnica permite análise da estrutura secundária de proteínas em solução e é amplamente usada para validação de aspectos conformacionais. Os ensaios foram realizados em um equipamento J-810 (JASCO) acoplado a um sistema interno de controle de temperatura (*Peltier Type Control System PFD 425S – JASCO*) e as medidas foram realizadas a temperatura constante de 20°C. O programa utilizado para registro de dados foi o *Spectra Manager®* (JASCO). Os dados foram coletados na faixa de 195 a 260 nm (em alguns espectros esta faixa se estreitou a fim de que a voltagem ficasse abaixo de 700 volts) e a cubeta de quartzo utilizada possuía 1 mm de caminho óptico.

As proteínas foram analisadas com a finalidade de investigação de suas estruturas secundárias. Foram ainda realizadas desnaturações térmicas de 10°C a 90°C, sendo coletadas 5 aquisições a cada 5°C. As desnaturações térmicas ocorreram tanto com as proteínas sozinhas, quanto adicionadas a sua parceira. No caso de p18 e MP1/p14 foi realizado espectro de ambas sozinhas e das mesmas juntas. No caso de FtsZ e MinC foram realizadas medidas de cada uma delas (incluindo os mutantes) e de MinC adicionada separadamente a FtsZ selvagem e mutantes.

### ***4.9. Microscopia Eletrônica***

As análises microscópicas foram realizadas no Laboratório Nacional de Nanotecnologia – LNNano/CNPEM, no Microscópio Eletrônico de Transmissão TEM - FEG (JEM 2100F), com colaboração dos pesquisadores Dr. Rodrigo Villares Portugal e Dr. Jefferson Bettini. A amostra da p18 foi dialisada com e tratada com acetato de uranila, um material pesado utilizado como contraste ou “coloração negativa”, o que permite uma acentuação apreciável do contraste entre a solução e a proteína.

Em uma primeira análise foi utilizada apenas uma amostra de p18, que foi dialisada no tampão 5 (**Tabela 6**) e se encontrava com 50 mM Tris-HCl pH 8,8, 150 mM NaCl, 5 mM  $\beta$  ME, 0,1% Tween 20 e 5% glicerol.

Foi realizada uma nova análise de microscopia eletrônica da p18 renaturada, desta vez com o tampão 2 (**Tabela 6**). Foram retiradas alíquotas em todas as etapas da diálise, tanto da amostra, quanto do tampão presente no béquer. As amostras foram observadas na microscopia eletrônica a fim de se investigar o grau de agregação da p18.

### ***4.10. Espectros de RMN***

Os experimentos de RMN foram realizados no espectrômetro de 600 MHz no laboratório de RMN do LNBio.

Os dados foram coletados no espectrômetro de RMN de 600 MHz com sonda criogênica, e analisados com os programas NMRDraw (NMRPipe) e NMRview. Foram realizados experimentos do tipo 1D e 2D ( $^{15}\text{N}$ -HSQC) [65]. O experimento de prótons de

Hidrogênio 1D foi realizado afim de verificar se a proteína se encontrava estruturada e experimento 2D ( $^{15}\text{N}$ -HSQC) [65] foi realizado para identificar interações entre proteínas. Para isso, uma proteína era marcada com  $^{15}\text{N}$  e a outra proteína, expressa em LB, era titulada em diferentes proporções. Os sinais NH da cadeia principal de proteína marcada que apresentaram mudanças em seu deslocamento químico durante a titulação com a proteína ligante foram responsáveis pela confirmação da interação entre as proteínas. Estes sinais correspondem aos aminoácidos da proteína marcada que interagiram com a proteína expressa em LB.

### ***4.11. Cross-linking do Ragulator***

A amostra proveniente da diálise sob a condição 7 da **Tabela 6**, contendo MP1/p14 e p18, foi submetida a cromatografia de afinidade a Ni. Uma das frações, ao invés de ser aplicada a filtração em gel, foi submetida a uma reação de *Cross-linking*.

O experimento de *Cross-linking* foi realizado adicionando-se o *cross-linker* suberato de N-N'-disuccinimidila (DSS) dissolvido em dimetil formamida (1 mg/mL) às proteínas do complexo *Ragulator* em uma razão equimolar. A durante 4 h a reação ficou sob agitação, sendo realizada uma reação a temperatura ambiente e outra a 4 h.

Decorrido este tempo, a reação foi finalizada com tampão de corrida e submetida a uma análise em gel SDS-PAGE corado com *Comassie Blue*.

# Capítulo 5

## Conclusões

## 5. CONCLUSÕES

### *O complexo Ragulator*

Para solubilização da proteína p18, foi necessária a padronização de uma diálise onde a concentração de uréia foi gradativamente retirada. Diversas diálises da p18 sozinha foram testadas e ela só pode ser obtida solúvel na forma de grandes agregados, exceto em uma única condição onde foi obtida uma quantidade mínima de p18 que supomos ser monomérica por se ligar à coluna de Ni, entretanto esta condição ainda necessita ser confirmada. A adição de seus parceiros, MP1/p14 possibilitou a obtenção de um maior rendimento da proteína p18 na suposta condição monomérica.

O fato de p18 apresentar predição estrutural de ser intrinsecamente desordenada pode teoricamente estar diminuindo a estabilidade desta proteína, por ser um alvo mais frequente de proteases. A grande instabilidade desta proteína dificultou estudos biofísicos e estruturais.

A presença de duas Cisteínas na proteína p18 pode auxiliar a formação de agregados protéicos por pontes de dissulfeto. A adição do agente redutor TCEP à diálise e aos tampões de purificação auxiliou a evitar a formação destes agregados.

A co-expressão da p18 com MP1/p14 (pETDuet) foi realizada pela aluna de iniciação científica Débora Camilotti. Com esta co-expressão esperávamos obter a p18 solúvel ligada aos seus parceiros MP1/p14, entretanto apenas MP1/p14 foi encontrada na fração solúvel (dados não mostrados).

Os agregados formados pela p18 foram caracterizados por espalhamento de luz dinâmico e microscopia eletrônica, e esta caracterização foi crucial para aumentar o rendimento da produção p18 em sua possível forma monomérica através de diálise. A proteína p18 agregada é incapaz de se ligar a seus parceiros MP1/p14 e a caracterização do agregado nos permitiu encontrar um ponto em que a p18 não se encontrava na forma de fibras em uma concentração de uréia mínima, a fim de manter as proteínas MP1/p14 complexadas.

Os resultados obtidos por CD sugerem uma variação no perfil de estrutura secundária entre a proteína p18 em sua forma supostamente “monomérica” e em sua forma

agregada. A p18 supostamente “monomérica” renaturada com e sem MP1/p14 apresentam espectros de CD semelhantes. O espectro de CD contendo p18 agregada, somada a MP1/p14 não sugere interação entre estas proteínas.

O complexo *Ragulator* apresenta um alto grau de dissociação [35] o que dificulta sua caracterização por RMN. A fim de aumentar a estabilidade do complexo foi realizado um *cross-linking*, onde foram formadas ligações covalentes entre as aminas primárias livres das cadeias laterais das proteínas próximas. Em um gel SDS-PAGE pode ser observado que apenas o heterodímero MP1/p14 foi estabilizado. Se as ligações covalentes tivessem sido formadas entre as três proteínas do *Ragulator*, o *cross-linking* seria analisado por espectrometria de Massas, onde seria possível caracterizar a região das proteínas onde ocorre a interação.

### ***A proteína FtsZ e sua moduladora MinC***

Foi realizada a padronização da expressão e purificação destas proteínas em meio mínimo M9, onde foi possível marcá-las com  $^{15}\text{N}$ .

Estudos de SAXS realizados por nosso colaborador Dr. Wanius José Garcia da Silva (UFABC), não apresentados neste trabalho, sugerem que a proteína MinC se encontra tetramérica e FtsZ dimérica. Nesta conformação o complexo apresentaria um peso molecular de aproximadamente 160 kDa, ou seja, muito além do limite possível para análise detalhada de cada um dos resíduos de aminoácidos da proteína por RMN.

Foi possível observar, por Dicroísmo Circular, diferenças entre as estruturas secundárias de FtsZ selvagem e seus mutantes pontuais, que foram divididos em três perfis de espectro, onde FtsZ e o mutante 111 apresentaram um perfil muito semelhante, típico de proteínas com maior conteúdo de  $\alpha$  hélice. Os mutantes de FtsZ 255, 376 apresentam um maior conteúdo de folhas  $\beta$  quando comparado com FtsZ selvagem, e o mutante FtsZ $\Delta$ C apresenta uma conformação intermediária entre os demais perfis. O desenovelamento térmico analisado por Dicroísmo Circular não confirmou a interação entre MinC e FtsZ Selvagem, FtsZ 111, FtsZ 255, FtsZ 376 e FtsZ  $\Delta$ C.

A mudança de alguns picos (resíduos de aminoácidos) no espectro  $^{15}\text{N}$ -HSQC da proteína MinC, quando titulada por FtsZ, confirmou uma interação, ainda que fraca, entre

estas proteínas. A identificação precisa dos picos/resíduos de aminoácidos envolvidos nesta interação foi impossibilitada, devido à aglomeração de picos centrais que prejudicaram a interpretação do espectro. Estes agregados se formaram devido ao alto peso molecular do complexo.

## 6. PERSPECTIVAS

### *O complexo Ragulator*

A região de interação entre p18 e MP1/p14 ainda não é conhecida detalhadamente. A construção do clone de p18 utilizado neste trabalho deu origem à proteína p18 inteira, entretanto instável com tendência a formar agregados. A co-expressão deste clone com MP1/p14 não apresentou uma melhora significativa a este quadro. Entretanto em um artigo publicado recentemente pelo pesquisador Dr. Mirosław Cygler, foi possível obter as três proteínas solúveis e complexadas através de uma construção em que as regiões N e C terminal de p18 foram deletadas, preservando apenas os resíduos de aminoácido 47 a 138 [35]. Este clone foi solicitado e com ele pretendemos realizar experimentos com a intenção de mapear, através de cross-linking e Espectrometria de Massas, os resíduos envolvidos na interação entre as proteínas do complexo *Ragulator*.

### *A proteína FtsZ e sua moduladora MinC*

Após confirmar a interação entre as proteínas de *Bacillus subtilis* MinC e FtsZ, pretendemos mapear os resíduos de aminoácidos envolvidos na interação, o que é crucial para o entendimento do papel destas proteínas na divisão celular bacteriana. Para isso será realizada a marcação das proteínas FtsZ e MinC, separadamente, com deutério ( $^2\text{H}$ ) e esta marcação nos possibilitará realizar experimentos HSQC-TROSY, que possivelmente melhorará o perfil de sinais agregados no centro do espectro.

A proteína MinC será dividida em duas metades (C e N terminal) e titulada com FtsZ selvagem e mutantes. A diminuição do tamanho da MinC e possivelmente o fato de

que apenas uma metade de MinC não polimerize, aumentam significativamente a resolução do espectro.

Adicionalmente serão realizados estudos com os mutantes de FtsZ e alguns mutantes de MinC que o grupo do prof. Dr. Frederico Gueiros Filho já está desenvolvendo. A reunião de todos estes dados pode nos fornecer informações preciosas sobre o papel de MinC e FtsZ na divisão celular.

## 6. REFERÊNCIAS BIBLIOGRÁFICAS

1. Haydon, D.J., et al., *An inhibitor of FtsZ with potent and selective anti-staphylococcal activity*. Science, 2008. **321**(5896): p. 1673-5.
2. Laplante, M. and D.M. Sabatini, *mTOR signaling at a glance*. J Cell Sci, 2009. **122**(Pt 20): p. 3589-94.
3. Hay, N. and N. Sonenberg, *Upstream and downstream of mTOR*. Genes Dev, 2004. **18**(16): p. 1926-45.
4. Sabatini, D.M., *mTOR and cancer: insights into a complex relationship*. Nat Rev Cancer, 2006. **6**(9): p. 729-34.
5. Stan, R., et al., *Interaction between FKBP12-rapamycin and TOR involves a conserved serine residue*. J Biol Chem, 1994. **269**(51): p. 32027-30.
6. Brown, E.J., et al., *A mammalian protein targeted by G1-arresting rapamycin-receptor complex*. Nature, 1994. **369**(6483): p. 756-8.
7. Sabatini, D.M., et al., *RAFT1: a mammalian protein that binds to FKBP12 in a rapamycin-dependent fashion and is homologous to yeast TORs*. Cell, 1994. **78**(1): p. 35-43.
8. Kim, D.H., et al., *mTOR interacts with raptor to form a nutrient-sensitive complex that signals to the cell growth machinery*. Cell, 2002. **110**(2): p. 163-75.
9. Sarbassov, D.D., et al., *Rictor, a novel binding partner of mTOR, defines a rapamycin-insensitive and raptor-independent pathway that regulates the cytoskeleton*. Curr Biol, 2004. **14**(14): p. 1296-302.
10. Hara, K., et al., *Raptor, a binding partner of target of rapamycin (TOR), mediates TOR action*. Cell, 2002. **110**(2): p. 177-89.
11. Guertin, D.A., et al., *Ablation in mice of the mTORC components raptor, rictor, or mLST8 reveals that mTORC2 is required for signaling to Akt-FOXO and PKCalpha, but not S6K1*. Dev Cell, 2006. **11**(6): p. 859-71.
12. Fang, Y., et al., *Phosphatidic acid-mediated mitogenic activation of mTOR signaling*. Science, 2001. **294**(5548): p. 1942-5.
13. Porstmann, T., et al., *SREBP activity is regulated by mTORC1 and contributes to Akt-dependent cell growth*. Cell Metab, 2008. **8**(3): p. 224-36.
14. Kim, J.E. and J. Chen, *regulation of peroxisome proliferator-activated receptor-gamma activity by mammalian target of rapamycin and amino acids in adipogenesis*. Diabetes, 2004. **53**(11): p. 2748-56.
15. Timothy R. Peterson, S.S.S., Thurl E. Harris, Anne E. Carmack, Seong A. Kang, Eric Balderas, and K.L.M. David A. Guertin, Anne E. Carpenter, Brian N. Finck, and David M. Sabatini, *mTOR Complex 1 Regulates Lipin 1 Localization to Control the SREBP Pathway*. Cell, 2011. **146**(3): p. 408-420.
16. Sarbassov, D.D., et al., *Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex*. Science, 2005. **307**(5712): p. 1098-101.
17. Sancak Y, B.-P.L., Zoncu R, Markahard AL, Nada S and Sabatini, DM., *Ragulator-RagComplex Targets mTORC1 to the Lysosomal Surface and Is Necessary for Its Activation by Amino Acids*. Cell, 2010. **141**: p. 290-303.
18. Kim, E., et al., *Regulation of TORC1 by Rag GTPases in nutrient response*. Nat Cell Biol, 2008. **10**(8): p. 935-45.

