

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL  
FACULDADE DE BIOCÊNCIAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

**Bruno Lopes Abbadi**

**Validação da proteína MDP2 de *Mycobacterium tuberculosis* H37Rv como uma proteína intrinsecamente desordenada e geração de uma cepa nocaute para o gene *hns***

Porto Alegre

2014

BRUNO LOPES ABBADI

**VALIDAÇÃO DA PROTEÍNA MDP2 DE *Mycobacterium tuberculosis* H37Rv COMO UMA PROTEÍNA INTRINSECAMENTE DESORDENADA E GERAÇÃO DE UMA CEPA NOCAUTE PARA O GENE *hns***

Dissertação apresentada ao Programa de Pós-graduação em Biologia Celular e Molecular da PUCRS como requisito para a obtenção do grau de Mestre em Biologia Celular e Molecular.

Orientador: Prof. Dr. Cristiano Valim Bizarro

Porto Alegre

2014

BRUNO LOPES ABBADI

**VALIDAÇÃO DA PROTEÍNA MDP2 DE *Mycobacterium tuberculosis* H37Rv COMO UMA PROTEÍNA INTRINSECAMENTE DESORDENADA E GERAÇÃO DE UMA CEPA NOCAUTE PARA O GENE *hns***

Dissertação apresentada ao Programa de Pós-graduação em Biologia Celular e Molecular da PUCRS como requisito para a obtenção do grau de Mestre em Biologia Celular e Molecular.

BANCA EXAMINADORA:

---

Mauricio Reis Bogo - PUCRS

---

Carlos Alexandre Sanchez Ferreira - PUCRS

---

Giancarlo Pasquali – UFRGS

Porto Alegre  
2014

## AGRADECIMENTOS

Aos professores Cristiano V. Bizarro, Diógenes S. Santos e Luiz A. Basso por possibilitarem o desenvolvimento deste trabalho no CPBMF e pelos ensinamentos que me passaram ao longo destes anos.

Aos colegas do CPBMF que me ajudaram a desenvolver e concluir este trabalho, e que me apoiaram nos momentos mais *tormentosos* da pesquisa.

À Zilpa Adriana Sanchez Quitian, que me acompanhou e me auxiliou desde o meu período de iniciação científica.

À Anne D. Villela, que me auxiliou na execução de boa parte deste trabalho.

A minha família pelo apoio dado durante estes dois anos.

A CAPES pelo apoio financeiro fornecido.

*“Chaos isn’t a pit. Chaos is a ladder. Many who try to climb it fail and never get to try again. The fall breaks them. And some are given a chance to climb, but they refuse. They cling to the realm, or the gods, or love. Illusions. Only the ladder is real. The climb is all there is.”*

***Littlefinger (George R. R. Martin)***

## RESUMO

A proteína micobacteriana ligadora de DNA 2 (MDP2) de *Mycobacterium tuberculosis* é uma proteína pequena e de caráter básico, conhecida por se ligar ao DNA de uma forma não específica e de ancorar o nucleóide à membrana plasmática promovendo o seu desempacotamento. Motivados por uma predição de desordem intrínseca *in silico* prévia, nós usamos um conjunto de técnicas complementares para caracterizar esta proteína, tais como determinação da estabilidade ao calor e a desnaturantes químicos, gel filtração, proteólise limitada e mobilidade eletroforética. Nossos resultados sugerem que a MDP2 é estruturalmente desordenada, uma vez que ela (1) foi predita por ser uma proteína intrinsecamente desordenada (IDP), possuindo 86 % da sua estrutura desordenada; resistiu à desnaturação induzida por temperatura de fervura e pela ação química do ácido tricloroacético (TCA); (2) migrou anormalmente em gel de poliacrilamida (SDS-PAGE), aparentando ser 1,4 vezes maior que o seu peso molecular calculado; e (3) foi altamente sensível à proteólise realizada pela protease K, em comparação às proteínas globulares. Nós também realizamos dois experimentos para determinar a estrutura quaternária da proteína, como a gel filtração e o *crosslinking* por glutaraldeído. Estes dois experimentos mostraram que a proteína tem uma tendência a auto agregação, formando grandes aglomerados em solução. Também foi realizada a mutagênese sítio-dirigida por troca alélica no gene *hns*, com o intuito de desenvolver uma cepa de *M. tuberculosis* sem a proteína MDP2. O resultado deste nocaute mostrou que o gene *hns* não é essencial para a sobrevivência da micobactéria. Combinado com resultados prévios de outros trabalhos relacionados à MDP2, nós especulamos que esta proteína usa sua região N-terminal altamente desordenada para se ligar ao DNA por meio das suas sequências repetitivas KAAK e PAKK de uma forma não específica. Desta forma, a MDP2 deve atuar como uma proteína promíscua dentro da célula, ligando-se a diferentes regiões do nucleóide pela formação de grandes aglomerados de proteína, com o objetivo de realizar o desempacotamento do DNA e de regular diversos genes micobacterianos. Nós esperamos que este trabalho possa contribuir para um melhor entendimento do metabolismo micobacteriano, uma vez que o *M. tuberculosis* ainda é uma grande ameaça global.

**Palavras-chave:** Migração Anômala; Crosslinking; Proteínas Intrinsecamente Desordenadas; Gel filtração; Resistência ao calor; Nocaute; Proteólise Limitada; *Mycobacterium tuberculosis*; Espectrometria de Massa; Purificação de Proteínas.

## ABSTRACT

The mycobacterial DNA-binding protein 2 (MDP2) of *Mycobacterium tuberculosis* is a small, basic protein known to bind to DNA in a non-specific manner and to anchor the nucleoid to the plasma membrane, promoting its unpacking. Motivated by a prior intrinsic disorder prediction *in silico*, we have used a complementary set of techniques to characterize this protein, such as heat and chemical stability, gel filtration, limited proteolysis and electrophoretic mobility. Our results suggest that the MDP2 is structurally disordered, since it (1) was predicted to be an IDP, having 86 % of its structure disordered; (2) resisted to denaturation induced by boiling temperature and to the chemical trichloroacetic acid (TCA); (3) migrated anomalously on SDS-PAGE, appearing to be 1.4-fold higher its calculated molecular weight; and (4) was highly sensitive to proteolysis performed by proteinase K in comparison to globular proteins. We also accomplished two experiments to determine the quaternary structure of the MDP2, such as gel-filtration and glutaraldehyde crosslinking. These two experiments showed that the protein has a tendency of self-aggregation, forming large clusters of protein. We also performed the site-directed mutagenesis by allelic exchange in the *hns* gene, in order to develop a *M. tuberculosis* strain lacking the protein MDP2. The result of this knockout showed that the *hns* gene is not essential for the survival of the mycobacteria, confirming previous results performed by transposon mutagenesis. Combined with previous results related to MDP2 from other works, we speculate that this protein uses its highly disorder N-terminal region to bind to DNA through its KAAK and PAKK sequences in a non-specific manner. Thereby the MDP2 may act as a promiscuous protein inside the cell, binding to different regions of nucleoid by forming large clusters of proteins, in order to perform the DNA unpacking and to regulate several genes of the mycobacteria. We hope this work may contribute to a better understanding of the mycobacterial metabolism, since the *M. tuberculosis* still stands a major global threat.

**Keywords:** Anomalous Migration; Crosslinking; Intrinsically Disordered Proteins; Gel filtration; Heat resistance; Knockout; Limited Proteolysis; *Mycobacterium tuberculosis*; Mass Spectrometry; Protein Purification.

## LISTA DE FIGURAS

Figura 1. Microscopia de alta resolução (x 1000) do <i>Mycobacterium tuberculosis</i> .....	11
Figura 2. Imagem esquemática da parede celular da micobactéria. ....	12
Figura 3. Cromossomo bacteriano com os seus domínios topológicos superenovelados. ....	15
Figura 4. Alterações estruturais do DNA ocasionadas pela ligação de algumas NAPs. ....	17
Figura 5. Estruturas cristalográficas das proteínas HU de <i>Anabaena</i> e IHF de <i>E. coli</i> complexadas com DNA. ....	19
Figura 6. Alterações estruturais no DNA ocasionadas pela ligação das proteínas HU e H-NS.....	20
Figura 7. Papel da proteína IHF de <i>E. coli</i> na modulação da estrutura do DNA.....	21
Figura 8. Esquema de um cromossomo bacteriano superenovelado com proteínas FIS. ....	22
Figura 9. Organização estrutural e funcional da proteína H-NS de <i>E. coli</i> . ....	24
Figura 10. Mecanismo proposto para a repressão gênica ocasionada pelas proteínas H-NS em alguns promotores. ....	25
Figura 11. Enovelamento do domínio fosforilado induzível por quinase (pKID) induzido pela ligação ao domínio de ligação KID (KIX) da proteína de ligação a CREB (CBP).....	33
Figura 12. Plasmídeo termosensível pPR27 contendo o marcador contra-seletivo <i>sacB</i> .....	56
Figura 13. Sequência $\Delta hns:Can^R$ utilizada como sítio para a mutagênese por troca alélica.....	57
Figura 14. Eletroforese em gel de agarose dos produtos de PCR das regiões flangeadoras <i>upstream</i> e <i>downstream</i> ao gene <i>hns</i> . ....	57
Figura 15. Eletroforese em gel de agarose mostrando a liberação da sequência $\Delta hns:Can^R$ a partir do vetor pPR27: <i>xylE</i> pela clivagem da enzima <i>SpeI</i> . ....	59
Figura 16. Colônias amarelas de <i>M. tuberculosis</i> transformadas com o vetor pPR27: <i>xylE</i> : $\Delta hns:Can^R$ após serem testadas com catecol. ....	60
Figura 17. Colônias brancas de <i>M. tuberculosis</i> após serem testadas com 1% catecol.....	61
Figura 18. Eletroforese em gel de agarose do produto de PCR dos 10 clones selecionados de <i>M. tuberculosis</i> que provavelmente sofreram o processo de recombinação homóloga.....	63
Figura 19. Eletroforese em gel de agarose do produto de PCR do gene <i>hns</i> selvagem flangeado pelas regiões <i>upstream</i> e <i>downstream</i> . ....	64
Figura 20. Colônias amarelas da cepa complementada de <i>M. tuberculosis</i> contendo o plasmídeo pPR27: <i>xylE</i> : $\Delta hns:Can^R$ . ....	65
Figura 21. Colônias brancas de <i>M. tuberculosis</i> mutantes para o gene <i>hns</i> , mas complementadas com uma cópia extra que foi integrada em outra porção do genoma. ....	66
Figura 22. Eletroforese em gel de agarose do produto de PCR dos quatro clones selecionados de <i>M. tuberculosis</i> , complementado com a cópia extra do gene <i>hns</i> , e que sofreram o processo de recombinação homóloga. ....	67

## SUMÁRIO

CAPÍTULO 1 – INTRODUÇÃO E OBJETIVOS.....	10
1 INTRODUÇÃO .....	11
1.1 O bacilo de Koch e a tuberculose humana.....	11
1.2 A estrutura do nucleóide bacteriano .....	14
1.3 Proteínas associadas ao nucleóide bacteriano .....	16
1.4 Proteína MDP2 de <i>Mycobacterium tuberculosis</i> .....	26
1.5 Proteínas Intrinsecamente Desordenadas (IDPs).....	31
2 JUSTIFICATIVA .....	35
3 OBJETIVOS.....	37
3.1 Objetivo Geral .....	37
3.2 Objetivos Específicos.....	37
CAPÍTULO 2 – ARTIGO CIENTÍFICO .....	38
CAPÍTULO 3 – RESULTADOS PRELIMINARES E PERSPECTIVAS .....	54
CAPÍTULO 4 – CONSIDERAÇÕES FINAIS.....	71
REFERÊNCIAS.....	76
ANEXO A.....	84

## **CAPÍTULO 1**

### **INTRODUÇÃO E OBJETIVOS**

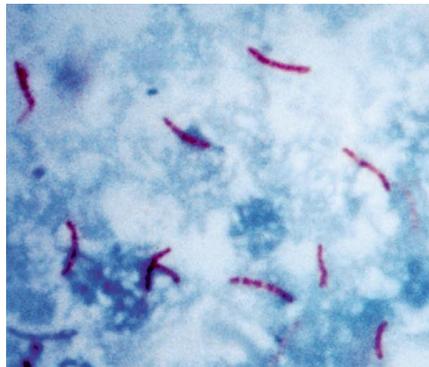
## 1 INTRODUÇÃO

### 1.1 O bacilo de Koch e a tuberculose humana

Em 1882, Robert Koch apresentou ao mundo um dos patógenos mais bem sucedidos que a humanidade já conheceu: o *Mycobacterium tuberculosis*, também conhecido como bacilo de Koch, o principal agente causador da tuberculose em humanos (Lawn e Zumla, 2011). Esta bactéria faz parte de um complexo também chamado de *M. tuberculosis*, no qual estão contidas ainda as demais bactérias que também causam a tuberculose humana, como o *Mycobacterium africanum*, *M. bovis*, *M. microti* e *M. canetti* (Lawn e Zumla, 2011).

O *M. tuberculosis* [Figura 1] é um patógeno intracelular que habita principalmente os macrófagos alveolares do hospedeiro, devido a sua preferência por ambientes ricos em oxigênio. Além disso, esta bactéria não apresenta mobilidade, não é capaz de formar uma cápsula e também não forma esporos (Lawn e Zumla, 2011), características que poderiam conferir-lhe certa proteção em ambientes hostis. A capacidade de formar esporos, entretanto, já foi sugerida para algumas espécies do gênero, como *M. marinum* e *M. bovis* (BCG) (Singh, Ghosh *et al.*, 2010).

**Figura 1.** Microscopia de alta resolução (x 1000) do *Mycobacterium tuberculosis*.



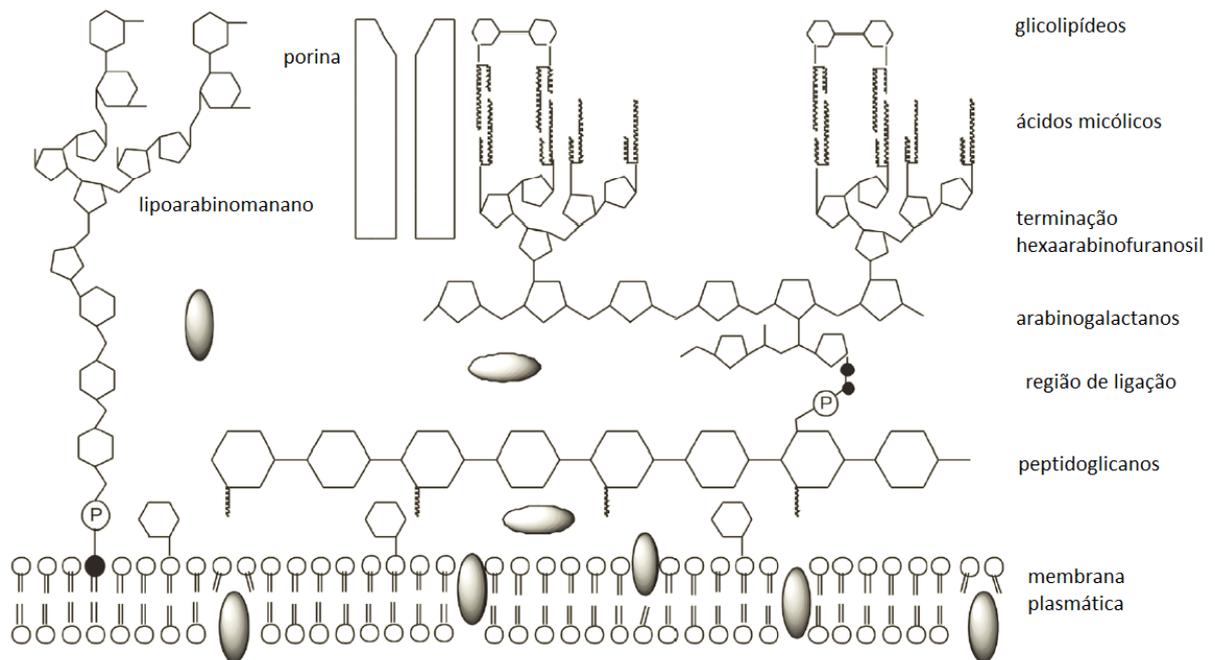
**Fonte:** Lawn e Zumla, 2011.

**Nota:** os bacilos foram coletados a partir do escarro de um paciente contaminado e corados pela técnica de Ziehl-Neelsen.

No entanto, o *M. tuberculosis*, assim como as demais espécies do gênero, apresenta uma característica bastante peculiar: sua parede celular exibe uma composição de lipídeos que a torna bastante impermeável, se comparada a de outras bactérias (Lawn e Zumla, 2011) [Figura 2]. A parede celular da micobactéria possui duas regiões distintas: uma mais interna, próxima à membrana plasmática, composta por peptidoglicanos covalentemente conecta-

dos a arabinogalactanos, os quais, por sua vez, estão ligados aos ácidos micólicos; e uma mais externa, composta por lipídios livres de cadeias longas e curtas. Intercalados entre estas suas camadas estão presentes diversas proteínas (p. ex.: porinas, HLP), que participam da montagem e organização desta estrutura, além de desempenhar outras funções importantes ao bacilo (Brennan, 2003; Katsube, Matsumoto *et al.*, 2007). Graças a sua parede celular incomum, a micobactéria resiste à dessecação, à alcalinidade e também a desinfetantes químicos (Ducati, Ruffino-Neto *et al.*, 2006). Além disso, a parede dificulta a entrada de diversos fármacos anti-tuberculose e permite a sobrevivência do bacilo dentro dos macrófagos alveolares (Ducati, Ruffino-Neto *et al.*, 2006).