19. Sancak, Y., et al., *The Rag GTPases bind raptor and mediate amino acid signaling to mTORC1*. Science, 2008. **320**(5882): p. 1496-501.
20. Avruch, J., et al., *Amino acid regulation of TOR complex 1*. Am J Physiol Endocrinol Metab, 2009. **296**(4): p. E592-602.
21. Takahashi, Y., Nada, S., Mori, S., Nagae, T., Oneyama, T. and Okada, M. , *The late endosome/lysosome-anchored p18-mTORC1 pathway controls terminal maturation of lysosomes*. Biochemical Biophysical Research Communications 2012.
22. Kalender A, S.A., Kim SY, Gulati P, Brûlé S, Viollet B, Kemp BE, Bardeesy N, Dennis P, Schlager JJ, Marette A, Kozma SC, Thomas G., *Metformin, independent of AMPK, inhibits mTORC1 in a rag GTPase-dependent manner*. Cell Metab. , 2010. **11**(5): p. 390-401.
23. Nicklin, P., et al., *Bidirectional transport of amino acids regulates mTOR and autophagy*. Cell, 2009. **136**(3): p. 521-34.
24. Jung, C.H., et al., *mTOR regulation of autophagy*. FEBS Lett, 2010. **584**(7): p. 1287-95.
25. Binda, M., et al., *The Vam6 GEF controls TORC1 by activating the EGO complex*. Mol Cell, 2009. **35**(5): p. 563-73.
26. Kogan, K., et al., *Structural conservation of components in the amino acid sensing branch of the TOR pathway in yeast and mammals*. J Mol Biol, 2010. **402**(2): p. 388-98.
27. Nakamura, N., et al., *Vam2/Vps41p and Vam6/Vps39p are components of a protein complex on the vacuolar membranes and involved in the vacuolar assembly in the yeast Saccharomyces cerevisiae*. J Biol Chem, 1997. **272**(17): p. 11344-9.
28. Dubouloz, F., et al., *The TOR and EGO protein complexes orchestrate microautophagy in yeast*. Mol Cell, 2005. **19**(1): p. 15-26.
29. Caplan, S., et al., *Human Vam6p promotes lysosome clustering and fusion in vivo*. J Cell Biol, 2001. **154**(1): p. 109-22.
30. Kurzbauer, R., et al., *Crystal structure of the p14/MPI scaffolding complex: how a twin couple attaches mitogen-activated protein kinase signaling to late endosomes*. Proc Natl Acad Sci U S A, 2004. **101**(30): p. 10984-9.
31. Lunin, V.V., et al., *The structure of the MAPK scaffold, MPI, bound to its partner, p14. A complex with a critical role in endosomal map kinase signaling*. J Biol Chem, 2004. **279**(22): p. 23422-30.
32. Nada, S., et al., *The novel lipid raft adaptor p18 controls endosome dynamics by anchoring the MEK-ERK pathway to late endosomes*. EMBO J, 2009. **28**(5): p. 477-89.
33. Qian, C., et al., *Structure of the adaptor protein p14 reveals a profilin-like fold with distinct function*. J Mol Biol, 2005. **347**(2): p. 309-21.
34. Davis, I.W., et al., *MolProbity: all-atom contacts and structure validation for proteins and nucleic acids*. Nucleic Acids Res, 2007. **35**(Web Server issue): p. W375-83.
35. Magee, J. and M. Cygler, *Interactions between kinase scaffold MPI/p14 and its endosomal anchoring protein p18*. Biochemistry, 2011. **50**(18): p. 3696-705.
36. Guillaumot, P., et al., *Pdro, a protein associated with late endosomes and lysosomes and implicated in cellular cholesterol homeostasis*. PLoS One 2010. **5**(6): p. e10977.

37. Dosztanyi, Z., et al., *IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content*. *Bioinformatics*, 2005. **21**(16): p. 3433-4.
38. Romero, Obradovic, and K. Dunker, *Sequence Data Analysis for Long Disordered Regions Prediction in the Calcineurin Family*. *Genome Inform Ser Workshop Genome Inform*, 1997. **8**: p. 110-124.
39. Li, X., et al., *Predicting Protein Disorder for N-, C-, and Internal Regions*. *Genome Inform Ser Workshop Genome Inform*, 1999. **10**: p. 30-40.
40. Romero, P., et al., *Sequence complexity of disordered protein*. *Proteins*, 2001. **42**(1): p. 38-48.
41. Marzzoco, A., Torres, B.B. , *Bioquímica Básica*. 2010.
42. Foss, M.H., Y.J. Eun, and D.B. Weibel, *Chemical-biological studies of subcellular organization in bacteria*. *Biochemistry*, 2011. **50**(36): p. 7719-34.
43. Rothfield, L., A. Taghbalout, and Y.L. Shih, *Spatial control of bacterial division-site placement*. *Nat Rev Microbiol*, 2005. **3**(12): p. 959-68.
44. Pichoff, S. and J. Lutkenhaus, *Tethering the Z ring to the membrane through a conserved membrane targeting sequence in FtsA*. *Mol Microbiol*, 2005. **55**(6): p. 1722-34.
45. Pichoff, S. and J. Lutkenhaus, *Unique and overlapping roles for ZipA and FtsA in septal ring assembly in Escherichia coli*. *EMBO J*, 2002. **21**(4): p. 685-93.
46. Addinall, S.G. and J. Lutkenhaus, *FtsZ-spirals and -arcs determine the shape of the invaginating septa in some mutants of Escherichia coli*. *Mol Microbiol*, 1996. **22**(2): p. 231-7.
47. Oliva, M.A., D. Trambaiolo, and J. Lowe, *Structural insights into the conformational variability of FtsZ*. *J Mol Biol*, 2007. **373**(5): p. 1229-42.
48. Bisson Filho, A.W., *Estudo genético da interação entre FtsZ e o modulador de divisão ZapA em Bacillus subtilis*. Instituto de Química, 2009. **Mestrado**.
49. Adler, H.I., et al., *MINIATURE escherichia coli CELLS DEFICIENT IN DNA*. *Proc Natl Acad Sci U S A*, 1967. **57**(2): p. 321-6.
50. de Boer, P.A., R.E. Crossley, and L.I. Rothfield, *A division inhibitor and a topological specificity factor coded for by the minicell locus determine proper placement of the division septum in E. coli*. *Cell*, 1989. **56**(4): p. 641-9.
51. Margolin, W., *Themes and variations in prokaryotic cell division*. *FEMS Microbiol Rev*, 2000. **24**(4): p. 531-48.
52. Levin, P.A., J.J. Shim, and A.D. Grossman, *Effect of minCD on FtsZ ring position and polar septation in Bacillus subtilis*. *J Bacteriol*, 1998. **180**(22): p. 6048-51.
53. Cha, J.H. and G.C. Stewart, *The divIVA minicell locus of Bacillus subtilis*. *J Bacteriol*, 1997. **179**(5): p. 1671-83.
54. Raskin, D.M. and P.A. de Boer, *The MinE ring: an FtsZ-independent cell structure required for selection of the correct division site in E. coli*. *Cell*, 1997. **91**(5): p. 685-94.
55. Hu, Z. and J. Lutkenhaus, *Topological regulation of cell division in Escherichia coli involves rapid pole to pole oscillation of the division inhibitor MinC under the control of MinD and MinE*. *Mol Microbiol*, 1999. **34**(1): p. 82-90.

56. Hu, Z. and J. Lutkenhaus, *A conserved sequence at the C-terminus of MinD is required for binding to the membrane and targeting MinC to the septum*. Mol Microbiol, 2003. **47**(2): p. 345-55.
57. Szeto, T.H., et al., *Membrane localization of MinD is mediated by a C-terminal motif that is conserved across eubacteria, archaea, and chloroplasts*. Proc Natl Acad Sci U S A, 2002. **99**(24): p. 15693-8.
58. Dajkovic, A., A. Mukherjee, and J. Lutkenhaus, *Investigation of regulation of FtsZ assembly by SulA and development of a model for FtsZ polymerization*. J Bacteriol, 2008. **190**(7): p. 2513-26.
59. Shiomi, D. and W. Margolin, *The C-terminal domain of MinC inhibits assembly of the Z ring in Escherichia coli*. J Bacteriol, 2007. **189**(1): p. 236-43.
60. Bernhardt, T.G. and P.A. de Boer, *SlmA, a nucleoid-associated, FtsZ binding protein required for blocking septal ring assembly over Chromosomes in E. coli*. Mol Cell, 2005. **18**(5): p. 555-64.
61. Wu, L.J. and J. Errington, *Coordination of cell division and chromosome segregation by a nucleoid occlusion protein in Bacillus subtilis*. Cell, 2004. **117**(7): p. 915-25.
62. Shen, B. and J. Lutkenhaus, *Examination of the interaction between FtsZ and MinCN in E. coli suggests how MinC disrupts Z rings*. Mol Microbiol. **75**(5): p. 1285-98.
63. Li, M.X., D.C. Corson, and B.D. Sykes, *Structure determination by NMR. Isotope labeling*. Methods Mol Biol, 2002. **173**: p. 255-65.
64. Russel., S.a., *Molecular Cloning: A Laboratory Manual*. . 2006.
65. Bodenhausen, G.a.R., D. , *Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy*. Chemical Physics Letters, 1980. **69**: p. 185-189.

## Anexo I – Declaração CIBio/IB

Uso exclusivo da CIBio:

Número de projeto / processo

ACM 7.2.2

### Formulário de encaminhamento de projetos de pesquisa para análise pela CIBio - Comissão Interna de Biossegurança da ABTLuS – Associação Brasileira de Tecnologia de Luz Síncrotron

Titulo do projeto: Expressão e caracterização de proteínas envolvidas na via da quinase mTOR e da divisão celular bacteriana

Pesquisador responsável: Ana Carolina de Mattos Zeri

Experimentador: Maria Luiza Caldas Nogueira

Nível do treinamento do experimentador:  -Iniciação científica,  -mestrado,  -doutorado,  -doutorado direto,  -pós-doutorado,  -nível técnico,  -outro, especifique: \_\_\_\_\_

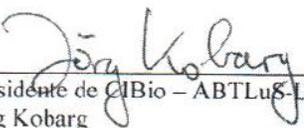
#### Resumo do projeto:

A mTOR é uma via de sinalização muito conservada que controla o crescimento celular em resposta à presença de nutrientes e fatores de crescimento. A desregulação dessa via em humanos está relacionada a doenças como câncer e diabetes. A quinase TOR é ativada na presença de aminoácidos e recentemente descobriu-se que as pequenas GTPases da família Rag são mediadoras da sinalização por Leucina. Essas GTPases são ancoradas na superfície do lisossomo por meio da interação com um complexo de três proteínas denominado *Ragulator*. Esse complexo também ancora um braço da via das MAPKs (MEK-ERK) aos lisossomos. O entendimento deste complexo pode nos ajudar elucidar doenças em que a via da mTOR se encontra desregulada. Neste trabalho obtivemos o complexo *Ragulator*, através da expressão da proteína p18 em corpos de inclusão e sua renaturação através da adição de suas parceiras Mp1/p14 a diálise. Foram realizados estudos biofísicos com a intenção de caracterização do complexo, entretanto o alto grau de dissociação do mesmo nos propiciou certa dificuldade em caracterizá-lo.

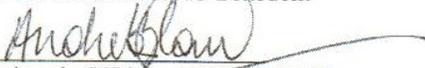
O estudo da divisão bacteriana é centralmente dependente de FtsZ, um homólogo procariótico das tubulinas. FtsZ desencadeia a divisão ao formar o "anel Z", uma estrutura supramolecular constituída por polímeros de FtsZ que circunda o interior da célula e funciona como arcabouço do aparato de divisão. A formação do anel Z é regulada por moduladores, proteínas que afetam tanto negativamente como positivamente a capacidade de FtsZ polimerizar-se. A proteína MinC é um inibidor da polimerização de FtsZ, recrutada por MinD para a face interna da membrana plasmática, onde o complexo MinCD exerce sua função. MinCD representa um inibidor sítio-específico da polimerização da FtsZ, prevenindo a formação do anel Z nos pólos das células mas permitindo que isto aconteça na região central. A elucidação deste processo seria de grande valia para o desenho racional de inibidores da divisão bacteriana. Neste trabalho, comprovamos a interação entre MinC e FtsZ por Ressonância Magnética Nuclear. Estes proteínas não se encontravam em sua forma monomérica e o alto peso molecular do complexo nos impossibilitou de identificar os aminoácidos necessários para esta interação, devido a limites da técnica. No momento, a proteína MinC está sendo expressa em presença de deutério, o que aumenta significativamente o limite da técnica de NHSQC. Foram realizados ainda estudos biofísicos com intuito de caracterização da interação.

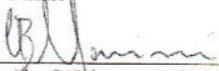
A CIBio analisou este projeto em reunião realizada no dia: 11.1.12.

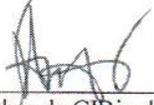
Parecer final:  -projeto aprovado,  -projeto recusado,  -projeto com deficiências, favor comentários anexo.

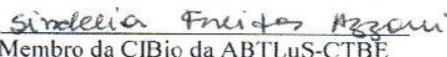
  
Presidente de CIBio – ABTLuS-LNBio  
Jörg Kobarg

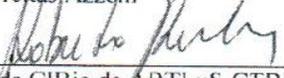
  
Membro da CIBio – ABTLuS-LNBio  
Prof. Dr. Celso Eduardo Benedetti

  
Membro da CIBio – ABTLuS-LNBio  
Andrea Balan

  
Membro da CIBio – ABTLuS-SGT  
Carolina Barbosa Marini

  
Membro da CIBio da ABTLuS-CTBE  
Fabio Squina

  
Membro da CIBio da ABTLuS-CTBE  
Sindelia Freitas Azzoni

  
Membro da CIBio da ABTLuS-CTBE  
Roberto Ruller

## DECLARAÇÃO

Declaro para os devidos fins que o conteúdo de minha dissertação de Mestrado/tese de Doutorado intitulada "Expressão e caracterização de proteínas envolvidas na via da quinase mTOR e na divisão celular bacteriana"

( ) não se enquadra no § 3º do Artigo 1º da Informação CCPG 01/08, referente a bioética e biossegurança.