**Figura 2.** Imagem esquemática da parede celular da micobactéria.



**Fonte:** Adaptado de Ducati, Ruffino-Neto *et al.*, 2006.

**Nota:** neste esquema é possível observar que a parede possui uma região mais interna, composta por peptidoglicanos e arabinogalactanos, e uma mais externa formada principalmente pelos ácidos micólicos. Nesta figura também está representada a porina, uma proteína que funciona como canal catiônico seletivo da parede.

A contaminação pelo bacilo da tuberculose ocorre principalmente por meio de aerossol expelido por um indivíduo contaminado. Uma vez dentro do hospedeiro, a micobactéria chega aos pulmões, onde é fagocitada pelos macrófagos alveolares (Russell, 2007). Contudo, as chances de uma pessoa contaminada vir a desenvolver a doença na sua forma ativa, ou seja, aquela em que o bacilo consegue se replicar dentro da célula e causar lesões ao

hospedeiro, são bastante baixas (aproximadamente 5 %) (Ducati, Ruffino-Neto *et al.*, 2006; Lawn e Zumla, 2011). Na maior parte dos casos, o bacilo permanece dentro do hospedeiro num estado não-replicativo, causando a forma latente da doença (Ducati, Ruffino-Neto *et al.*, 2006; Koul, Arnoult *et al.*, 2011). Apesar de não causar danos ao hospedeiro, nem provocar sintomas da doença, os indivíduos com tuberculose latente representam um importante reservatório da doença, visto que o número de pessoas nesta situação chega a dois bilhões em todo o mundo (Lawn e Zumla, 2011).

A grande demanda por pesquisas que envolvam o estudo da tuberculose e do seu agente causador se dá principalmente pelos alarmantes números que acompanham a doença. Segundo a Organização Mundial da Saúde (OMS), estima-se que no ano de 2012 ocorreram 8,6 milhões de novos casos da tuberculose e 1,3 milhões de mortes causadas por esta doença globalmente (WHO, 2013). Entre os casos incidentes, aproximadamente 13 % estavam contaminados pelo vírus da imunodeficiência humana (HIV), o qual é considerado um dos principais fatores para o ressurgimento da tuberculose no final do século passado (Ducati, Ruffino-Neto *et al.*, 2006; WHO, 2013). Diante destes números preocupantes, a tuberculose é considerada a segunda maior causa de mortes ocasionadas por uma doença infecciosa mundialmente, atrás apenas da AIDS (WHO, 2013).

Outro grande obstáculo para o combate à tuberculose é o surgimento de cepas resistentes aos fármacos anti-tuberculose disponíveis no mercado. Atualmente, o tratamento da doença nos sistemas públicos de saúde segue as recomendações da OMS, a qual lançou uma nova estratégia denominada “*Stop TB Strategy*”, a qual também utiliza o tratamento diretamente observado pelos agentes de saúde (do inglês *DOTS - Directly Observed Treatment*) e que visa um mundo livre de tuberculose (Lawn e Zumla, 2011, WHO, 2013). A fase inicial (também chamada de ataque) do *DOTS* envolve a administração das drogas isoniazida, rifampicina, pirazinamida e etambutol por dois meses; já a segunda fase do tratamento segue por mais quatro meses, com as drogas isoniazida e rifampicina (Koul, Arnoult *et al.*, 2011). Visto que o tratamento é longo, requer o uso de diversos fármacos em conjunto e confere alguns efeitos colaterais aos pacientes, muitos indivíduos o abandonam ou o fazem de forma incorreta, ocasionando a sobrevivência e a prevalência de micobactérias persistentes (Ducati, Ruffino-Neto *et al.*, 2006; Lawn e Zumla, 2011). Até o momento, existem três principais cepas resistentes de *M. tuberculosis*: a MDR-TB (tuberculose multirresistente a drogas),

que apresenta resistência a pelo menos isoniazida e rifampicina; a XDR-TB (tuberculose extensivamente resistente a drogas), resistente a isoniazida, rifampicina, qualquer fluoroquinolona e qualquer droga anti-tuberculose injetável de segunda-linha (amicacina, canamicina, capreomicina); e a TDR-TB (totalmente resistente a drogas), a qual é resistente a qualquer droga utilizada no tratamento da tuberculose (Gandhi, Nunn *et al.*, 2010; Koul, Arnoult *et al.*, 2011; Udwadia, Amale *et al.*, 2012).

De posse destes dados, torna-se fundamental o contínuo estudo da tuberculose e do próprio *M. tuberculosis*, para que novas estratégias de combate à doença possam ser delineadas futuramente. Diversas abordagens têm sido realizadas para entender a biologia deste patógeno, como, por exemplo, o estudo das suas enzimas, com a finalidade de se desenvolver inibidores enzimáticos que possam comprometer a sobrevivência do bacilo (Ducati, Breda *et al.*, 2011); da sua parede e membrana celular, a fim de buscar antígenos imunodominantes, alvos para drogas e biomarcadores da doença (Wolfe, Mahaffey *et al.*, 2010); e também do seu nucleóide e das suas proteínas associadas, com o objetivo de entender a expressão gênica do bacilo em diferentes situações e encontrar alvos genéticos e proteicos que possam afetar o crescimento do *M. tuberculosis* (Mdluli e Ma, 2007; Werlang, Schneider *et al.*, 2009).

## 1.2 A estrutura do nucleóide bacteriano

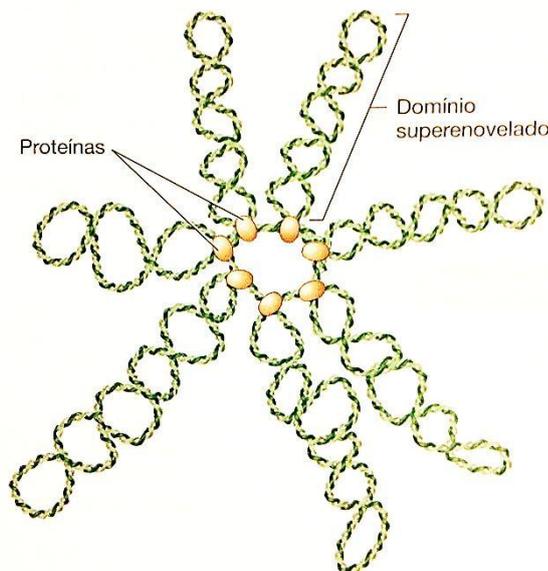
Até recentemente as bactérias eram consideradas seres de baixa complexidade celular, devido em parte a sua origem evolutiva ancestral - em torno de 3,8 bilhões de anos atrás -, ao seu diminuto genoma - 4,6 milhões de pares de bases em *Escherichia coli*, por exemplo - e as suas pequenas dimensões celulares - com diâmetros que variam de 0,2 até 700  $\mu\text{m}$  (Alberts, Johnson *et al.*, 2010). No entanto, diferente do que se pensava há algumas décadas atrás, as bactérias são organismos dotados de uma altíssima organização celular, a qual envolve uma complexa rede de circuitos regulatórios, estruturas citoesqueléticas dinâmicas e até mesmo localização espacial e temporal de proteínas - uma característica que era atribuída somente às células eucarióticas (Thanbichler, Wang *et al.*, 2005; Madigan, Martinko *et al.*, 2010).

Esta mudança de paradigma foi induzida pelos recentes avanços na área de biologia celular microbiana, impulsionados principalmente pelos estudos da organização e dinâmica da

cromatina bacteriana. Sabe-se que as bactérias são dotadas, na sua maioria, de um cromossomo circular altamente condensado, o qual não é separado do citoplasma por um compartimento membranosos, como ocorre nas células eucarióticas (Egan, Fogel *et al.*, 2005; Thanbichler, Wang *et al.*, 2005). Entretanto, apesar da cromatina não estar confinada em uma membrana, ela não fica totalmente livre no citoplasma; esta estrutura é encontrada em uma região livre de ribossomos denominada de nucleóide - uma área funcionalmente equivalente ao núcleo eucariótico (Robinow e Kellenberger, 1994; Thanbichler, Wang *et al.*, 2005).

Estudos envolvendo o nucleóide de *E. coli* mostraram que a cromatina é compactada por superenovelamentos, que formam diversos domínios cromossomais topologicamente não relacionados (Worcel e Burgi, 1972; Thanbichler, Wang *et al.*, 2005; Madigan, Martinko *et al.*, 2010). Entende-se por domínios topológicos qualquer segmento de DNA que não possui uma rotação livre das suas extremidades (Mirkin, 2001). Estima-se que no cromossomo de *E. coli* existam mais de 100 domínios topológicos independentes (Madigan, Martinko *et al.*, 2010). Além disso, consta na literatura que o cromossomo bacteriano é organizado em voltas plectonemicamente enroladas de DNA (do inglês *plectonemically wound loops of DNA*), as quais irradiam de maneira radial a partir de uma região central (Thanbichler, Wang *et al.*, 2005) [Figura 3].

**Figura 3.** Cromossomo bacteriano com os seus domínios topológicos superenovelados.



**Fonte:** Madigan, Martinko *et al.*, 2010.

**Nota:** nesta figura estão representados, de forma esquemática, sete domínios independentes, os quais são mantidos por diversas proteínas ancoradas ao DNA. Esta conformação permite que modificações topológicas em um domínio específico não interfiram nos seus domínios adjacentes.

Sabe-se que esta organização cromossômica em voltas influencia diretamente a atividade transcricional em determinadas regiões da cromatina. Por exemplo, bactérias que estão em crescimento acelerado têm sua atividade transcricional mais ativa, principalmente dos genes que codificam componentes ribossomais e demais elementos da maquinaria traducional, e apresentam mais domínios de DNA em voltas; por outro lado, quando o crescimento diminui, devido à chegada da fase estacionária de crescimento ou à falta de nutrientes no meio, ocorre uma redução da atividade transcricional, a qual está relacionada ao relaxamento dos domínios em volta do nucleóide (Deng, Stein *et al.*, 2005; Dillon e Dorman, 2010).

Já foi proposto que a organização da cromatina em domínios topológicos traz algumas vantagens à célula bacteriana. Caso haja uma lesão ao cromossomo, ocasionada por danos ao DNA ou eventos de reparo e replicação, por exemplo, haverá somente o relaxamento do domínio topológico que contém a sequência lesionada, não interferindo na super-helicidade dos demais domínios. Esta característica é extremamente importante para a sobrevivência da célula, visto que mudanças globais na super-helicidade do cromossomo são letais (Gellert, O'dea *et al.*, 1976; Thanbichler, Wang *et al.*, 2005). Além disso, este arranjo em domínios permite que a cromatina seja extremamente compactada, por volta de 1000 vezes, caso contrário não seria possível manter o material genético dentro da célula (Thanbichler, Wang *et al.*, 2005; Jin e Cabrera, 2006).

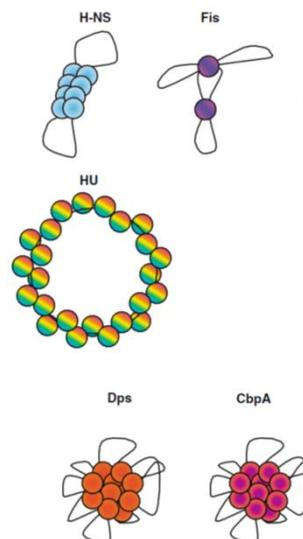
### **1.3 Proteínas associadas ao nucleóide bacteriano**

Visto que o cromossomo bacteriano é uma estrutura dinâmica, alguns fatores podem alterar sua estrutura direta ou indiretamente. Por exemplo, durante a transcrição, o complexo da RNA polimerase gera domínios de superenrolamentos positivos e negativos à frente e atrás da bolha transcricional; conseqüentemente, isso altera indiretamente a estrutura da cromatina em determinados pontos (Thanbichler, Wang *et al.*, 2005). Por outro lado, diversas proteínas atuam diretamente na modulação da estrutura da cromatina bacteriana. Estas

proteínas eram conhecidas inicialmente como “proteínas semelhantes a histonas” (do inglês *histone-like proteins*), devido ao seu papel análogo às histonas eucarióticas. Contudo, atualmente este termo foi alterado para “proteínas associadas ao nucleóide” (do inglês NAPs – *Nucleoid-Associated Proteins*), que reflete claramente a sua localização celular na bactéria (Dillon e Dorman, 2010). Estas proteínas são geralmente básicas e pequenas (entre 16-20 kDa de peso molecular) e estão presentes em grandes quantidades dentro da célula bacteriana (Anuchin, Goncharenko *et al.*, 2011).

A maioria das NAPs é capaz de se ligar à molécula de DNA, podendo alterar sua conformação por diferentes mecanismos, como: dobra (*bending*), enrolamento (*wrapping*) ou formação de pontes (*bridging*) entre hélices-duplas de DNA (Dillon e Dorman, 2010) [Figura 4]. Por meio destes mecanismos, as NAPs alteram o nível de superenrolamentos do DNA, tanto positivos quanto negativos, podendo, assim, regular a expressão de determinados genes, a recombinação e, até mesmo, a replicação do DNA (Dillon e Dorman, 2010; Anuchin, Goncharenko *et al.*, 2011).

**Figura 4.** Alterações estruturais do DNA ocasionadas pela ligação de algumas NAPs.



**Fonte:** Browning, Grainger *et al.*, 2010.

**Nota:** as NAPs podem realizar diversas alterações estruturais no nucleóide; enquanto as proteínas HU são capazes de abrir o DNA circular, as proteínas H-NS reduzem o tamanho desta estrutura e aproximam regiões distantes, por exemplo.

O número de NAPs conhecidas está constantemente aumentando e diferentes classes destas proteínas têm sido propostas. Existem quatro grandes grupos de NAPs bem esta-

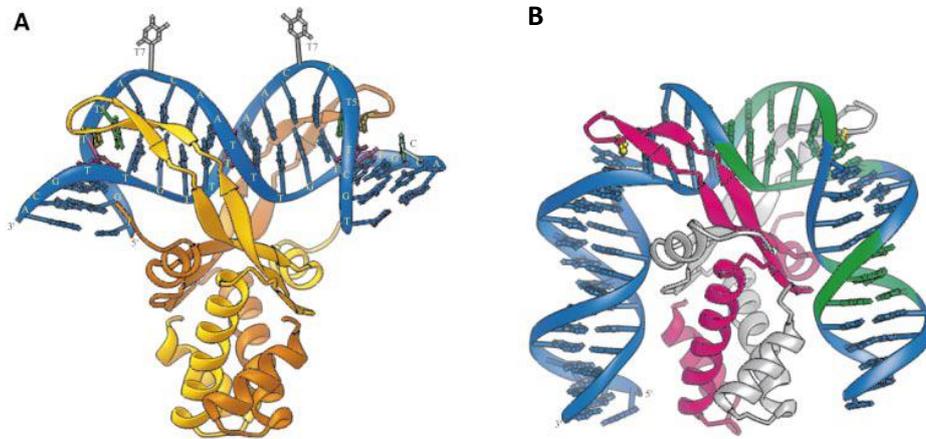
belecidos, os quais são constituídos por proteínas que possuem alta identidade nas suas sequências proteicas. São eles: HU (proteínas semelhantes a histonas de *E. coli* U93), H-NS (proteínas estruturantes semelhantes a histonas), IHF (fator de integração ao hospedeiro) e FIS (fator para a estimulação da inversão) (Dillon e Dorman, 2010; Anuchin, Goncharenko *et al.*, 2011). Existem também outros grupos menores de NAPs que foram recentemente reportados, como: CbpA (proteínas de ligação ao DNA curvado), EbfC (proteína de ligação ao DNA), Dps (proteína de proteção ao DNA durante a falta de nutrientes), Crp (proteína regulatória de AMP cíclico), Lrp (proteína regulatória responsiva à leucina), MukB (proteína de partição do cromossomo) (Dillon e Dorman, 2010). Além destas proteínas, existem outras que não se encaixam nestes grupos, por falta de identidade entre suas sequências ou por exercerem funções ainda não reportadas, as quais podem representar novas classes de NAPs (Dillon e Dorman, 2010; Anuchin, Goncharenko *et al.*, 2011). A seguir serão apresentadas algumas características das proteínas dos quatro grupos principais de NAPs, evidenciando o seu importante papel na organização do nucleóide e na sobrevivência da célula bacteriana.

### 1.3.1 HU

As proteínas termoestáveis da família HU são proteínas básicas, pequenas - possuem um peso molecular de aproximadamente 20 kDa - e estão presentes em grandes quantidades dentro da bactéria (10.000 moléculas por célula) (Rouviere-Yaniv e Gros, 1975). Já foi reportado que estas proteínas são capazes de se ligar ao DNA e estimular a transcrição de DNA  $\lambda$  pela RNA polimerase em ensaios *in vitro* (Rouviere-Yaniv e Gros, 1975).

Consta na literatura que as proteínas HU são predominantemente encontradas como heterodímeros, formados pelas subunidades HU $\alpha$  e HU $\beta$  (Dame e Goosen, 2002). A estrutura do heterodímero consiste em um “corpo” formado por uma hélice- $\alpha$ , a qual é “coberta” por folhas- $\beta$  estendidas em dois “braços” de fitas que se ligam à ranhura menor do DNA (Swinger, Lemberg *et al.*, 2003; Swinger e Rice, 2004) [Figura 5].