Tem autorização da(s) seguinte(s) Comissão(ões):

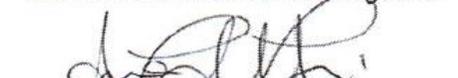
(  ) CIBio – Comissão Interna de Biossegurança , projeto No. ACM7 2.2, Instituição:

( ) CEUA – Comissão de Ética no Uso de Animais , projeto No. \_\_\_\_\_, Instituição:

( ) CEP - Comissão de Ética em Pesquisa, protocolo No. \_\_\_\_\_, Instituição:

*\* Caso a Comissão seja externa ao IB/UNICAMP, anexar o comprovante de autorização dada ao trabalho. Se a autorização não tiver sido dada diretamente ao trabalho de tese ou dissertação, deverá ser anexado também um comprovante do vínculo do trabalho do aluno com o que constar no documento de autorização apresentado.*

  
Aluno: Maria Luiza Caldas Nogueira

  
Orientador: Ana Carolina de Mattos Zeri

Para uso da Comissão ou Comitê pertinente:

(  ) Deferido ( ) Indeferido

Carimbo e assinatura

Prof. Dr. MARCELO LANCELLOTTI  
Presidente da Comissão Interna de Biossegurança  
Instituto de Biologia - UNICAMP

Para uso da Comissão ou Comitê pertinente:

( ) Deferido ( ) Indeferido

Carimbo e assinatura

## Anexo II – Artigo

Santos CR, Paiva JH, Sforça ML, Neves JL, Navarro RZ, Cota J, Akao PK, Hoffmam ZB, Meza AN, Smetana JH, Nogueira ML, Polikarpov I, Xavier-Neto J, Squina FM, Ward RJ, Ruller R, Zeri AC, Murakami MT (2012). “Dissecting structure-function-stability relationships of a thermostable GH5-CBM3 cellulase from *Bacillus subtilis* 168”. *Biochem J.* 441(1):95-104.

O artigo a seguir corresponde a uma colaboração realizada entre vários grupos de pesquisa, coordenados pelo prof. Dr. Mario T. Murakami, onde minha participação se enquadrou na purificação da CBM3 de *Bacillus subtilis*, marcada  $^{15}\text{N}$  e  $^{13}\text{C}$  para estudos de RMN.

## Dissecting structure–function–stability relationships of a thermostable GH5-CBM3 cellulase from *Bacillus subtilis* 168

Camila R. SANTOS<sup>\*1</sup>, Joice H. PAIVA<sup>\*1</sup>, Maurício L. SFORÇA<sup>\*</sup>, Jorge L. NEVES<sup>\*</sup>, Rodrigo Z. NAVARRO<sup>\*</sup>, Júnio COTA<sup>†</sup>, Patrícia K. AKAO<sup>\*</sup>, Zaira B. HOFFMAM<sup>†</sup>, Andréia N. MEZA<sup>\*</sup>, Juliana H. SMETANA<sup>\*</sup>, Maria L. NOGUEIRA<sup>\*</sup>, Igor POLIKARPOV<sup>‡</sup>, José XAVIER-NETO<sup>\*</sup>, Fábio M. SQUINA<sup>†</sup>, Richard J. WARD<sup>§</sup>, Roberto RULLER<sup>†</sup>, Ana C. ZERI<sup>\*</sup> and Mário T. MURAKAMI<sup>\*2</sup>

<sup>\*</sup>Laboratório Nacional de Biotecnologia, Centro Nacional de Pesquisa em Energia e Materiais, Campinas SP, Brazil, <sup>†</sup>Laboratório Nacional de Ciência e Tecnologia do Bioetanol, Centro Nacional de Pesquisa em Energia e Materiais, Campinas SP, Brazil, <sup>‡</sup>Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, Brazil, and <sup>§</sup>Departamento de Química, Universidade de São Paulo, Ribeirão Preto, Brazil

Cellulases participate in a number of biological events, such as plant cell wall remodelling, nematode parasitism and microbial carbon uptake. Their ability to depolymerize crystalline cellulose is of great biotechnological interest for environmentally compatible production of fuels from lignocellulosic biomass. However, industrial use of cellulases is somewhat limited by both their low catalytic efficiency and stability. In the present study, we conducted a detailed functional and structural characterization of the thermostable BsCel5A (*Bacillus subtilis* cellulase 5A), which consists of a GH5 (glycoside hydrolase 5) catalytic domain fused to a CBM3 (family 3 carbohydrate-binding module). NMR structural analysis revealed that the *Bacillus* CBM3 represents a new subfamily, which lacks the classical calcium-binding

motif, and variations in NMR frequencies in the presence of cellopentaose showed the importance of polar residues in the carbohydrate interaction. Together with the catalytic domain, the CBM3 forms a large planar surface for cellulose recognition, which conducts the substrate in a proper conformation to the active site and increases enzymatic efficiency. Notably, the manganese ion was demonstrated to have a hyper-stabilizing effect on BsCel5A, and by using deletion constructs and X-ray crystallography we determined that this effect maps to a negatively charged motif located at the opposite face of the catalytic site.

**Key words:** accessory domain, cellulase 5A, carbohydrate-binding module, kinetics, structure, thermal stability.

### INTRODUCTION

The production of ethanol from lignocellulosic biomass is well-placed among other possibilities to produce energy, owing to its potential sustainability and agro-economic benefits [1]. Different chemical and enzymatic strategies have been proposed for the saccharification of lignocellulosic biomass [2,3]. The latter has been implemented at different organization levels, from point mutations of cellulolytic enzymes to the engineering of whole metabolic pathways in micro-organisms [4]. Furthermore, these two approaches have been employed synergistically to increase the yields of fermentable sugars from lignocellulosic biomass. Although enzymatic hydrolysis has been successfully implemented as a biomass-to-bioenergy technology, it is widely regarded as an expensive and wasteful link in the whole chain of bioethanol production, mainly because of the large amounts of enzymes required to compensate for their low catalytic efficiency and stability [5,6]. Driven by this need, a number of enzymes with biological and commercial value have been systematically modified through structure-based rational approaches, or by intelligent use of serendipity [7–9]. In spite of these efforts, the molecular basis of protein stability and its intricate correlation with catalysis is still elusive, putting a premium on studies that address structure–function–stability

relationships and the discovery of new alternative pathways to improve thermal stability of the enzymes.

Endo- $\beta$ -1,4-glucanases (EC 3.2.1.4), also referred to as endoglucanases, are the major enzymes responsible for the breakdown of internal glycosidic bonds of cellulose chains. Several endoglucanases from different *Bacillus subtilis* strains (BsEgls) have been cloned and characterized, aiming at their potential applications in the biofuels industry [10–17]. BsEgls encompass a catalytic domain belonging to family 5A, containing a CBM3 (family 3 carbohydrate-binding module) appended to their C-terminus. BsEgls show maximum activity at approximately pH 6 and 60 °C [10,14] and are capable of hydrolysing CMC (carboxymethylcellulose) and lichenan, but not xylan, chitosan or laminarin [10,18]. Typically, BsEgls are thermostable enzymes, retaining 90 % of activity after incubation for 2 h at 65 °C, 70 % after 30 min at 75 °C, and 12 % after 10 min at 80 °C [11,13,14].

It has been proposed that the non-catalytic domain of BsEgls is important to bind insoluble substrates [13], increasing the catalytic efficiency by a mechanism that involves the disruption of the compact cellulose structure and delivering of the substrate to the CC (catalytic core) in a favoured configuration for catalysis [19–21]. As it is often the case with GHs (glycoside hydrolases), deletion of the accessory domain decreases the thermal stability

Abbreviations used: BsCel5A, *Bacillus subtilis* cellulase 5A; CBM, carbohydrate-binding module; BsEgl, *Bacillus subtilis* endoglucanase; CC, catalytic core; CMC, carboxymethylcellulose; CtCipACBM, *Clostridium thermocellum* CBM3; GH, glycoside hydrolase; HSQC, heteronuclear single-quantum coherence; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser enhancement spectroscopy; PEG, poly(ethylene glycol); SAXS, small-angle X-ray scattering;  $T_m$ , melting temperature; TmCel5A, *Thermotoga maritima* Cel5A; WT, wild-type.

<sup>1</sup> These authors contributed equally to this work.

<sup>2</sup> To whom correspondence should be addressed (email mario.murakami@inbio.org.br).

Structural factors and atomic co-ordinates of the *Bacillus subtilis* cellulase 5A catalytic core have been deposited in the Protein Data Bank under accession codes 3PZV (Form I), 3PZU (Form II) and 3PZT (Form II\*). NMR data of the *Bacillus subtilis* cellulase 5A CBM3 were deposited in the Biological Magnetic Resonance Bank and Protein Data Bank under codes 17399 and 2L8A respectively.

of the catalytic domain [22–30]; however, in BsEgls it has a contrary effect, by increasing enzyme stability [13]. Interestingly, the thermal stability conferred by a CBM to its parental protein is not transferred when it is recombinantly fused to a new CC, suggesting a complex molecular modulation and co-evolution of both modules [25,31–34]. In addition, experimental and *in silico* studies regarding the influence of the linker region [35] and domain–domain interactions [26,29] in the thermal stability of GHs are inadequate. Thus, in spite of extensive functional and structural studies of CBMs, the consequences of the fusion of a given CBM to a specific GH are as yet unpredictable [22,24,36,37].

In the present study we have described in detail the role exerted by the accessory domain in the operation and stability of BsEgls. Furthermore, by studying the specific nature of BsEgl thermal stability, we uncovered a novel stabilization mechanism based on metal ion co-ordination. This in-depth understanding of the modulation of the function and stability of BsEgls at a molecular level is a required step towards its optimization and use in biotechnological processes.

## MATERIALS AND METHODS

### Cloning, expression and purification

Three constructions of BsCel5A (*B. subtilis* cellulase 5A) were amplified by PCR using *B. subtilis* 168 genomic DNA and specific oligonucleotides [5'-GCTAGCGCAGCAGGGACAAAACG-3' and 5'-GGATCCCTTAGGTGCCGAGAATGTTTCTC-3' for CC; 5'-GCTAGCATTTCTGTACAGTACAGAGCA-3' and 5'-GGA-TCCCTAATTGGTTCTGTGCC-3' for CBM3; and 5'-GCT-AGCGCAGCAGGGACAAAACG-3' and 5'-GGATCCCTAA-TTTGGTTCTGTGCC-3' for the WT (wild-type) protein]. The amplified sequences were cloned into pGEM-T Easy vector (Promega) and further subcloned into the pET28a expression vector using the restriction sites present in the oligonucleotides (underlined sequences). The proteins were expressed in BL21(DE3)*styD*<sup>-</sup> cells at 37 °C for 4 h after induction with 0.5 mM IPTG (isopropyl  $\beta$ -D-thiogalactopyranoside) in Luria–Bertani broth. For NMR experiments, the BL21(DE3)*styD*<sup>-</sup> cells transformed with pET28a-CBM3 plasmid were grown in M9 minimal medium supplemented with 2 g/l [<sup>13</sup>C]glucose and 1 g/l [<sup>15</sup>N]ammonium chloride [38] using the same conditions described above. The harvested cells were resuspended in lysis buffer (50 mM sodium phosphate, pH 7.2, 100 mM NaCl, 1 mM PMSF and 5 mM benzamidine) and disrupted by lysozyme treatment (80  $\mu$ g/ml, 30 min, on ice), followed by sonication (10 pulses of 10 s on ice using a Vibracell VCX 500, Sonics & Materials). The solutions were centrifuged at 10000 *g* for 30 min and the supernatants were loaded on to nickel-charged 5 ml HiTrap Chelating columns (GE Healthcare) using a flow rate of 1 ml/min. The proteins were eluted using a non-linear gradient of 0–500 mM imidazole. For WT construction, an anionic-exchange chromatography using a 5 ml HiTrap SP HP column (GE Healthcare), a flow rate of 1 ml/min and a non-linear gradient of 0–1 M NaCl was used to improve sample purity. All constructions were concentrated further using Ultrafree-CL (Millipore) and submitted to size-exclusion chromatography using a Superdex 75 16/60 column (GE Healthcare) and a flow rate of 0.5 ml/min. All chromatographic steps were performed using an ÄKTA FPLC system (GE Healthcare). Sample homogeneity was confirmed by SDS/PAGE under denaturing conditions [39]. Protein concentration was estimated by absorbance at 280 nm using molar extinction coefficients derived from the sequence of the constructs.

### Enzymatic characterization and optimization of endo- $\beta$ -1,4-glucanase activity

The standard enzymatic assays for evaluation of CC and WT constructs were performed in triplicate using different substrates at 0.25% and 100 ng of enzyme in 50 mM citrate-phosphate buffer, pH 6.0. The mixture was incubated at 60 °C for 60 min to determine activity in a set of 12 polysaccharides including  $\beta$ -glucan, CMC, larch arabinogalactan, laminarin, galactomannan, xyloglucan, debranched arabinan, 1,4- $\beta$ -D-mannan, rye arabinoxylan, xylan from beechwood, pectin and lichenan (purchased from Megazyme and Sigma–Aldrich). The enzymatic activity was determined by the amount of reducing sugar liberated from different polysaccharides using the DNS (3,5-dinitrosalicylic acid) method [40]. One unit of endoglucanase activity was defined as the quantity of enzyme to release 1  $\mu$ mol of reducing sugars per min under standard conditions. A response surface methodology was performed to optimize the reaction conditions for both CC and WT constructs of BsCel5A in the absence or presence of manganese ion (10 mM). The variables analysed in the present study were the pH together with the temperature. A central composite design ( $k = 2$ ) with four central points and a total of 12 experiments was considered for optimization of these variables. Details concerning the statistical approaches for these experiments have been described by Myers and Montgomery [41]. The regression and graphical analysis of the data were performed using the software Statistica™ version 8.0, considering a *P* value of 0.05 as significant. The significance of the regression coefficients was given by Student's *t* test. The second-order model equation was determined by Fisher's test and the multiple coefficient of determination ( $R^2$ ) gave the variance explained by the model. All glucanase activity assays were carried out in triplicate using 10  $\mu$ l of enzyme (100 ng), 40  $\mu$ l of 0.1 M citrate-phosphate buffer and 50  $\mu$ l of  $\beta$ -glucan (stock 0.5% in water). The enzymatic reaction was performed at different pH values in the citrate-phosphate buffer system and various temperatures as indicated in Supplementary Table S1 (at <http://www.BiochemJ.org/bj/441/bj4410095add.htm>). Metal ions, chelating agent (EDTA), surfactants (SDS and Triton X-100) and organic solvents (methanol, ethanol and DMSO) were added to the enzyme assay to test their effects on activity. Control experiments using a citrate-free buffer were carried out, and the results were statistically identical with those obtained in the citrate buffer, indicating no chelating effect of the citrate over the manganese ion under the tested conditions. The kinetic study was performed with  $\beta$ -glucan as substrate at 60 °C and pH 6.0 according to the method described above. The kinetic parameters  $k_{cat}$  and  $K_m$  were determined by a non-linear regression fit of the saturation curves using the Michaelis–Menten equation.