**Figura 5.** Estruturas cristalográficas das proteínas HU de *Anabaena* e IHF de *E. coli* complexadas com DNA.

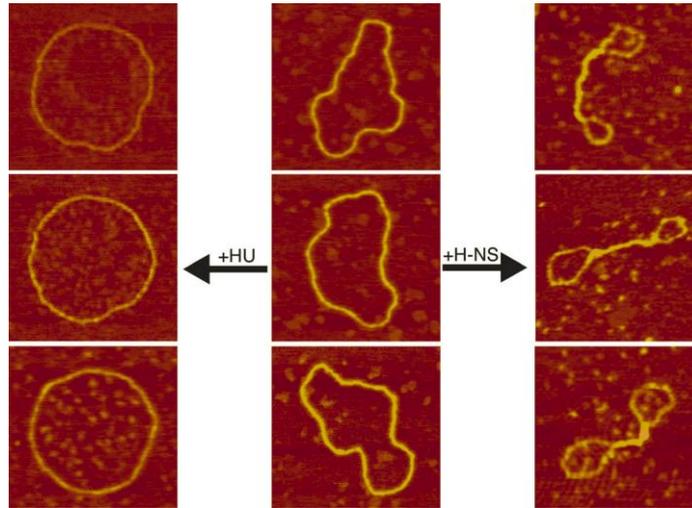


**Fonte:** Swinger, Lemberg *et al.*, 2003.

**Nota:** (A) Homodímero da proteína HU (em ouro e laranja) ligadas à ranhura menor do DNA (azul). (B) Heterodímero da proteína IHF (subunidade  $\alpha$  em branco e  $\beta$  em rosa) ligado à ranhura menor do DNA (azul e verde).

As proteínas HU têm preferência por moléculas de DNA superenroladas negativamente e com conformações atípicas, como nas junções de quatro fitas ou nas lesões de fita-simples, porém a ligação é independente de sequência (Dame e Goosen, 2002; Dorman e Deighan, 2003). Ao interagir com estas estruturas, elas são capazes de influenciar a superhelicidade do DNA, por meio da formação de dobras na dupla-hélice (Dillon e Dorman, 2010). Entretanto, diferente de outras NAPs, como a H-NS, as proteínas HU não são capazes de compactar o DNA bacteriano; ao contrário, em estudos utilizando plasmídeos, verificou-se que elas abrem moléculas de DNA circulares quando está em altas concentrações, deixando-os no formato de um anel (Dame e Goosen, 2002; Dorman e Deighan, 2003; Thanbichler, Wang *et al.*, 2005) [Figura 6].

**Figura 6.** Alterações estruturais no DNA ocasionadas pela ligação das proteínas HU e H-NS.



**Fonte:** Dame e Goosen, 2002.

**Nota:** experimento realizado com o plasmídeo pUC19 na presença de HU (à esquerda) e de H-NS (à direita), mostrando os papéis antagônicos que estas proteínas exercem no nucleóide bacteriano. Enquanto a HU, em altas concentrações, abre moléculas de DNA circular, a H-NS as compacta, deixando o DNA no formato de halteres.

As proteínas HU também são essenciais para o funcionamento da divisão celular e do crescimento bacteriano. Já foi reportado que não é possível obter células de *Streptococcus intermedius* nocautes para o gene que codifica a HLP, uma proteína homóloga a HU de *E. coli*, visto que a perda da função deste gene leva à incapacidade do crescimento e divisão celular, além de alterar a morfologia da célula e da parede celular (Liu, Yumoto *et al.*, 2008; Anuchin, Goncharenko *et al.*, 2011).

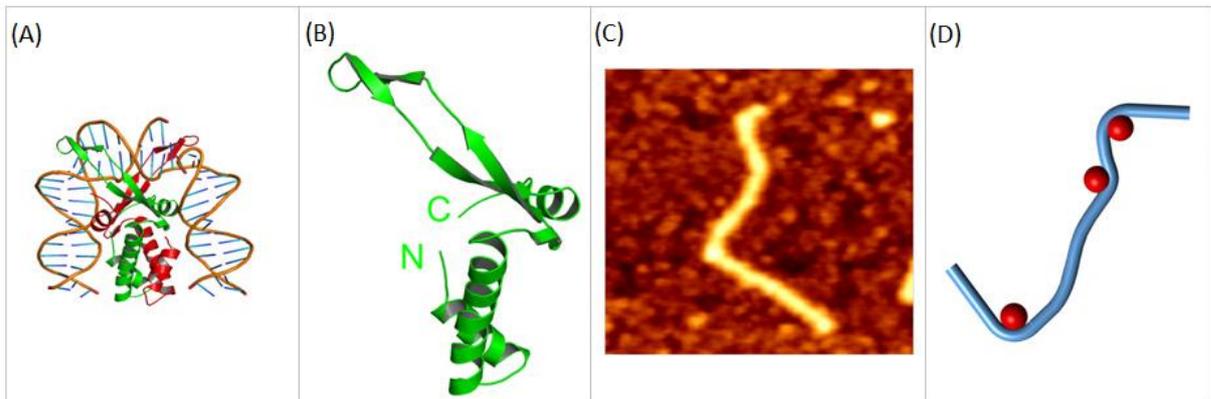
### 1.3.2 IHF

Outra família de NAPs encontrada nas bactérias apresenta homologia estrutural com as proteínas HU (de 30 a 40 % de identidade). Estas proteínas, chamadas de fatores de integração ao hospedeiro (IHF), são pequenas (20,5 kDa) e formam tanto heterodímeros quanto homodímeros com suas duas subunidades ( $\alpha$  e  $\beta$ ), assim como suas homólogas (Dhavan, Crothers *et al.*, 2002; Dillon e Dorman, 2010; Anuchin, Goncharenko *et al.*, 2011) [Figura 7]. A IHF possui uma estrutura semelhante à HU: um corpo compacto, formado pelas hélices- $\alpha$  de ambas as subunidades interenoveladas, coberto com folhas- $\beta$  que se estendem em dois braços flexíveis de fitas- $\beta$ ; são estes braços que fazem contato com o DNA pela ranhura me-

nor, por meio dos resíduos de prolina localizados na ponta de cada braço (Luijsterburg, Noom *et al.*, 2006).

As IHFs são proteínas associadas ao nucleóide que se ligam fortemente em sítios específicos do DNA, com aproximadamente 30 pares de base, como as sequências WATCARXXXTTR (onde W = A ou T; X = A, T, C ou G; e R = A ou G) (Luijsterburg, Noom *et al.*, 2006; Anuchin, Goncharenko *et al.*, 2011). Entretanto, elas também são capazes de se ligarem de forma inespecífica ao DNA, assim como as proteínas HU, mas com uma força de ligação bem mais fraca (Thanbichler, Wang *et al.*, 2005). Ao se ligar a sítios específicos de regiões promotoras, a IHF induz a formação de voltas no DNA (em inglês *U-turns*) que permitem a comunicação entre proteínas distantes [Figura 7]. Este fenômeno facilita, por exemplo, a interação da RNA polimerase, ligada a um gene, com outras proteínas regulatórias mais distantes (Dhavan, Crothers *et al.*, 2002; Dillon e Dorman, 2010; Anuchin, Goncharenko *et al.*, 2011).

**Figura 7.** Papel da proteína IHF de *E. coli* na modulação da estrutura do DNA.



**Fonte:** Adaptado de Luijsterburg, Noom *et al.*, 2006.

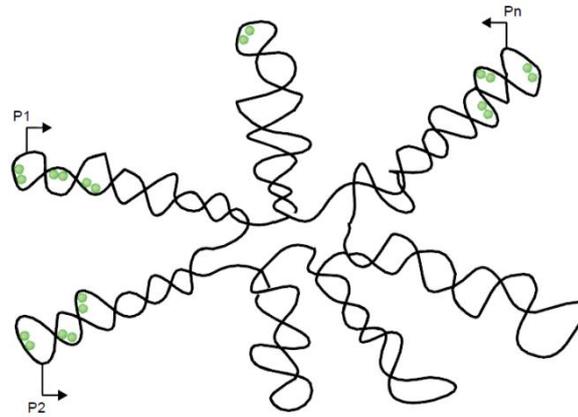
**Legenda:** (A) Complexo DNA-IHF. (B) Subunidade da IHF. (C) Microscopia de um fragmento de DNA sofrendo curvaturas pela interação com a IHF. (D) Representação esquemática da alteração da estrutura do DNA pela ligação com a IHF.

### 1.3.3 FIS

As proteínas FIS são formadas por duas subunidades idênticas, com 98 resíduos de aminoácidos cada. A porção N-terminal consiste em quatro hélices- $\alpha$ , que se ligam a duas hélices na porção C-terminal, sendo esta última a responsável por se ligar ao DNA; já foi verificado que mutações na região C-terminal da FIS de *E. coli* dificultam a ligação desta proteína ao DNA (Koch, Ninnemann *et al.*, 1991; Anuchin, Goncharenko *et al.*, 2011). Dentro da célula, estas proteínas se ligam a sítios específicos no DNA, onde promovem dobras nesta estru-

tura [Figura 8]. Entretanto, em altas concentrações, as proteínas FIS também são capazes de se ligarem ao DNA de maneira inespecífica (Anuchin, Goncharenko *et al.*, 2011).

**Figura 8.** Esquema de um cromossomo bacteriano superenovelado com proteínas FIS.



**Fonte:** Dorman e Deighan, 2003.

**Nota:** estas proteínas (em verde) são conhecidas por gerar dobras no DNA e por manter a estabilidade de contorções nos domínios topológicos. Nesta figura também estão representadas alguns promotores (P1, P2... Pn) nas regiões apicais dos domínios associados às proteínas FIS, as quais facilitam a interação destes com a RNA polimerase.

A produção da FIS dentro da célula bacteriana está sujeita a vários níveis de regulação, com a finalidade de manter a concentração desta proteína em quantidades apropriadas. A FIS de *E. coli*, por exemplo, está em altas concentrações durante a fase de crescimento exponencial e é a NAP mais abundante neste período, enquanto que sua concentração diminui radicalmente durante a fase estacionária (Schneider, Lurz *et al.*, 2001; Bradley, Beach *et al.*, 2007).

As proteínas FIS também estão envolvidas na alteração do superenrolamento do DNA, visto que elas aumentam o número de superenrolamentos negativos nesta estrutura; esta propriedade também permite que elas compactem o DNA (Schneider, Lurz *et al.*, 2001; Anuchin, Goncharenko *et al.*, 2011) [Figura 4]. Além do mais, estas NAPs também modulam a topologia do DNA indiretamente: por meio da repressão dos genes *gyrB* e *gyrA*, reduzindo a atividade da DNA girase, e pela modulação da expressão de outras NAPs, como a H-NS, HU e CRP (Bradley, Beach *et al.*, 2007).

Assim como outras NAPs, as proteínas FIS também atuam como proteínas regulatórias globais, modulando a expressão de diferentes genes, incluindo genes relacionados a RNA estáveis. Um estudo que utilizou microarranjos de DNA, por exemplo, mostrou que em célu-

las de *E. coli*, a proteína FIS modula, direta ou indiretamente, a expressão de pelo menos 330 genes, dos quais fazem parte operons de rRNA, tRNA, e promotores da família dos fatores  $\sigma_{70}$ ,  $\sigma_{28}$  e  $\sigma_{38}$  (Bradley, Beach *et al.*, 2007). Entre os genes superregulados pela FIS em *S. typhimurium*, encontram-se aqueles localizados nas ilhas de patogenicidade (SPI-1, SPI-2, SPI-3, SPI-4 e SPI-5), o que sugere que esta NAP está envolvida nos eventos que desencadeiam a virulência desta bactéria (Bradley, Beach *et al.*, 2007). Em *Shigella*, por exemplo, já foi demonstrado que a FIS inibe a repressão gênica induzida pela H-NS através do desligamento desta proteína ao DNA, deste modo, ativando genes de virulência, como o *virF* (Anuchin, Goncharenko *et al.*, 2011).

#### 1.3.4 H-NS

As proteínas H-NS são amplamente encontradas dentro do grupo das bactérias Gram-negativas, do qual fazem parte, por exemplo, *E. coli*, *S. typhimurium*, *Proteus vulgaris*, *Erwinia chrysanthemi*, *Serratia marcescens* e *Shigella flexneri* (Tendeng e Bertin, 2003). No grupo das Enterobacteriaceae, são encontradas algumas das proteínas semelhantes às proteínas H-NS (do inglês *H-NS-like*), como a BpH3 de *Bordetella pertussis*, a vicH de *Vibrio cholerae*, e a Pm70 de *Pausterella multocida*, as quais compartilham certa identidade com as H-NS (Tendeng e Bertin, 2003). É interessante observar que o número de cópias do gene *hns* ou *hns-like* varia de acordo com a espécie; em *E. coli* são encontradas três cópias, enquanto que em *Salmonella typhimurium* são encontradas apenas duas. Estes genes podem ser ligeiramente diferentes um do outro, permitindo que diferentes formas da H-NS possam ser geradas (Anuchin, Goncharenko *et al.*, 2011). Além disso, atribui-se às bactérias que tem uma alta diversidade ecológica e metabólica a característica de terem várias cópias destes genes no seu genoma (Tendeng e Bertin, 2003).

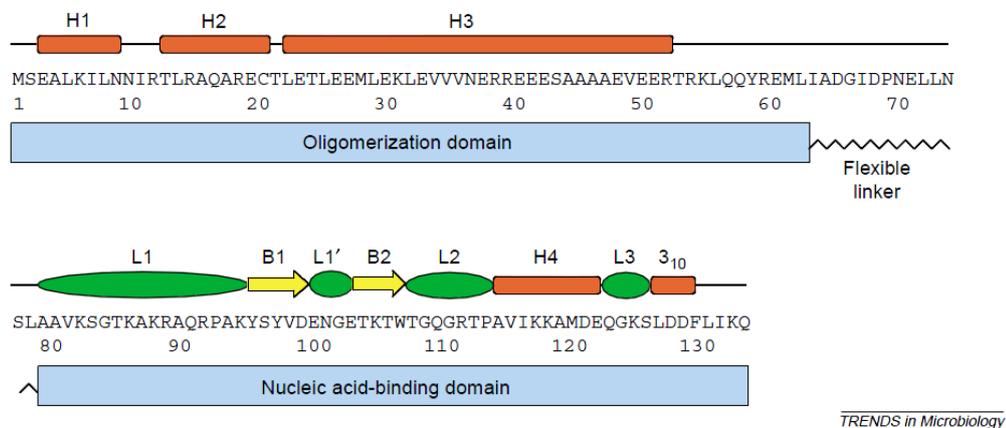
Análises *in silico* mostraram que em espécies de bactérias mais distantes filogeneticamente das Proteobacteria, os genes *hns* ou *hns-like* não estão presentes. Entre estas espécies, por exemplo, encontram-se *Bacillus subtilis*, *Mycoplasma genitalium*, *Streptococcus pneumoniae* e *Mycobacterium tuberculosis* (Tendeng e Bertin, 2003).

Assim como as proteínas HU, as HNS são bastante abundantes no nucleóide bacteriano (aproximadamente 20 mil cópias), onde elas influenciam a topologia desta estrutura e a expressão de diversos genes (Dorman, 2007). Estas proteínas são pequenas (em torno de 15,5

kDa), básicas, e consistem em dois domínios distintos: a porção N-terminal, com aproximadamente 64 resíduos, é responsável pela formação de dímeros (com outras proteínas, como a StpA, ou com a própria H-NS); já a porção C-terminal, formada por uma folha- $\beta$  antiparalela, uma hélice- $\alpha$  e uma hélice-3<sub>10</sub>, é responsável pela ligação ao DNA (Anuchin, Goncharenko *et al.*, 2011) [Figura 9]. Além de formar dímeros, estas NAPs são capazes de formar trímeros ou oligômeros ainda maiores (Dame e Goosen, 2002).

Estudos *in vitro* mostraram que a H-NS de *S. typhimurium* é capaz de formar superenrolamentos negativos no DNA plasmidial, ligando-se preferencialmente em regiões curvadas de DNA (Owen-Hughes, Pavitt *et al.*, 1992; Tupper, Owen-Hughes *et al.*, 1994). Estes superenrolamentos negativos facilitam a formação de pontes entre regiões distantes de DNA (*bridging*), permitindo que a H-NS influencie simultaneamente a estrutura do nucleóide e a expressão gênica (Dillon e Dorman, 2010) [Figura 4].

**Figura 9.** Organização estrutural e funcional da proteína H-NS de *E. coli*.



**Fonte:** Tendeng e Bertin, 2003.

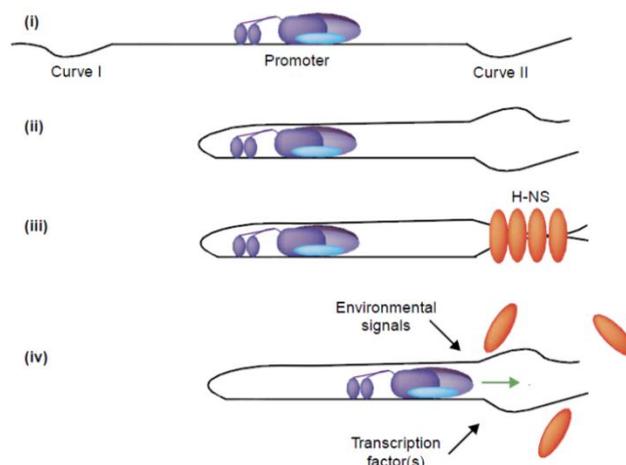
**Nota:** representação dos dois domínios da proteína: o domínio de oligomerização (N-terminal), formado por três hélices- $\alpha$  (laranja), o qual é ligado ao domínio de ligação ao DNA (C-terminal) por uma alça flexível. Este último domínio é formado por quatro *loops* flexíveis (L1, L1', L2, L3, em verde), sendo que o L1 e L2 interagem diretamente com o DNA, e também por duas folhas- $\beta$  (B1 e B2, em amarelo), uma hélice- $\alpha$  (H4, em laranja) e uma hélice- $\alpha$  3<sub>10</sub> (em laranja).