### CD spectroscopy and thermal unfolding studies

Far-UV CD spectra (190–260 nm) were recorded on a Jasco J-810 spectropolarimeter (Jasco International Co.) coupled to a Peltier temperature controller using a 1 mm quartz cuvette. Proteins were used at a concentration of 10  $\mu$ M. Data collection parameters were set to scan rate of 50 nm/min, response time of 4 s, sensitivity of 100 mdeg, accumulation of 10, heating rate of 1 °C/min and delay time for spectrum collection of 60 s. Results were expressed as mean residue ellipticity ( $\text{deg} \cdot \text{cm}^2 \cdot \text{dmol}^{-1} \cdot \text{residue}^{-1}$ ). All thermal unfolding experiments were monitored at 222 nm, except for the CBM3 construct, which was monitored at 230 nm.

### SAXS (small-angle X-ray scattering)

SAXS data for CC, CBM3 and WT constructs were collected using a 165 mm MarCCD detector at the D02A/SAXS2 beamline (Brazilian Synchrotron Light Laboratory, Campinas, Brazil). The radiation wavelength was set to 1.48 Å (1 Å = 0.1 nm) and the sample-to-detector distance was set to 1559.1 mm to give a scattering vector range from 0.1 nm<sup>-1</sup> to 2.3 nm<sup>-1</sup> ( $q = 4\pi \sin\theta/\lambda$ , where  $2\theta$  is the scattering angle). The samples at 2, 4 or 8 mg/ml, in 150 mM NaCl and 20 mM Tris/HCl, pH 7.5, were centrifuged for 10 min at 20000 g (4°C) and then filtered to remove any aggregates. Measurements were carried out at room temperature (~23°C). Buffer scattering was recorded and subtracted from the corresponding sample scattering. The integration of SAXS patterns was performed using Fit2D software [42] and the curves were scaled by protein concentration. Data analysis was performed using the GNOM program [43]. Low-resolution envelopes of each construction were calculated from the experimental SAXS data using an *ab initio* procedure implemented in the DAMMIN program [44]. Averaged models were generated from 10 runs using the DAMAVER package [45]. SAXS envelopes and NMR or XTAL structures were superimposed using the SUPCOMB 2.0 program [46]. BUNCH software [47], which employs rigid body model and simulated annealing routines, was used to find the best relative positions of NMR and XTAL structures in relation to the SAXS envelope. The CRY SOL program [48] was used to calculate the scattering curve from high-resolution structures, including the WT model obtained from BUNCH.

### NMR spectroscopy

NMR experiments were performed using a Varian Inova spectrometer at the National Biosciences Laboratory (LNBio/CNPEM), operating at a <sup>1</sup>H Larmor frequency of 599.887 MHz and temperature of 20°C. For these experiments, the spectrometer was equipped with a triple resonance cryogenic probe and a Z pulse-field gradient unit. The <sup>15</sup>N-<sup>13</sup>C-labelled CBM sample was dissolved in 20 mM phosphate buffer, pH 7.2, containing 50 mM NaCl and 5% (v/v) <sup>2</sup>H<sub>2</sub>O, at a final concentration of approximately 0.3 mM. Water suppression was achieved by low-power continuous wave irradiation over the relaxation delay or using the WATERGATE method [49]. All data were processed using NMRPipe and NMRVIEW software [50,51]. Prior to Fourier transformation, the time domain data were zero-filled in all dimensions. When necessary, a fifth-order polynomial baseline correction was applied after transformation and phasing. To obtain distance constraints, cross-peak volumes were estimated from NOESY (nuclear Overhauser enhancement spectroscopy). Protein backbone resonance peaks were assigned using <sup>1</sup>H-<sup>15</sup>N-HSQC (heteronuclear single-quantum coherence) and standard triple resonance experiments, including HNCACB, CBCACONH, HNCO and HNCACO [52] of the <sup>15</sup>N-<sup>13</sup>C-labelled CBM3 sample. All spectra for resonance assignment were recorded by setting up a resolution of 2048 × 128 × 96 points and spectral window of 8.00 kHz, 12.07 kHz and 1.95 kHz for <sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N respectively. In order to assign the side chain resonances, the <sup>15</sup>N-<sup>13</sup>C-labelled CBM sample was freeze-dried and dissolved in 100% <sup>2</sup>H<sub>2</sub>O, followed by acquisition of HCCH-TOCSY, hCCH-TOCSY spectra. NOE (nuclear Overhauser effect)-derived distance restraints were obtained from the <sup>15</sup>N-HSQC-NOESY and <sup>13</sup>C-HSQC-NOESY (separately optimized for aliphatics and aromatics), both collected with 80 ms mixing time. The CBM3 structure was calculated in a semi-automated iterative manner with the CYANA program (version 2.1) [53], using 100

starting conformers. The CYANA 2.1 protocol was applied to calibrate and assign NOE cross-peaks. After the first few rounds of automatic calculations, the NOESY spectra were analysed again to identify additional cross-peaks consistent with the structural model and to correct misidentified NOEs. The structures obtained were further refined by restrained minimization and molecular dynamic studies using the CNS software [54]. The 20 structures with the lowest target function were selected to represent the ensemble of protein structures. The quality of the structures was analysed with PROCHECK-NMR [55]. The summary of NMR data is presented in Supplementary Table S2 (at <http://www.BiochemJ.org/bj/441/bj4410095add.htm>).

### Proton/deuterium exchange

Proton/deuterium exchange measurements were carried out with the <sup>15</sup>N-<sup>13</sup>C-labelled CBM3 sample dissolved in 100% <sup>2</sup>H<sub>2</sub>O. In order to monitor the intensity loss of labile protons, a series of <sup>1</sup>H-<sup>15</sup>N-HSQC spectra were collected first at 15 min (time required to dissolve the protein and perform the first experiment) and subsequently at 40 min time intervals over two days. Additionally, a last spectrum was collected 10 days after the protein dissolution.

### X-ray crystallography

The highly purified CC sample was concentrated to 75 mg/ml and its buffer was exchanged to water by ultrafiltration using Vivaspin6 centrifugal filters (GE Healthcare). Crystallization experiments were performed by the vapour diffusion method using a HoneyBee 963 robot (Genomic Solutions). Sitting drops were prepared by mixing 0.5 μl of the protein solution with an equal volume of mother liquor and equilibrated against 80 μl of the reservoir solution at 18°C. Formulations based on the following commercial crystallization kits were tested: SaltRX, Crystal Screen, Crystal Screen 2 (Hampton Research), Precipitant Synergy, Wizard I and II (Emerald BioSystems), PACT and JCSG+ (Qiagen/Nextal), with a total of 544 solutions. Large crystals grew within 45 days in the following conditions: 100 mM Bis-Tris-propane pH 6.5, 20% (w/v) PEG [poly(ethylene glycol)] 3350 and 200 mM sodium nitrate (form I); 100 mM SPG (succinic acid, sodium dihydrogen phosphate monohydrate, glycine) pH 6.0 and 25% (w/v) PEG 1500 (form II). Crystals were transferred to new drops containing the mother liquor and 20% (v/v) glycerol for cryoprotection and then flash-cooled in a 100 K nitrogen gas stream. In order to obtain an ion-protein complex, crystals belonging to form II were soaked for 3 min in a cryoprotectant solution containing 50 mM of manganese chloride. X-ray diffraction data were collected on the W01B-MX2 beamline at the Brazilian Synchrotron Light Laboratory (Campinas, Brazil). The wavelength used was 1.4586 Å and the intensities were recorded in a Mar Mosaic 225 mm CCD (charged-coupled-device) detector. Data were indexed, integrated and scaled using the HKL2000 package [56]. The structure of the crystal form II\* was solved by the molecular replacement method using the MOLREP program [57] and the atomic co-ordinates of Cel5A from *Bacillus agaradhaerens* (PDB code 7A3H), which shares 70% identity with the catalytic domain of BsCel5A as the search model. Initial cycles of refinement involved a restrained and overall B-factor refinement in the REFMAC5 program [58] using automatic weighting between X-ray and geometry terms. After each cycle of refinement, the model was inspected and manually adjusted into the ( $2F_o - F_c$ ) and ( $F_o - F_c$ ) electron density maps using the COOT program [59]. Solvent molecules were manually added at positive peaks above 3.0  $\sigma$  in the difference Fourier

**Table 1** Kinetic parameters of BsCel5A and its catalytic core using  $\beta$ -glucan as substrate

Parameter	CC		WT	
	0 mM Mn <sup>2+</sup>	10 mM Mn <sup>2+</sup>	0 mM Mn <sup>2+</sup>	10 mM Mn <sup>2+</sup>
$V_{max}$	143.7 ± 6.6	92.3 ± 5.9	46.2 ± 3.1	35.7 ± 1.8
$K_m$	1.4 ± 0.2	1.6 ± 0.3	0.32 ± 0.2	0.5 ± 0.2
$k_{cat}$	78.7	50.5	40.4	31.2
$k_{cat}/K_m$	56.2	31.6	127.8	58.8

maps, taking into consideration hydrogen bonding potential. This structure was extensively refined and then used as the search model for solving other crystalline forms.

#### Accession numbers

Structure factors and atomic co-ordinates of the BsCel5A CC were deposited in the PDB under accession codes 3PZV (Form I), 3PZU (Form II) and 3PZT (Form II\*). NMR data of the BsCel5A CBM3 were deposited in the Biological Magnetic Resonance Bank and PDB under codes 17399 and 2L8A respectively.

## RESULTS AND DISCUSSION

### Accessory domain enhances catalytic efficiency

The WT protein and its truncated domains were expressed in *Escherichia coli* cells and purified by nickel-affinity and size-exclusion chromatographies (see Supplementary Figure S1 at <http://www.BiochemJ.org/bj/441/bj4410095add.htm>). Biochemical characterization showed that the CBM3 deletion does not alter the substrate specificity (see Supplementary Figure S2 at <http://www.BiochemJ.org/bj/441/bj4410095add.htm>) and optimal conditions for catalysis (see Supplementary Figure S3 at <http://www.BiochemJ.org/bj/441/bj4410095add.htm>). Despite the fact that the physiological substrate of BsCel5A is cellulose, we have used soluble substrates in kinetic studies owing to the very slow rate of hydrolysis of insoluble substrates, such as Avicel® (results not shown). In addition, it has been shown that BsCel5A is only able to bind Avicel® when CBM3 is appended to its C-terminus [13].

The kinetics study showed that CBM3 increased BsCel5A affinity to the substrate owing to a reduced  $K_m$ , but the enzyme turnover ( $V_{max}$ ) is reduced (Table 1). However, the reduction in  $K_m$  largely compensates for the lower  $V_{max}$  and, as a result, CBM3 enhances catalytic efficiency ( $k_{cat}/K_m$ ) (Table 1). We determined that BsCel5A displays an extended substrate binding surface owing to the strategically three-dimensional arrangement of CBM3 in relation to CC (see details below), which favours additional protein-carbohydrate interactions. These additional interactions are reflected in a higher substrate affinity and may also explain the lower turnover of the WT enzyme ( $V_{max}$ ).

### Deletion of the accessory domain reduces enzyme stability

The thermal stability of the BsCel5A constructions was assessed by unfolding studies using CD spectroscopy (Figure 1). As expected, CD spectra of CC and CBM3 displayed characteristic profiles of  $\alpha/\beta$  and all- $\beta$  proteins respectively (Figure 1A). It is worth noting that BsCel5A and its truncated forms showed a single structural transition in the thermal unfolding experiments (Figure 1B), despite the fact that BsCel5A is a two-domain protein and therefore two transitions corresponding to each domain were

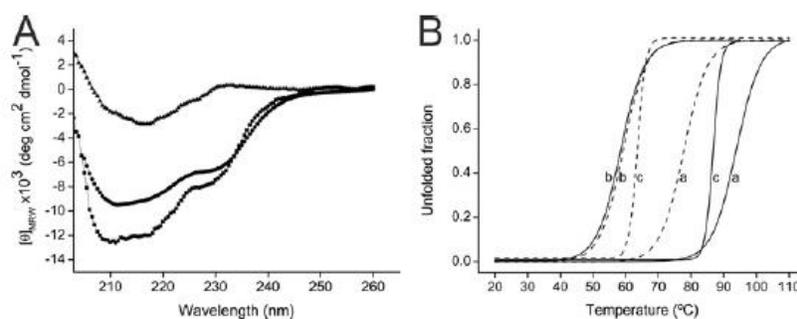
expected. Perhaps the 23-residue-long linker and inter-domain contacts drive the thermal denaturation to a single-transition process. Interestingly, unfolding studies revealed that deletion of the accessory domain increases the thermal stability from 64 to 78 °C ( $T_m$ , melting temperature). The reduced thermal stability of the WT protein is probably influenced by the significantly lower stability of CBM3 ( $T_m = 59$  °C) in relation to the parental catalytic domain ( $T_m = 78$  °C). As shown by NMR structural studies, the lower stability of BsCel5ACBM is explained by the absence of a structural calcium ion. Moreover, deletion of CBM3 promotes exposure of polar residues at the CC surface, which becomes favourably hydrated, thus increasing stability [60].

Both biochemical and biophysical studies clearly indicate the importance of CBM3 in the regulation of overall BsCel5A function, in spite of its apparent negative effect on stability.

### *Bacillus* CBM3 is a member of a novel calcium-independent subfamily

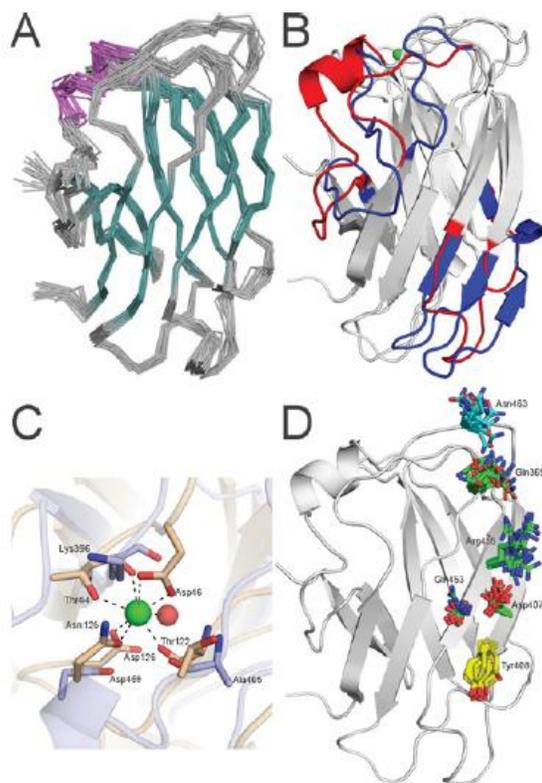
To gain insights into the structural determinants of the stability of *Bacillus* CBM3, we performed a NMR study. The <sup>1</sup>H-<sup>15</sup>N-HSQC spectrum of CBM3 showed high-quality signal dispersion (see Supplementary Figure S4 at <http://www.BiochemJ.org/bj/441/bj4410095add.htm>), indicating a well-folded protein. The most common classification, based on sequence similarities, defines BsCel5A CBM3 as a member of the CBM3 family (<http://www.cazy.org/Carbohydrate-Binding-Modules.html>). So far, five members of CBM3 have been solved, sharing from 13 to 36% of identity with the BsCel5A CBM3. BsCel5A CBM3 folds into a  $\beta$ -sandwich, a common scaffold for CBMs, encompassing a core formed by two  $\beta$ -sheets of four and five  $\beta$ -strands, and one short  $\alpha$ -helix at the C-terminal region (Figure 2A). Structural comparisons showed that the  $\beta$ -sandwich core is conserved among CBMs, but the surface loops are different in both composition and conformation (Figure 2B). Canonically, the CBM3 members with solved structures have a highly conserved calcium-binding site, which is not present in the BsCel5A CBM3 structure (Figures 2B and 2C). A search for sequences sharing similarity with BsCel5A CBM3 in GenBank® resulted in a set of CBMs that do not have the structural requirements for calcium binding (see Supplementary Figure S5 at <http://www.BiochemJ.org/bj/441/bj4410095add.htm>). Interestingly, all of these CBMs are from *Bacillus* spp., suggesting that the ancestor of this bacterium genus has lost the CBM3 calcium-binding site during evolution.