As proteínas H-NS são conhecidas pelos seus efeitos pleiotrópicos, atuando como repressores globais da transcrição. Em *E. coli*, estas proteínas, por exemplo, diminuem consideravelmente a atividade da RNA polimerase do fago T7, enquanto que as proteínas HU têm um efeito contrário (Anuchin, Goncharenko *et al.*, 2011). Já foi demonstrado em *E. coli* e *S.*

*typhimurium* que a H-NS tem um papel importante na organização do nucleóide e também na modulação da expressão de diversos genes envolvidos na adaptação às modificações ambientais e ao combate ao estresse (Tendeng e Bertin, 2003; Ono, Goldberg *et al.*, 2005). Em *S. typhimurium*, foi verificado que a maioria dos genes que tem sua expressão modulada pela temperatura também depende da ação da própria H-NS. Na ilha de patogenicidade SPI-1, por exemplo, a H-NS impede a expressão destes genes a uma temperatura de 25 °C, mas a 37 °C estes genes são expressos devido à dificuldade desta proteína em se ligar ao DNA nesta temperatura (Ono, Goldberg *et al.*, 2005).

Estudos em *V. cholerae* mostraram que a H-NS também atua como um repressor de diversos genes de virulência neste patógeno. Foi verificado que células que não possuíam o gene *hns* tiveram um aumento na expressão dos genes *toxT*, *tcpA* e *ctx* - os quais estão envolvidos nos processos de virulência -, sugerindo que na presença desta proteína, os promotores destes genes encontram-se reprimidos por ela (Nye, Pfau *et al.*, 2000). Os meios pelos quais as proteínas H-NS reprimem a expressão gênica podem ser, principalmente, de duas formas: impedindo diretamente a ligação da RNA polimerase ao promotor, visto que estas NAPs têm afinidade por regiões ricas em nucleotídeos AT, geralmente localizadas nas regiões promotoras; ou se ligando em regiões curvadas, também ricas em AT, fora do promotor, onde elas alteram a topologia do DNA (Tupper, Owen-Hughes *et al.*, 1994; Nye, Pfau *et al.*, 2000) [Figura 10].

**Figura 10.** Mecanismo proposto para a repressão gênica pelas proteínas H-NS em alguns promotores.



**Fonte:** Dorman e Deighan, 2003.

**Nota:** (I) A região promotora em destaque está flanqueada por duas regiões de DNA curvado (I e II), as quais são reconhecidas pelas H-NS. (II) As regiões curvas são aproximadas pela curvatura da dupla-fita em torno da

polimerase. (III) A H-NS reconhece as regiões curvadas e impede a movimentação da polimerase pelo gene. (IV) Alguns sinais externos são capazes de desfazer os complexos DNA-proteína, permitindo a transcrição gênica.

Visto que os genes de virulência encontrados em *V. cholerae* e *S. typhimurium* foram obtidos por transferência horizontal (entre bactérias) e que a H-NS se liga com alta eficiência às regiões promotoras destes genes, impedindo a sua expressão, foi proposto que esta proteína está envolvida no silenciamento de genes transferidos horizontalmente (Dorman, 2007). Esta propriedade permitiria que o gene estranho pudesse ser incorporado ao genoma de forma inerte e sem causar danos ao hospedeiro.

Diferentemente das proteínas FIS, a H-NS tem uma expressão constante dentro da célula. No entanto, ambas são reguladas por diversos fatores. A expressão da H-NS de *E. coli*, por exemplo, pode ser reprimida por ela mesma ou pela sua paróloga StpA. Já a proteína FIS atua como um ativador da transcrição do gene *hns*, bem como a proteína CspA, durante eventos de choque por frio (Dorman, 2007).

#### **1.4 Proteína MDP2 de *Mycobacterium tuberculosis***

Os primeiros estudos envolvendo a proteína MDP2 (Rv3852, NP\_218369.1) de *Mycobacterium tuberculosis* foram realizados por Werlang e colaboradores (Werlang, Schneider *et al.*, 2009), com o intuito de caracterizar bioquimicamente esta proteína de 134 aminoácidos. A partir da sequência do gene *hns*, reportada pelo sequenciamento completo do genoma da cepa H37Rv de *M. tuberculosis* (Cole, Brosch *et al.*, 1998), o grupo primeiramente expressou e purificou a proteína por cromatografia líquida.

Por meio de ensaios de alteração de mobilidade eletroforética (EMSA), utilizando-se a região promotora *proU* de *M. tuberculosis* como substrato, verificou-se que a MDP2 é capaz de se ligar ao DNA, apresentando maior afinidade por regiões curvadas. Além disso, foi possível identificar que esta proteína é capaz de formar múltiplos complexos DNA-proteína, por meio das interações entre os seus dímeros (estado oligomérico em solução desta proteína) ligados ao DNA (Werlang, Schneider *et al.*, 2009; Sharadamma, Harshavardhana *et al.*, 2010).

Em outro estudo no qual também foi utilizado o EMSA, verificou-se que a MDP2 se liga com alta afinidade em regiões do DNA com baixa quantidade de bases GC (40 %), enquanto que em regiões com alta concentração destas bases (70 %) a afinidade da proteína diminuiu consideravelmente. Neste mesmo experimento também foi possível verificar que a proteína

interage de forma distinta com o DNA fita-dupla e simples com baixo conteúdo GC, formando diferentes complexos DNA-proteína; entretanto, a ligação desta proteína à fita simples de DNA é bastante fraca. Estes mesmos resultados foram visualizados para a proteína H-NS de *E. coli* (Sharadamma, Harshavardhana *et al.*, 2010). Além disso, foi proposto que a MDP2 produz supertorções negativas no DNA, tornando-o mais compacto, o que poderia evidenciar o seu papel na organização estrutural do DNA bacteriano e também no silenciamento de genes (Sharadamma, Harshavardhana *et al.*, 2010).

Por meio de alinhamentos da sequência da MDP2 de *M. tuberculosis* com sequências de proteínas homólogas de outras espécies do gênero, constatou-se a conservação de repetições formadas pelos aminoácidos prolina (P), alanina (A) e lisina (K), agrupados em tetrapeptídeos (PAKK ou KAAK) (Churchill e Travers, 1991; Kasinsky, Lewis *et al.*, 2001; Werlang, Schneider *et al.*, 2009; Sharadamma, Harshavardhana *et al.*, 2010). No entanto, estas sequências aparecem em quantidades diferentes em cada espécie: *M. tuberculosis* e *M. bovis* possuem quatro repetições, *M. marinum* e *M. ulcerans* possuem sete e *M. leprae* apenas uma (Werlang, Schneider *et al.*, 2009; Sharadamma, Harshavardhana *et al.*, 2010). Devido à presença destas repetições, sugeriu-se que a MDP2 possui certa similaridade, na sua porção N-terminal, com a histona de ligação H1 de células eucarióticas (Churchill e Travers, 1991; Kasinsky, Lewis *et al.*, 2001; Werlang, Schneider *et al.*, 2009; Sharadamma, Harshavardhana *et al.*, 2010).

Nas histonas H1, as sequências repetitivas de prolina, serina e outros aminoácidos básicos, são encontradas na porção C-terminal, diferentemente da MDP2 (Churchill e Travers, 1991; Kasinsky, Lewis *et al.*, 2001). Por meio de estudos de mutagênese, verificou-se que este domínio é o responsável pela ligação desta proteína à superfície dos nucleossomos, mediante a interação de alta afinidade das suas sequências repetitivas com a cromatina (Hendzel, Lever *et al.*, 2004; Werlang, Schneider *et al.*, 2009; Sharadamma, Harshavardhana *et al.*, 2010). Além do seu papel na organização estrutural da cromatina, já foi relatado que a histona H1 funciona como um repressor geral da transcrição, tendo sua presença reduzida em genes que estão sendo transcritos (Kamakaka e Thomas, 1990; Bresnick, Bustin *et al.*, 1992); atua no silenciamento epigenético e na diferenciação celular (Jedrusik e Schulze, 2001; Gabrilovich, Cheng *et al.*, 2002); e também modula a expressão de genes específicos (Crane-Robinson, 1999; Folco, Freitag *et al.*, 2003).