Therefore, we propose a new classification of CBM3 into two groups: the calcium-binding proteins, which include the already defined subfamilies CBM3a, CBM3b, CBM3b' and CBM3c; and the calcium-independent proteins, termed in the present study as the CBM3d subfamily, with BsCel5A CBM3 being its first structurally characterized member. The classification of CBM3s in a, b and c refers to the type of substrate that is preferentially recognized by the CBM [61]. The a and b groups bind crystalline cellulose; a is specific for cellulosomal-scaffolding proteins and b is found in soluble cellulases. Previously, it was shown that certain CBM3b members bind neither crystalline nor soluble cellulose and hence, they were collectively designated as CBM3b' [62]. The c group binds single-chain cellulose and is found in cellulosomal cellulases. The new CBM3d subfamily is closely related to the CBM3a subfamily, owing to its substrate preference; however, it lacks the calcium-binding site, which seems to play a structural role. Indeed, our biophysical studies showed that the *Bacillus* CBM3 is less stable than its parental protein, which can be attributed to the lack of structural calcium.



**Figure 1** Biophysical characterization of BsCel5A constructions

(A) Far-UV CD spectra of CC (●), CBM (▲) and WT (■). (B) Thermal denaturation in the absence (dashed lines) or presence of manganese (solid lines) of CC (a), CBM (b) and WT (c).



**Figure 2** NMR solution structure of BsCel5A CBM3

(A) View of the ensemble of 20 lowest energy structures, with  $\beta$ -strands in cyan,  $\alpha$ -helices in magenta and loops in grey. (B) Structural superposition of BsCel5A CBM3 structure on CiCipACBM (PDB code 1NBC) with structural differences coloured in blue and red respectively. The  $\text{Ca}^{2+}$  ion from the CiCipACBM structure is shown as a green sphere. (C) Calcium-binding site of CiCipACBM (wheat) and the corresponding region in the BsCel5A CBM3 structure (light blue). (D) BsCel5A CBM3 residues involved in carbohydrate recognition: residues conserved in CiCipACBM and identified by NMR CSP (chemical shift perturbation) analysis (carbon atoms in green); residues conserved but not identified by NMR CSP (carbon atoms in yellow); and residues not conserved but identified by NMR CSP (carbon atoms in cyan).

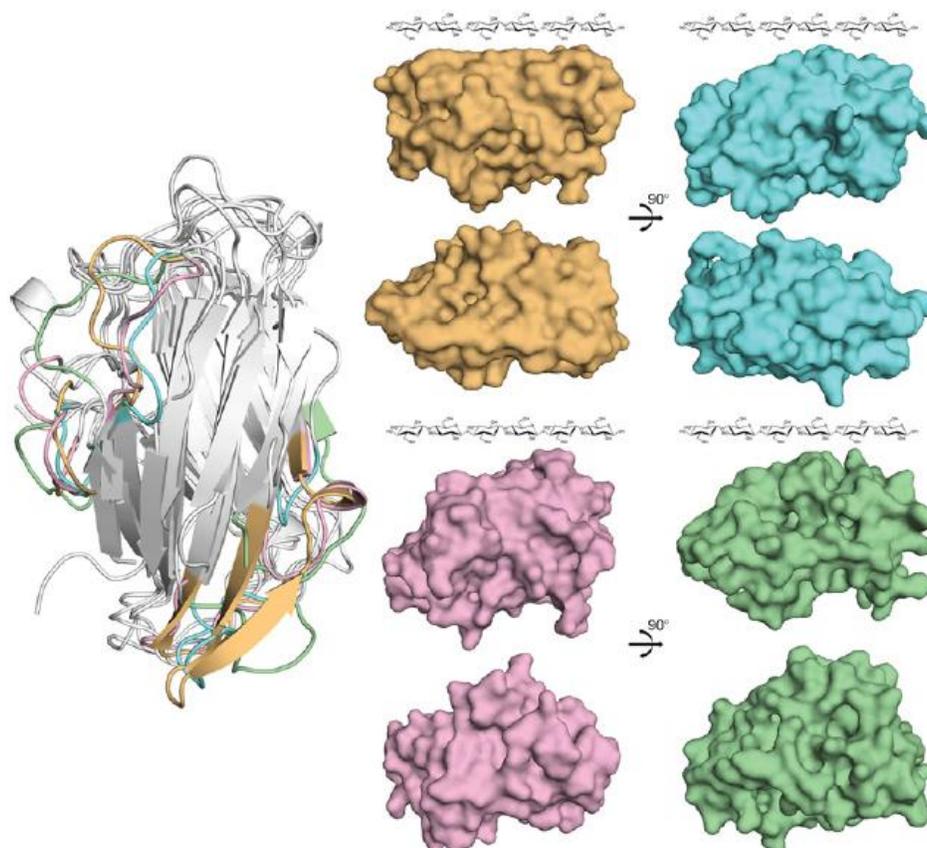
It is worth mentioning that there is another classification of CBMs taking into account both structural and functional similarities, which assorts these modules in types A, B and C

[19]. The type A binds crystalline cellulose and chitin through a flat surface, type B binds a single-chain polysaccharide via an extended groove or cleft, and type C binds mono-, di- or tri-saccharides. In this classification, the BsCel5A CBM3 is considered a type A member, owing to its ability to bind crystalline cellulose [13] and higher structural similarity with type A (PDB code 1NBC) than other type B (PDB code 1J84) or type C (PDB code 1I82) CBMs.

#### Polar residues play a pivotal role in oligosaccharide recognition by CBM3d

The molecular surface of CBMs has been associated with the type of substrate that they can bind, e.g. CBM3a presents a very flat surface compatible with its target, crystalline cellulose. Despite the fact that BsCel5A binds crystalline cellulose and the binding occurs only when the CBM3 is present [13], the *Bacillus* CBM3 surface is not as flat as those of CBM3a members (Figure 3). This unexpected surface topology of BsCel5A CBM3 can be attributed to the absence of the CBM3a  $\beta 4'$  strand (Figure 3) and to the inherent flexibility of its surface loops, as assessed by hydrogen–deuterium exchange (see Supplementary Figure S6 at <http://www.BiochemJ.org/bj/441/bj4410095add.htm>) and NMR relaxation experiments (results not shown). On the basis of these dynamics studies, we suggest that the BsCel5A CBM3 can undergo conformational changes, adapting its topology for cellulose recognition and binding. Moreover, the exact protein–carbohydrate interactions are not known, since sugar–complex structures of a CBM3a or even a type A CBM are not available. Thus, in order to map the residues involved in the recognition of cellulose, we analysed variations in NMR frequencies of BsCel5A CBM3 in the presence of cellopentaose. The major perturbations in chemical shifts occurred in Asn<sup>463</sup> and Arg<sup>455</sup>, and minor changes were observed in Gln<sup>369</sup>, Gln<sup>453</sup> and Asp<sup>407</sup> (Figure 2D and Supplementary Figure S7 at <http://www.BiochemJ.org/bj/441/bj4410095add.htm>). The residues Arg<sup>455</sup>, Gln<sup>369</sup>, Gln<sup>453</sup> and Asp<sup>407</sup> are conserved in CiCipACBM (*Clostridium thermocellum* CBM3, PDB code 1NBC), a well-characterized type A CBM [63]. Tyr<sup>408</sup> is also conserved, but it is not affected by the presence of cellopentaose, perhaps owing to the short length of this oligosaccharide.

The molecular recognition and affinity towards carbohydrates is typically achieved by van der Waals interactions arising from stacking of aromatic residues with the furanose/pyranose rings and by hydrogen bonds between polar groups of saccharides and amino acid side chains. The mapped interactions of the BsCel5A CBM3 with cellopentaose revealed that the carbohydrate-binding



**Figure 3** Structural comparison of BsCel5A CBM3 (CBM3d) with CBM3a, b' and c

The CBM3s from CtpA (CBM3a, PDB code 1NBC), E4 (CBM3c, PDB code 4TF4) and Cel9V (CBM3b', PDB code 2WOB) were superposed on the CBM3d structure and are shown as cartoon representations (left-hand panel). The main differences of the carbohydrate-binding face were coloured in green (CBM3d), orange (CBM3a), cyan (CBM3b') and pink (CBM3c). The molecular surfaces of CBM3a, b', c and d are shown as side (middle panels) and frontal (right-hand panels) views following the colours above.

face is rich in charged and polar residues, such as glutamine, arginine, aspartate, lysine and asparagine. The only aromatic residue on this face is Tyr<sup>408</sup>, whereas a set of aromatic residues is buried in the  $\beta$ -sandwich core making hydrophobic contacts. These findings indicate that the interactions of BsCel5A CBM3 with cellulose are mainly by polar side chains. Moreover, we suggest that these polar residues may not only work on the carbohydrate affinity/recognition, but could specifically replace inter-chain hydrogen bonds, acting to disrupt the tight cellulose structure [63]. Therefore it may result in the local separation of cellulose single chains to be delivered to the CC, thus contributing to a higher catalytic efficiency.

#### Three-dimensional domain architecture of BsCel5A reveals a flat cellulose-binding platform

BsCel5A and its separated domains were analysed by SAXS, showing that all forms are monomeric in solution with  $R_g$  values of 1.75, 1.63 and 3.04 nm, and  $D_{max}$  values of 4.65, 4.8, and 9.5 nm, for the CC, CBM3 and WT protein respectively (Table 2 and Supplementary Figure S8 at <http://www.BiochemJ.org/bj/441/bj4410095add.htm>). Both CC and CBM3 were nicely fitted into the corresponding envelopes (Figure 4A). The good

quality of SAXS data allowed us to observe the active-site cavity in the CC. The averaged low-resolution model of the WT protein calculated from 20 independent *ab initio* models gives a good notion of the three-dimensional arrangement of the accessory domain with respect to its parental catalytic domain (Figure 4B). The molecular envelope reveals a large continuous planar surface, which has been considered essential for crystalline cellulose binding. The atomic co-ordinates of both domains were fitted into the SAXS model of the WT protein, which showed that this planar surface is formed by the active-site interface of the CC in line with the cellulose recognition face of CBM3. This model is in full agreement with the substrate-recognition face mapped by NMR and with the crystallographic structures of the multidomain endoglucanase from *Clostridium cellulolyticum* (Cel9G) and cellulase E4 from *Thermobifida fusca* (Figure 4C).

#### Manganese metal ion hyper-stabilizes the CC of BsCel5A

Divalent ions are well-known cofactors of enzymes, either activating their catalytic machinery or having a structural role. In the present study we tested the effect of different divalent ions, including Mn<sup>2+</sup>, Mg<sup>2+</sup>, Ca<sup>2+</sup>, Zn<sup>2+</sup>, Ni<sup>2+</sup> and Co<sup>2+</sup>. None of them affected the stability, except the Mn<sup>2+</sup> ion, which drastically

**Table 2** Summary of SAXS calculations

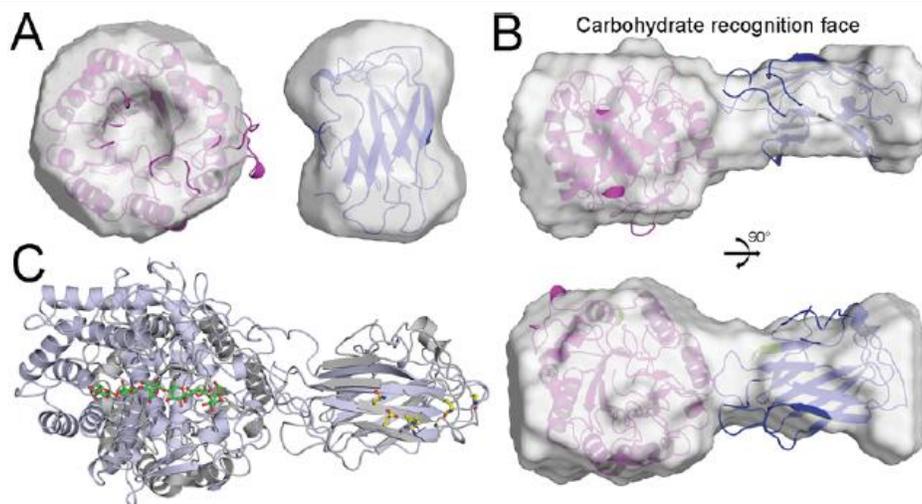
	CC			CBM			WT		
	Exper*	Envel†	XTAL‡	Exper*	Envel†	NMR‡	Exper*	Envel†	Assem‡
$R_g$ (nm)	1.75	1.73	1.86	1.63	1.63	1.65	3.04	2.96	2.87
$D_{max}$ (nm)	4.65			4.80			9.50		
$\chi^2$ §		2.13	2.89		1.89	2.75		2.37	2.01

\*Calculated from experimental data.

†Calculated from the SAXS-derived envelope.

‡Calculated from XTAL, NMR or assembly.

§Comparison between experimental and theoretical curves.

**Figure 4** SAXS analysis of BsCel5A constructions

(A) SAXS envelopes from CC (left-hand panel) and CBM3d (right-hand panel) are shown as transparent surfaces and the filled XTAL or NMR structures are shown as cartoon representations. (B) Different views of the CC and CBM3d arrangement into the SAXS envelope from the WT protein. (C) Overlay of the E4 structure from *T. fusca* (light blue) on the BsCel5A model (grey). The oligosaccharide bound in the E4 catalytic domain is shown in green and the BsCel5A CBM3 residues identified by NMR CSP are shown in yellow.

improved the thermal stability of BsCel5A, by elevation of the  $T_m$  by 23 °C, from 64 to 87 °C (Figure 1B). This result indicates that Cel5As (family 5A cellulases) display a very specific motif for recognition of manganese ions. Thus, in order to map the domains involved in the manganese ion interactions, truncated forms of the protein were analysed. As shown in Figure 1(B), the CC had its thermal stability improved by the  $Mn^{2+}$  ion, whereas that of CBM remained unaltered, mapping the specific interactions of the  $Mn^{2+}$  ion with the CC.

#### Structural basis for BsCel5A stabilization by manganese ion

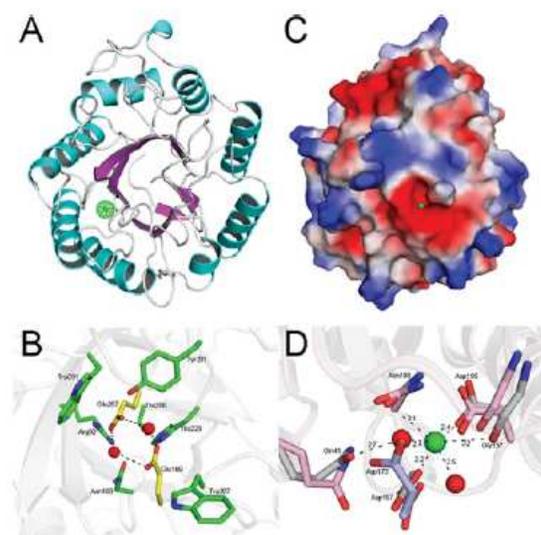
In order to identify the manganese-binding site and to delineate the structural basis for the enhanced thermal stability of BsCel5A by this ion, we undertook an X-ray crystallographic study. BsCel5A CC crystallized in two different crystalline forms (form I and form II), which diffracted up to 2.90 Å and 2.10 Å respectively. However, only the crystal form II was successfully soaked with manganese ions (Table 3). Crystals belonging to form I display a very particular crystallographic symmetry in which the manganese-binding site is concealed by an aspartate residue of the neighbouring molecule (Figure 5D), which prevented ion incorporation by soaking methods.

The BsCel5A CC structure encompasses a classical  $(\beta/\alpha)_8$ -barrel fold (Figure 5A), with the active site located at the crevice formed by the C-terminal ends of  $\beta$ -strands. The catalytic machinery involves two acid residues, Glu<sup>229</sup> and Glu<sup>169</sup>, the aromatic gate keepers Trp<sup>207</sup> and Trp<sup>291</sup>, and the polarizers His<sup>229</sup>, Arg<sup>92</sup> and Thr<sup>256</sup> (Figure 5B). From homology with other GH5 cellulases, Glu<sup>169</sup> is expected to be the proton donor, which protonates the glycosidic bond and catalyses the bond fission, and the Glu<sup>257</sup> is the nucleophile, which facilitates the reaction by stabilizing the resulting carbonium ion intermediate. Several studies demonstrated the key role of a histidine residue in the protonation of the acid/base glutamate residue, which is controlled by a catalytic triad, similar to the serine proteases [64]. In this scheme, a third element, which could be a serine/threonine residue, an aspartate residue or a water molecule, would protonate the histidine, permitting a classification of the glycosyl hydrolases according to this histidine proton donor [64]. BsCel5A contains a threonine residue in the equivalent position, which allows classification as type A catalytic triad. Cel5As from *Thermotoga maritima* and *C. thermocellum* have a water molecule in the corresponding position and are classified as type C enzymes (see Supplementary Figure S9 at <http://www.BiochemJ.org/bj/441/bj4410095add.htm>). Water-dependent catalytic machineries are labile at alkaline conditions, since the water molecule

**Table 3** Data collection and refinement statistics

Values in parentheses are for the highest-resolution shell. RMSD, root mean square deviation.

Parameter	Form I	Form II native	Form II* 10 mM Mn <sup>2+</sup>
<i>Data collection</i>			
Space group	C2	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Unit cell parameters (Å, °)	a = 155.36, b = 81.08, c = 114.27, β = 104.29	a = 50.24, b = 110.76, c = 121.38	a = 49.53, b = 110.62, c = 122.26
Resolution	50.00–2.90 (3.00–2.90)	50.00–2.10 (2.18–2.10)	50.00–1.97 (2.04–1.97)
Number of unique reflections	30 407 (2607)	39 069 (3393)	47 366 (4163)
Multiplicity	3.5 (2.7)	6.9 (6.2)	4.8 (4.3)
Completeness (%)	97.1 (83.7)	97.2 (85.9)	97.6 (87.8)
R <sub>merge</sub> (%)	8.3 (43.6)	11.6 (31.9)	8.4 (37.3)
I/σ (I)	11.3 (2.0)	13.5 (5.0)	14.7 (3.8)
<i>Refinement</i>			
R <sub>work</sub> /R <sub>free</sub> (%)	23.0/29.6	17.1/22.1	16.4/20.6
Average B-factor	54.5	32.7	29.1
RMSD for bond lengths (Å)	0.006	0.023	0.025
RMSD for bond angles (°)	1.261	1.903	1.938
Protein chains	4	2	2
Water molecules	30	330	452
<i>Ramachandran plot</i>			
Favoured region (%)	90.63	96.44	97.09
Additionally allowed region (%)	8.02	3.39	2.91
Outlier region (%)	1.35	0.17	0.00
PDB code	3PZV	3PZU	3PZT

**Figure 5** Crystal structure of the BsCel5A CC

(A) Cartoon representation of the protein structure, with the manganese ion shown as a green sphere. (B) Zoom of the active site showing the glutamate residues (carbon atoms in yellow), other catalytically relevant residues (carbon atoms in green) and two water molecules (red spheres). (C) Electrostatic surface coloured by charge, from red (–) to blue (+), indicating the presence of the manganese ion in an acidic groove. (D) Interactions of protein residues (grey) and water molecules (red spheres) with the manganese ion (green sphere). The 2.9 Å resolution structure (pink) has an aspartate residue of a neighbouring molecule (blue) occluding the manganese-binding site.

deprotonates at high pH, disrupting the hydrogen network required for catalysis. On the other hand, enzymes with the type A catalytic triad would remain active even in basic conditions, which is advantageous for biotechnological purposes.

In the manganese complex structure, the ion is located at the opposite face of the active site, near to the lid-like β-

hairpin (Figure 5C), and is hexa-co-ordinated by the two negatively charged residues Asp<sup>195</sup> and Asp<sup>197</sup>, the polar Asn<sup>198</sup>, the carbonyl oxygen from Gly<sup>157</sup> and two water molecules (Figure 5D). The strategic location of the manganese-binding site permits simultaneous water-mediated interactions between the N-terminus and distant regions of the protein, increasing the overall structural rigidity and consequently a notable gain in the protein stability. Comparison between forms II (without manganese ion) and II\* (with manganese ion) showed that the ion only substituted a water molecule in the highly negatively charged aperture without affecting either the cleft architecture or the conformational states of catalytically relevant residues. This result also suggests that the region comprising the lid-like β-hairpin, which is subsequent to the N-terminal end, can be a strategic motif that could be engineered with the aim of improving protein stability. Indeed, in the analogous hyperthermostable TmCel5A (*T. maritima* Cel5A) [65], the β-hairpin is replaced by an α-helical element, which enables additional polar contacts, supporting the fact that the stabilization of this region may implicate a higher overall stability.

#### The effect of the manganese ion on kinetics and specificity of BsCel5A

Our biochemical data showed that the presence of Mn<sup>2+</sup> ions prolonged the lifetime of the enzyme at high temperatures (90 °C) (see Supplementary Figure S10 at <http://www.BiochemJ.org/bj/441/bj4410095add.htm>) without interfering with the optimal conditions (pH and temperature) for activity (see Supplementary Figure S3). Although it was expected that the gain in thermal stability by Mn<sup>2+</sup> ions could produce higher relative activities at high temperatures, both univariate (Supplementary Figure S11 at <http://www.BiochemJ.org/bj/441/bj4410095add.htm>) and response surface (Supplementary Figure S3) experiments indicated no alteration in the temperature-dependent activity profile. This native-like behaviour of BsCel5A could be related to the unaltered configuration of the active site, since the ion binds to a remote site, ~30 Å away from the active-site cleft. In terms of protein dynamics, we suggest that, although the ion stabilizes the structural framework of the enzyme, the catalytic machinery remains susceptible to temperature-induced changes.

As expected, the presence of  $Mn^{2+}$  ions did not modify substrate preference, but it slightly reduced the total activity by approximately 30% (Supplementary Figure S2). This partial inhibition was initially attributed to a possible secondary binding site involving the catalytic glutamate residues, since the manganese-binding site is sterically and electrostatically similar to the active-site pocket, and in the crystal structure of TmCel5A, a divalent (cadmium) ion was found interacting with the catalytic glutamate residues [66]. However, kinetic experiments showed that the  $Mn^{2+}$  ion did not interfere with the substrate binding ( $K_m$ ) (Table 2), thus discarding the possibility of an inhibitory effect by a competitive mechanism. In fact, only  $V_{max}$  was decreased by the presence of  $Mn^{2+}$  ions, which reflected as a lower enzymatic efficiency (Table 2). It could be a consequence of the higher overall molecular rigidity, which prevents the conformational changes required for optimal catalysis. This hypothesis is supported by previous studies that described the pivotal role of ligand-induced conformational alterations in the catalytic activity of family 5A cellulases [66].

As a complement to the thorough characterization of the binding and effects of the  $Mn^{2+}$  ions, other ligands were tested, including several ions ( $Fe^{3+}$ ,  $Cu^{2+}$ ,  $Mn^{2+}$ ,  $Zn^{2+}$ ,  $Ni^{2+}$ ,  $Co^{2+}$ ,  $Mg^{2+}$  and  $Ca^{2+}$ ), a chelating agent (EDTA), surfactants (SDS and Triton X-100) and organic solvents (methanol, ethanol and DMSO) (see Supplementary Figure S12 at <http://www.BiochemJ.org/bj/441/bj4410095add.htm>). Briefly, none of the additives increased the enzymatic activity;  $Fe^{3+}$  and  $Cu^{2+}$  almost abolished the activity; methanol and ethanol did not reduce the enzymatic activity at concentrations up to 10%; and SDS caused a decrease of 30% in enzymatic activity.

#### Biotechnological implications

Modern biotechnology lies at the interface between basic and applied science, and continuously demands novel enzymes with improved catalytic efficiency and stability under extreme and non-natural conditions. Cellulases have many industrial applications, including food processing, finishing of textiles and generation of cellulosic bioethanol, one of the sustainable energy sources. These industrial processes require cellulolytic activity under a broad range of pH, temperature and ionic strengths. In this context, the understanding of enzymes at a molecular level and the mechanisms to improve their stability and to modify their catalytic efficiency and specificity is essential for rational protein design aiming to ameliorate their use in biotechnological processes.

In the present study, we investigated the role of the accessory domain in the stability and function of family 5 cellulases, and a novel mechanism for thermal stabilization was proposed. Our results showed that *Bacillus* CBM3 represents a novel subfamily, which lacks the calcium-binding site and its binding to carbohydrate is mediated by polar contacts. Moreover, *Bacillus* CBM3 decreases the overall enzyme stability, which can be attributed to the lack of the structural calcium ion, but on the other hand it improves the catalytic efficiency. The discovery of a specific manganese-binding site that is conserved among 5A cellulases suggests a novel approach for thermal stability enhancement without enzyme engineering.

#### AUTHOR CONTRIBUTION

Camila Santos and Mário Murakami designed the research. Camila Santos, Igor Polikarpov, José Xavier-Neto, Richard Ward, Ana Zeri and Mário Murakami conceived and designed the experiments. Crystal structure determination was carried out by Camila Santos and Joice Paiva. Spectroscopic analyses were performed by Joice Paiva, Patricia Akao and Andréia Meza. Functional and kinetic studies were performed by Júnio Cota, Zaira Hoffmann, Fábio Squina and Roberto Ruller. NMR studies were carried out by Maurício Sforça, Jorge

Neves, Rodrigo Navarro, Juliana Smetana, Maria Nogueira and Ana Zeri. Camila Santos and Mário Murakami analysed the data and wrote the paper. Mário Murakami supervised the project. All authors read and approved the final version of the paper.

#### ACKNOWLEDGEMENTS

We are thankful for the provision of time on the MX2 beamline (LNLS), Robolab (LNBio), LEC (LNBio) and NMR (LNBio) at the National Center for Research in Energy and Materials (Campinas, Brazil).

#### FUNDING

This research was supported by grants from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) [grant number 10/51890-8] Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) [grant number 476043/2011-5] and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

#### REFERENCES

- Farrell, A. E., Plevin, R. J., Turner, B. T., Jones, A. D., O'Hare, M. and Kammen, D. M. (2006) Ethanol can contribute to energy and environmental goals. *Science* **311**, 506–508
- Lynd, L. R., Laser, M. S., Bransby, D., Dale, B. E., Davison, B., Hamilton, R., Himmel, M., Keller, M., McMillan, J. D., Sheehan, J. and Wyman, C. E. (2008) How biotech can transform biofuels. *Nat. Biotechnol.* **26**, 169–172
- Binder, J. B. and Raines, R. T. (2010) Fermentable sugars by chemical hydrolysis of biomass. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 4516–4521
- Shaw, A. J., Podkaminer, K. K., Desai, S. G., Bardsley, J. S., Rogers, S. R., Thome, P. G., Hogsett, D. A. and Lynd, L. R. (2008) Metabolic engineering of a thermophilic bacterium to produce ethanol at high yield. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 13769–13774
- Margeot, A., Hahn-Hagerdal, B., Edlund, M., Slade, R. and Monot, F. (2009) New improvements for lignocellulosic ethanol. *Curr. Opin. Biotechnol.* **20**, 372–380
- Wen, F., Nair, N. U. and Zhao, H. (2009) Protein engineering in designing tailored enzymes and microorganisms for biofuels production. *Curr. Opin. Biotechnol.* **20**, 412–419
- Heinzlman, P., Snow, C. D., Wu, I., Nguyen, C., Villalobos, A., Govindarajan, S., Minshull, J. and Arnold, F. H. (2009) A family of thermostable fungal cellulases created by structure-guided recombination. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 5610–5615
- Morais, S., Heyman, A., Barak, Y., Caspi, J., Willson, D. B., Lamed, R., Shoseyov, O. and Bayer, E. A. (2010) Enhanced cellulose degradation by nano-complexed enzymes: synergism between a scaffold-linked exoglucanase and a free endoglucanase. *J. Biotechnol.* **147**, 205–211
- van der Veen, B. A., Skov, L. K., Potocki-Véronèse, G., Gajhede, M., Monsan, P. and Remaud-Simeon, M. (2006) Increased amylosucrase activity and specificity, and identification of regions important for activity, specificity and stability through molecular evolution. *FEBS J.* **273**, 673–681
- Liu, M., Wang, J., Liu, J., Yao, J. and Yu, Z. (2006) Expression of *Bacillus subtilis* JA18 endo- $\beta$ -1,4-glucanase gene in *Escherichia coli* and characterization of the recombinant enzyme. *Ann. Microbiol.* **56**, 41–45
- Li, W., Zhang, W. W., Yang, M. M. and Chen, Y. L. (2008) Cloning of the thermostable cellulase gene from newly isolated *Bacillus subtilis* and its expression in *Escherichia coli*. *Mol. Biotechnol.* **40**, 195–201
- Li, W., Huan, X., Zhou, Y., Ma, Q. and Chen, Y. (2009) Simultaneous cloning and expression of two cellulase genes from *Bacillus subtilis* newly isolated from Golden Takin (*Budorcas taxicolor bedfordi*). *Biochem. Biophys. Res. Commun.* **383**, 397–400
- Wang, Y., Yuan, H., Wang, J. and Yu, Z. (2009) Truncation of the cellulose binding domain improved thermal stability of endo- $\beta$ -1,4-glucanase from *Bacillus subtilis* JA18. *Bioresour. Technol.* **100**, 345–349
- Yang, D., Weng, H., Wang, M., Xu, W., Li, Y. and Yang, H. (2010) Cloning and expression of a novel thermostable cellulase from newly isolated *Bacillus subtilis* strain 115. *Mol. Biol. Rep.* **37**, 1923–1929
- Deka, D., Bhargavi, P., Sharma, A., Goyal, D., Jawed, M. and Goyal, A. (2011) Enhancement of cellulase activity from a new strain of *Bacillus subtilis* by medium optimization and analysis with various cellulosic substrates. *Enzyme Res.* **2011**, 151656
- Liao, H., Zhang, X. Z., Rollin, J. A. and Zhang, Y. H. (2011) A minimal set of bacterial cellulases for consolidated bioprocessing of lignocellulose. *Biotechnol. J.*, doi:10.1002/biot.201100157
- Zhang, X. Z. and Zhang, Y. H. (2011) Simple, fast and high-efficiency transformation system for directed evolution of cellulase in *Bacillus subtilis*. *Microb. Biotechnol.* **4**, 98–105
- Woli, M., Geczi, A., Simon, O. and Borriss, R. (1995) Genes encoding xylan and  $\beta$ -glucan hydrolysing enzymes in *Bacillus subtilis*: characterization, mapping and construction of strains deficient in lichenase, cellulase and xylanase. *Microbiology* **141**, 281–290

- 19 Boraston, A. B., Bolam, D. N., Gilbert, H. J. and Davies, G. J. (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem. J.* **382**, 769–781
- 20 Shoseyov, O., Shani, Z. and Levy, I. (2006) Carbohydrate binding modules: biochemical properties and novel applications. *Microbiol. Mol. Biol. Rev.* **70**, 283–295
- 21 Hashimoto, H. (2006) Recent structural studies of carbohydrate-binding modules. *Cell. Mol. Life Sci.* **63**, 2954–2967
- 22 Fontes, C. M., Hazlewood, G. P., Morag, E., Hall, J., Hirst, B. H. and Gilbert, H. J. (1995) Evidence for a general role for non-catalytic thermostabilizing domains in xylanases from thermophilic bacteria. *Biochem. J.* **307**, 151–158
- 23 Winterhalter, C., Heinrich, P., Candussio, A., Wich, G. and Liebl, W. (1995) Identification of a novel cellulose-binding domain within the multidomain 120 kDa xylanase XynA of the hyperthermophilic bacterium *Thermotoga maritima*. *Mol. Microbiol.* **15**, 431–444
- 24 Irwin, D., Shin, D. H., Zhang, S., Barr, B. K., Sakon, J., Karplus, P. A. and Wilson, D. B. (1998) Roles of the catalytic domain and two cellulose binding domains of *Thermomonospora fusca* E4 in cellulose hydrolysis. *J. Bacteriol.* **180**, 1709–1714
- 25 Kalaeva, I. A., Blum, D. L., Li, X. L. and Ljungdahl, L. G. (2001) Do domain interactions of glycosyl hydrolases from *Clostridium thermocellum* contribute to protein thermostability? *Protein Eng.* **14**, 167–172
- 26 Xie, H., Gilbert, H. J., Charnock, S. J., Davies, G. J., Williamson, M. P., Simpson, P. J., Raghobama, S., Fontes, C. M., Dias, F. M., Ferreira, L. M. and Bolam, D. N. (2001) *Clostridium thermocellum* Xyn10B carbohydrate-binding module 22-2: the role of conserved amino acids in ligand binding. *Biochemistry* **40**, 9167–9176
- 27 Shin, E. S., Yang, M. J., Jung, K. H., Kwon, E. J., Jung, J. S., Park, S. K., Kim, J., Yun, H. D. and Kim, H. (2002) Influence of the transposition of the thermostabilizing domain of *Clostridium thermocellum* xylanase (XynX) on xylan binding and thermostabilization. *Appl. Environ. Microbiol.* **68**, 3496–3501
- 28 Zhao, G., Ali, E., Araki, R., Sakka, M., Kimura, T. and Sakka, K. (2005) Function of the family-9 and family-22 carbohydrate-binding modules in a modular  $\beta$ -1,3-1,4-glucanase/xylanase derived from *Clostridium stercovarium* Xyn10B. *Biosci. Biotechnol. Biochem.* **69**, 1562–1567
- 29 Zhang, S., Yin, Q. Y., Li, Y. H., Ding, M., Xu, G. J. and Zhao, F. K. (2007) Molecular and biochemical characterization of Ba-EGA, a cellulase secreted by *Bacillus* sp. AC-1 from *Ampullaria crosseana*. *Appl. Microbiol. Biotechnol.* **75**, 1327–1334
- 30 Sunna, A. (2010) Modular organisation and functional analysis of dissected modular  $\beta$ -mannanase CsMan26 from *Caldicellulosiruptor* R18B.4. *Appl. Microbiol. Biotechnol.* **86**, 189–200
- 31 Kittur, F. S., Mangala, S. L., Rus'd, A. A., Kitaoka, M., Tsujibo, H. and Hayashi, K. (2003) Fusion of family 2b carbohydrate-binding module increases the catalytic activity of a xylanase from *Thermotoga maritima* to soluble xylan. *FEBS Lett.* **549**, 147–151
- 32 Mahadevan, S. A., Wi, S. G., Lee, D. S. and Bae, H. J. (2006) Site-directed mutagenesis and CBM engineering of Cel5A (*Thermotoga maritima*). *FEMS Microbiol. Lett.* **287**, 205–211
- 33 Mamo, G., Hatti-Kaul, R. and Mattiasson, B. (2007) Fusion of carbohydrate binding modules from *Thermotoga neapolitana* with a family 10 xylanase from *Bacillus halodurans* S7. *Extremophiles* **11**, 169–177
- 34 Zhang, Y., Chen, S., Xu, M., Cavaco-Paulo, A., Wu, J. and Chen, J. (2010) Characterization of *Thermobifida fusca* cutinase-carbohydrate-binding module fusion proteins and their potential application in bioscouring. *Appl. Environ. Microbiol.* **76**, 6870–6876
- 35 Sonan, G. K., Receveur-Brechot, V., Duez, C., Aghajari, N., Czjzek, M., Haser, R. and Gerday, C. (2007) The linker region plays a key role in the adaptation to cold of the cellulase from an Antarctic bacterium. *Biochem. J.* **407**, 293–302
- 36 Morris, D. D., Gibbs, M. D., Chin, C. W., Koh, M. H., Wong, K. K., Allison, R. W., Nelson, P. J. and Bergquist, P. L. (1998) Cloning of the xynB gene from *Dicetylglomus thermophilum* RI46B.1 and action of the gene product on kraft pulp. *Appl. Environ. Microbiol.* **64**, 1759–1765
- 37 Blanco, A., Diaz, P., Zueco, J., Parascandola, P. and Javier Pastor, F. I. (1999) A multidomain xylanase from *aBacillus* sp. with a region homologous to thermostabilizing domains of thermophilic enzymes. *Microbiology* **145**, 2163–2170
- 38 Li, M. X., Corson, D. C. and Sykes, B. D. (2002) Structure determination by NMR isotope labeling. *Methods Mol. Biol.* **173**, 255–265
- 39 Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685
- 40 Miller, G. L. (1959) Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Chem.* **31**, 426–42
- 41 Myers, R. H. and Montgomery, D. C. (2002) *Response Surface Methodology: Process and Product Optimization Using Designed Experiments* (2nd edn), John Wiley & Sons, New York
- 42 Hammersley, A. P., Brown, K., Burmeister, W., Clautre, L., Gonzalez, A., McSweeney, S., Mitchell, E., Moy, J. P., Svensson, S. O. and Thompson, A. W. (1997) Calibration and application of an X-ray image intensifier/charge-coupled device detector for monochromatic macromolecular crystallography. *J. Synchrotron Radiat.* **4**, 67–77
- 43 Svergun, D. I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25**, 495–503
- 44 Svergun, D. I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.* **76**, 2879–2886
- 45 Volkov, V. V. and Svergun, D. I. (2003) Uniqueness of *ab initio* shape determination in small-angle scattering. *J. Appl. Crystallogr.* **36**, 860–864
- 46 Kozin, M. B. and Svergun, D. I. (2000) Automated matching of high- and low-resolution structural models. *J. Appl. Crystallogr.* **34**, 33–41
- 47 Petoukhov, M. V. and Svergun, D. I. (2005) Global rigid body modelling of macromolecular complexes against small-angle scattering data. *Biophys. J.* **89**, 1237–1250
- 48 Svergun, D. I., Barberato, C. and Koch, M. H. J. (1995) CRYSOLOG—a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **28**, 768–773
- 49 Plotto, M., Saudek, V. and Sklenar, V. (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR* **2**, 661–665
- 50 Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293
- 51 Johnson, B. A. and Blevins, R. A. (1994) NMRView: a computer program for the visualization and analysis of NMR data. *J. Biomol. NMR* **4**, 603–614
- 52 Sattler, M., Schleucher, J. and Griesinger, C. (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectrosc.* **34**, 93–158
- 53 Guntert, P., Mumenthaler, C. and Wolrich, K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* **273**, 283–298
- 54 Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Panu, N. S. et al. (1998) Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905–921
- 55 Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R. and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477–486
- 56 Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326
- 57 Vagin, A. and Teplyakov, A. (2010) Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 22–25
- 58 Murshudov, G. N., Vagin, A. A. and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255
- 59 Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132
- 60 Vogt, G., Woell, S. and Argos, P. (1997) Protein thermal stability, hydrogen bonds, and ion pairs. *J. Mol. Biol.* **269**, 631–643
- 61 Bayer, E. A., Chanzy, H., Lamed, R. and Shoham, Y. (1998) Cellulose, cellulases and cellulosomes. *Curr. Opin. Struct. Biol.* **8**, 548–557
- 62 Jindou, S., Xu, Q., Kenig, R., Shulman, M., Shoham, Y., Bayer, E. A. and Lamed, R. (2006) Novel architecture of family-9 glycoside hydrolases identified in cellular enzymes of *Acetivibrio cellulosilyticus* and *Clostridium thermocellum*. *FEMS Microbiol. Lett.* **254**, 308–316
- 63 Tormo, J., Lamed, R., Chirino, A. J., Morag, E., Bayer, E. A., Shoham, Y. and Steitz, T. A. (1996) Crystal structure of a bacterial family-III cellulose-binding domain: a general mechanism for attachment to cellulose. *EMBO J.* **15**, 5739–5751
- 64 Shaw, A., Bott, R., Vonrhein, C., Brice, G., Power, S. and Day, A. G. (2002) A novel combination of two classic catalytic schemes. *J. Mol. Biol.* **320**, 303–309
- 65 Pereira, J. H., Chen, Z., McAndrew, R. P., Sapra, R., Chhabra, S. R., Sale, K. L., Simmons, B. A. and Adams, P. D. (2010) Biochemical characterization and crystal structure of endoglucanase Cel5A from the hyperthermophilic *Thermotoga maritima*. *J. Struct. Biol.* **172**, 372–379
- 66 Varrot, A., Frandsen, T. P., Driguez, H. and Davies, G. J. (2002) Structure of the *Humicola insolens* cellobiohydrolase Cel6A D416A mutant in complex with a non-hydrolysable substrate analogue, methyl cellobiosyl-4-thio- $\beta$ -cellobioside, at 1.9 Å. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 2201–2204

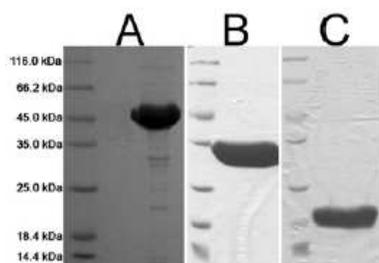
Received 16 May 2011/22 August 2011; accepted 1 September 2011  
Published as BJ Immediate Publication 1 September 2011, doi:10.1042/BJ20110869

SUPPLEMENTARY ONLINE DATA

## Dissecting structure–function–stability relationships of a thermostable GH5-CBM3 cellulase from *Bacillus subtilis* 168

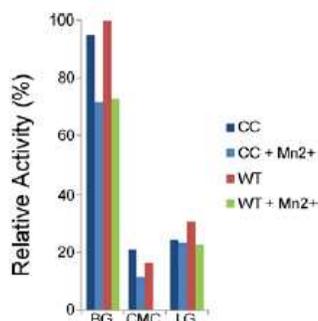
Camila R. SANTOS<sup>\*1</sup>, Joice H. PAIVA<sup>\*1</sup>, Maurício L. SFORÇA<sup>\*</sup>, Jorge L. NEVES<sup>\*</sup>, Rodrigo Z. NAVARRO<sup>\*</sup>, Júnio COTA<sup>†</sup>, Patrícia K. AKAO<sup>\*</sup>, Zaira B. HOFFMAM<sup>‡</sup>, Andréia N. MEZA<sup>\*</sup>, Juliana H. SMETANA<sup>\*</sup>, Maria L. NOGUEIRA<sup>\*</sup>, Igor POLIKARPOV<sup>‡</sup>, José XAVIER-NETO<sup>\*</sup>, Fábio M. SQUINA<sup>†</sup>, Richard J. WARD<sup>§</sup>, Roberto RULLER<sup>†</sup>, Ana C. ZERI<sup>\*</sup> and Mário T. MURAKAMI<sup>\*2</sup>

<sup>\*</sup>Laboratório Nacional de Biociências, Centro Nacional de Pesquisa em Energia e Materiais, Campinas SP, Brazil, <sup>†</sup>Laboratório Nacional de Ciência e Tecnologia do Bioetanol, Centro Nacional de Pesquisa em Energia e Materiais, Campinas SP, Brazil, <sup>‡</sup>Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, Brazil, and <sup>§</sup>Departamento de Química, Universidade de São Paulo, Ribeirão Preto, Brazil



**Figure S1** SDS/PAGE analysis of the purified BsCel5A constructions

(A) WT, (B) CC and (C) CBM3.



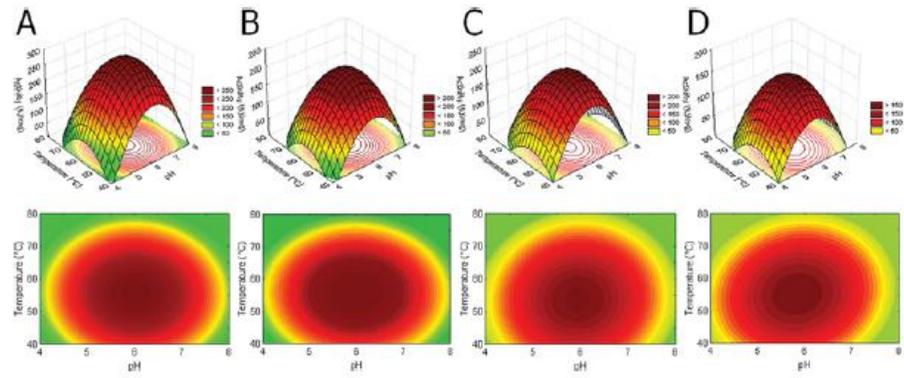
**Figure S2** Relative activity of BsCel5A (WT) and the CC on different substrates

Substrates used were  $\beta$ -glucan (BG), carboxymethylcellulose (CMC) and lichenan (LG).

<sup>1</sup> These authors contributed equally to this work.

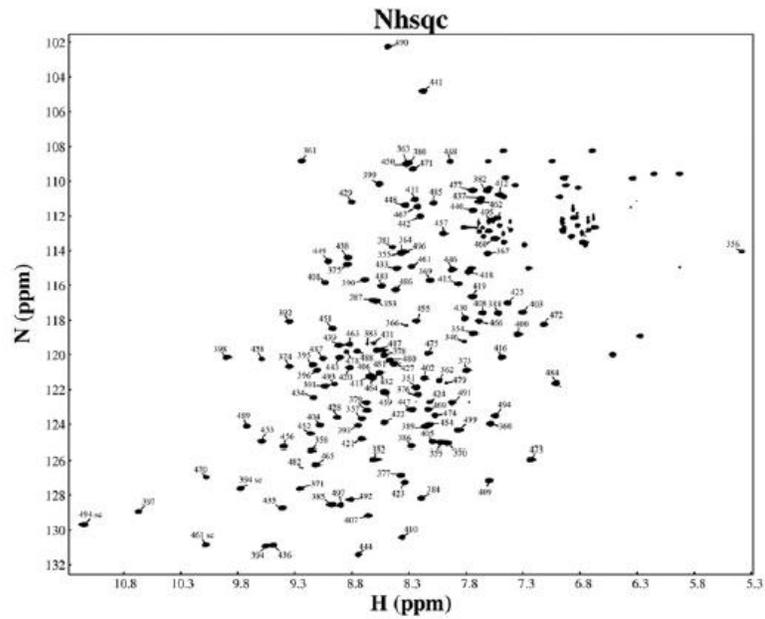
<sup>2</sup> To whom correspondence should be addressed (email mario.murakami@lnbio.org.br).

Structural factors and atomic co-ordinates of the *Bacillus subtilis* cellulase 5A catalytic core have been deposited in the Protein Data Bank under accession codes 3PZV (Form I), 3PZU (Form II) and 3PZT (Form II'). NMR data of the *Bacillus subtilis* cellulase 5A CBM3 were deposited in the Biological Magnetic Resonance Bank and Protein Data Bank under codes 17399 and 2L6A respectively.



**Figure S3 Response surface analysis**

Effect of pH and temperature on the glucanase activity of BsCel5A (**A**, without  $Mn^{2+}$  and **B**, with 10 mM  $Mn^{2+}$ ) and on the catalytic core (**C**, without  $Mn^{2+}$  and **D**, with 10 mM  $Mn^{2+}$ ).

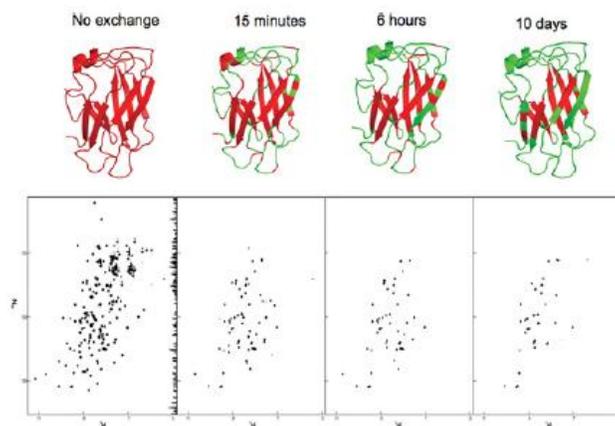


**Figure S4 Assigned  $^{15}N$ - $^1H$ -HSQC spectrum of the CBM3 domain**

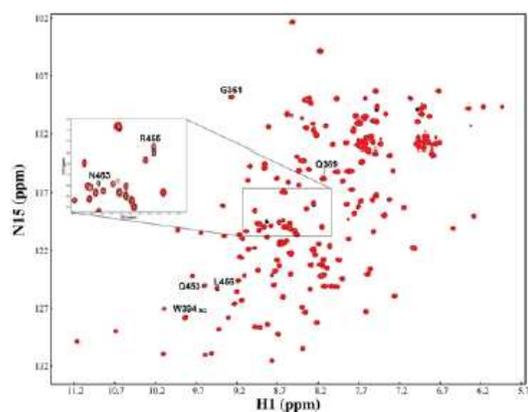
	392	406	453	473
<i>Bsu</i>	RWYV <b>Y</b> AN <b>Y</b> NRKGQN-PDC...QLRLHNDW <b>Y</b> SN-Y <b>Y</b> Q <b>S</b> SDY <b>S</b> FF			
<i>Bst</i>	RWYV <b>Y</b> AN <b>Y</b> NRKGQN-PDC...QLRLHNDW <b>Y</b> SN-Y <b>Y</b> Q <b>S</b> SDY <b>S</b> FF			
<i>Bli</i>	RWYV <b>Y</b> AN <b>Y</b> NRKGQN-PDC...QLRLHNDW <b>Y</b> SN-Y <b>Y</b> Q <b>S</b> SDY <b>S</b> FL			
<i>Bpu</i>	RWYV <b>Y</b> AN <b>Y</b> NRKGQN-VDC...QLRLHNDW <b>Y</b> SN-Y <b>Y</b> Q <b>S</b> SDY <b>S</b> FF			
<i>Bam</i>	RWYV <b>Y</b> AN <b>Y</b> NRKGQN-PDC...QLRLHNDW <b>Y</b> SN-Y <b>Y</b> Q <b>S</b> SDY <b>S</b> FF			
<i>4TF4</i>	RWYV <b>Y</b> LD <b>E</b> -----GV...Q <b>F</b> RIAG <b>G</b> PG-- <b>W</b> PS <b>S</b> HW <b>S</b> FP <b>Q</b>			
<i>1G87</i>	K <b>Y</b> FM <b>L</b> S <b>E</b> I <b>V</b> AA--GI...Q <b>F</b> RI <b>A</b> AP <b>Q</b> GT <b>T</b> Y <b>W</b> PK <b>N</b> DF <b>S</b> Y <b>D</b>			
<i>2WOB</i>	RWYV <b>Y</b> TS <b>G</b> NEQ <b>N</b> NPVC...P <b>F</b> RI <b>E</b> G-- <b>A</b> A <b>E</b> -Y <b>D</b> OT <b>T</b> D <b>Y</b> SY <b>N</b>			
<i>1G43</i>	R <b>Y</b> Y <b>Y</b> TS <b>Q</b> AD <b>K</b> EL <b>T</b> FW <b>C</b> ...Q <b>T</b> RF <b>A</b> R <b>N</b> D <b>W</b> SN- <b>F</b> D <b>Q</b> SN <b>D</b> W <b>S</b> Y <b>T</b>			
<i>1NBC</i>	R <b>Y</b> Y <b>Y</b> TS <b>V</b> GG <b>K</b> D <b>Q</b> T <b>F</b> W <b>C</b> ...Q <b>G</b> RF <b>A</b> R <b>N</b> D <b>W</b> SN- <b>Y</b> T <b>Q</b> SN <b>D</b> Y <b>S</b> Y <b>F</b> K			
	40	55	110	130

**Figure S5** Multiple sequence alignment of the region corresponding to the classical calcium-binding site in CBMs belonging to family 3

CBM3a (PDB codes 1NBC and 1G43), CBM3b' (PDB code 2WOB), CBM3c (PDB codes 1G87 and 4TF4) and CBM3d sequences: *Bsu*, *B. subtilis* (GB: NP\_389695.2); *Bst*, *Bacillus stearothermophilus* (GB: ADO21451.1); *Bli*, *Bacillus licheniformis* (GB: AAP51020.1); *Bpu*, *Bacillus pumilus* (GB: ACY72384.1); and *Bam*, *Bacillus amyloliquefaciens* (GB: ABS70711.1). The residues that interact with the calcium ion by side-chain or main-chain atoms are highlighted in green or cyan respectively. The residue conserved in all CBM3 subclasses is in yellow and those conserved only in CBM3d members are in red. The numbers at the top correspond to the *BsCel5A* numbering and the numbers at the bottom correspond to the 1NBC numbering.

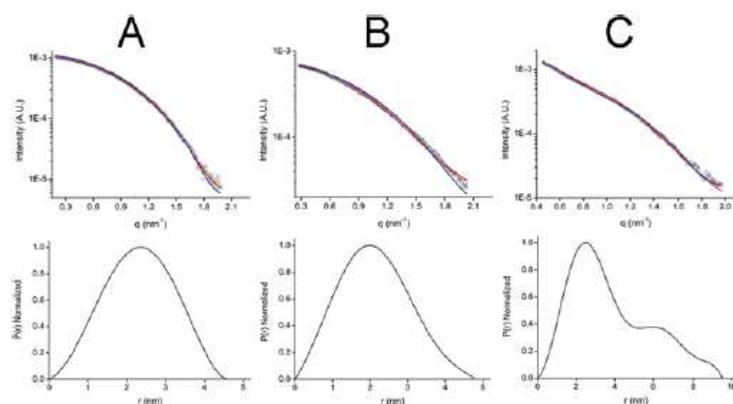


**Figure S6**  $^{15}\text{N}$ - $^1\text{H}$ -HSQC spectra (bottom panels) of hydrogen–deuterium exchange experiments at different time intervals and schematic representation of the structure (top panels) indicating more (green) or less (red) solvent-exposed residues



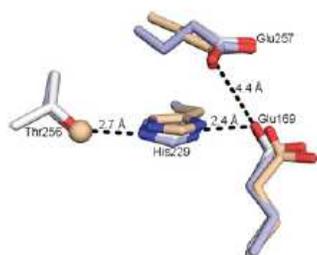
**Figure S7**  $^{15}\text{N}$ - $^1\text{H}$ -HSQC spectra of CBM3 as isolated (black) and in the presence of cellopentaose (overlaid in red)

The residues perturbed by the ligand are labelled. The inset shows a magnification of a representative spectral region with chemical shift changes.



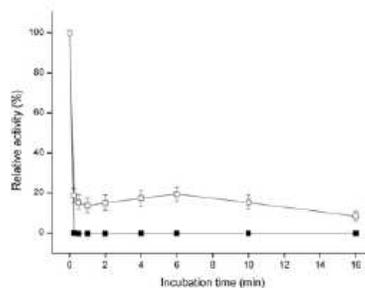
**Figure S8** SAXS scattering curves (top panels) and normalized distance distribution  $p(r)$  functions (bottom panels)

(A) Catalytic core, (B) CBM3 and (C) full-length protein.



**Figure S9** Schematic representation of the three possible catalytic triads in cellulases belonging to family 5A

Carbon atoms are coloured in white, blue and brown for type A, B and C respectively.



**Figure S10** Thermotolerance of BsCel5A at 90 °C in the absence (■) or in the presence (□) of manganese ion

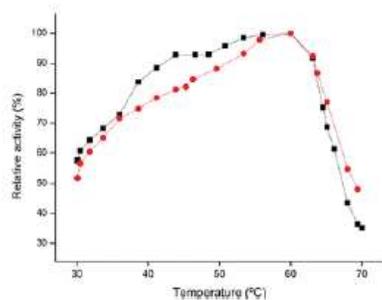


Figure S11 Temperature-dependent activity profile of BsCel5A in the presence (black) and in the absence (red) of manganese ion

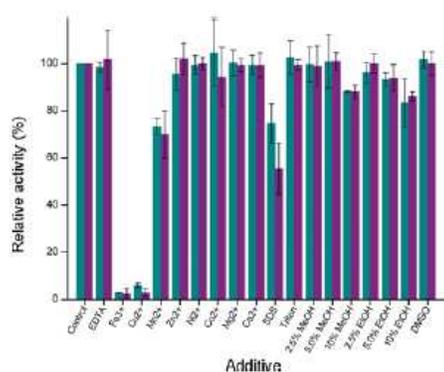


Figure S12 Effect of different additives on the endo-1,4-β-glucanase activity of BsCel5A

Table S2 Summary of the NMR data for *Bacillus* CBM3

RMSD, root mean square deviation.

Factor	Result
NOES	
Total	1373
Sequential	672
Medium-range distance	105
Long-range distance	596
Ramachandran plot	
Favoured region (%)	97.4
Generously allowed region (%)	2.6
Disallowed region (%)	0.0
RMSD (Å)	
Backbone atoms (full-length*)	0.91
Backbone atoms (structured†)	0.66
Non-hydrogen atoms (full-length)	1.66
Non-hydrogen atoms (structured)	1.53

\*Full-length CBM3d comprises the amino acid segment from 351 to 499.  
†Structured regions are the segments 354–359, 373–378, 388–395, 402–408, 417–422, 431–438, 452–457, 467–473, 484–488 and 491–496.

Table S1 Central composite design matrix (2<sup>2</sup>) and glucanase activity after 10 min of incubation

Run number	Coded levels ( $X_1 = \text{pH}; X_2 = T$ )		Actual levels ( $X_1 = \text{pH}; X_2 = T$ )		Glucanase activity (IU/mg protein)
	$X_1$	$X_2$	$X_1$	$X_2$	
1	-1	-1	4.6	74.2	139.8
2	1	-1	7.4	74.2	170.9
3	-1	1	4.6	94.8	175.2
4	1	1	7.4	94.8	204.2
5	-1.414	0	4.0	84.5	68.6
6	1.414	0	8.0	84.5	154.6
7	0	-1.414	6.0	70.0	159.7
8	0	1.414	6.0	99.0	238.9
9	0	0	6.0	84.5	234.6
10	0	0	6.0	84.5	241.3
11	0	0	6.0	84.5	247.8
12	0	0	6.0	84.5	247.1