Outra proteína de *Mycobacterium tuberculosis* que apresenta semelhanças com o domínio C-terminal da histona H1 é a HLP (ou HupB). Diferente da proteína MDP2, a HLP possui sete repetições de tetrapeptídeos na porção C-terminal, enquanto que a porção N-terminal apresenta certa identidade com as proteínas procarióticas HU. Análises por southern-blot mostraram que a HLP tem a propriedade de ligação ao DNA, assim como fazem as histonas H1, e, portanto, poderia estar envolvida na modulação desta estrutura na micobactéria (Prabhakar, Annapurna *et al.*, 1998).

Uma característica bastante importante da MDP2, também descoberta por meio de EMSA, é a capacidade de se ligar com alta afinidade a estruturas de DNA com quatro fitas chamadas de junções *Holliday* (Stahl, 1994), assim como as proteínas HU de *E. coli*. Esta propriedade pode indicar que a MDP2 está relacionada a eventos de reparação do DNA durante a replicação, ou de recombinação homóloga, como já foi verificado para outras NAPs (Sharadamma, Harshavardhana *et al.*, 2010). Ainda no âmbito da recombinação homóloga, foi demonstrado que a MDP2 é capaz de impedir a troca entre três fitas de DNA pela proteína RecA, podendo, portanto, regular negativamente as vias que levam à recombinação do DNA (Sharadamma, Harshavardhana *et al.*, 2010).

Em relação a sua regulação dentro da célula, já foi demonstrado em dois estudos diferentes que a MDP2 é superregulada no *M. tuberculosis* durante períodos de falta de nutrientes e também em linhagens duplo-mutantes para os genes *hspR* e *hrcA*, os quais codificam repressores de proteínas de choque térmico (Stewart, Wernisch *et al.*, 2002). Por meio de análise por espectrometria de massa associada à ionização/dessorção de matriz assistida por laser (MALDI) e espectrometria de massa acoplada à cromatografia líquida, observou-se que micobactérias submetidas a períodos de restrição de nutrientes superexpressam determinadas proteínas, incluindo a MDP2. Sabe-se que o *M. tuberculosis* é capaz de sobreviver por longos períodos de restrição de nutrientes, permitindo que ele permaneça em estado de latência dentro do hospedeiro. Propôs-se que as proteínas superreguladas detectadas neste estudo possam estar envolvidas na sobrevivência do bacilo, e poderiam, portanto, ser utilizadas como alvos para o desenvolvimento de novos fármacos (Betts, Lukey *et al.*, 2002).

De acordo com a base de dados *online* TubercuList, a proteína MDP2 é codificada a partir de um gene não-essencial (segundo estudos utilizando mutagênese mediada por transposon (Sasseti, Boyd *et al.*, 2003)) denominado *hns*, com 405 pares de bases. Visto que

a anotação do genoma para este gene remete às proteínas H-NS, apesar de predizer também que o gene pode expressar uma proteína HU, foram realizados estudos de complementação utilizando culturas de *Escherichia coli* mutantes para o gene *hns*. Por meio da complementação *in vivo* destas cepas suscetíveis à serina, com um plasmídeo contendo a cópia do gene *hns* de *M. tuberculosis*, observou-se que não houve complementação por parte da proteína MDP2. A partir deste experimento, levantou-se a hipótese de que a MDP2 não corresponde a uma proteína H-NS “clássica”, mas que seria uma nova proteína associada ao nucleóide com funções ainda desconhecidas (Werlang, Schneider *et al.*, 2009).

Existem outros fenótipos associados a mutações no gene *hns* que poderiam ser utilizados para evidenciar os papéis que esta proteína pode desempenhar na célula bacteriana, como: desrepressão do metabolismo do  $\beta$ -glicosídeo, aumento da resistência ao baixo pH e à alta osmolaridade, perda de motilidade, mucoidia e atividade hemolítica (Tendeng e Bertin, 2003). Curiosamente, já foi verificado que outra proteína de *M. tuberculosis*, a Lsr2, foi capaz de restaurar quatro fenótipos selvagens analisados a partir de complementação em cepas de *E. coli* mutantes para o gene *hns*, levando à constatação de que esta é uma nova classe de proteínas semelhante às H-NS (Gordon, Imperial *et al.*, 2008). No entanto, estes testes ainda não foram realizados para a MDP2.

Além de não complementar as células mutantes para o gene *hns*, a proteína MDP2 de *M. tuberculosis* apresenta baixa identidade de sequência com as proteínas H-NS de *E. coli* e *Salmonella* sp. (< 10 % de identidade), mostrando uma aparente ausência de homologia com estas proteínas (Sharadamma, Harshavardhana *et al.*, 2010). De fato, estudos *in silico* e métodos bioquímicos e genéticos apontaram que o *Mycobacterium tuberculosis* não apresenta o gene *hns* no seu genoma (Tendeng e Bertin, 2003). No entanto, apesar da sequência primária da MDP2 não apontar para uma homologia com outras proteínas H-NS, este parâmetro não é suficiente para constatar que esta proteína não apresenta um papel funcionalmente semelhante ao das H-NS (Sharadamma, Harshavardhana *et al.*, 2010). Já foi demonstrado, por exemplo, que as proteínas corismato mutase de *M. tuberculosis* e *E. coli* compartilham em torno de 20 % de identidade na sua sequência de aminoácidos. Contudo, apesar da baixa identidade, ambas desempenham a mesma atividade catalítica e possuem similaridades significativas na sua estrutura terciária (Sharadamma, Harshavardhana *et al.*, 2010). Por outro lado, algumas proteínas podem desempenhar funções similares em dois organismos distin-

tos, apesar de não possuírem identidade de sequência de aminoácidos e de estrutura terciária. Este é o caso, por exemplo, da proteína Lsr2 de *M. tuberculosis* em relação à proteína HNS de *E. coli* (Gordon, Li *et al.*, 2010).

Um estudo recente, que utilizou separação bidimensional baseada em gel e espectrometria de massa como ferramentas para estudar a parede celular do *M. tuberculosis*, mostrou que a proteína MDP2 está presente nesta estrutura. Contudo, os autores não atribuíram nenhuma função para a mesma (Wolfe, Mahaffey *et al.*, 2010). Em 2013, outro trabalho envolvendo a proteína MDP2 de *M. tuberculosis* foi realizado por Ghosh e colaboradores (Ghosh, 2013), no qual os autores se propuseram a compreender o papel funcional da proteína *in vivo*. Os principais resultados apresentados pelos autores indicam que a MDP2 está envolvida no processo de descompactação da estrutura do nucleóide micobacteriano, por meio do ancoramento desta estrutura à membrana plasmática da célula, e que a proteína atua como um regulador negativo da expressão de genes ligados à biossíntese de lipídeos, como o *groEL1*, e genes do operon *fasII*, como o *kasA* e *kasB* (Ghosh, 2013). Estes mesmos autores também realizaram ensaios de complementação *in vivo* utilizando cepas de *E. coli* mutantes para o gene *hns*, e verificaram que, ao se complementar a célula com a MDP2 de *M. tuberculosis*, a proteína não recuperava o estado selvagem da cepa de *E. coli*, corroborando os resultados relatados anteriormente por Werlang e colaboradores (Werlang, 2009).

No mesmo trabalho de Ghosh e colaboradores, anteriormente mencionado, foi observado que a proteína MDP2 recombinante apresentou uma mobilidade anormal em 15 % SDS-PAGE, aparentando ter uma massa molecular maior (~ 21 kDa) que a sua massa teórica de 13,8 kDa (Ghosh, 2013). Apesar dos autores não terem sugerido os motivos pelos quais a proteína migra de forma anômala em gel, existem diversos relatos na literatura de proteínas que apresentam comportamento anormal em SDS-PAGE, como: (1) glicoproteínas, causado pelos seus radicais glicanos hidrofílicos, que reduzem as interações hidrofóbicas entre a proteína e o SDS; (2) proteínas com alto conteúdo de resíduos ácidos, tais como a tropomiosina, devido à repulsão do SDS negativamente carregado pelos resíduos ácidos; (3) proteínas básicas, como as histonas e a troponina I, causado pela redução da razão carga/massa do complexo SDS-polipeptídeo; e (4) proteínas com alto conteúdo de prolina e com uma sequência incomum de aminoácidos nas suas sequências, tais como as proteínas intrinsecamente desordenadas (IDPs) (Tompa, 2002), que causam uma alteração na conformação do complexo

SDS-proteína (Stoyanov, 2001). Visto que este fenômeno também foi observado desde os primeiros experimentos realizados no presente trabalho, cogitou-se a possibilidade de que a MDP2 também poderia apresentar também outras características esperadas para IDPs, as quais serão destacadas a seguir.

### **1.5 Proteínas Intrinsecamente Desordenadas (IDPs)**

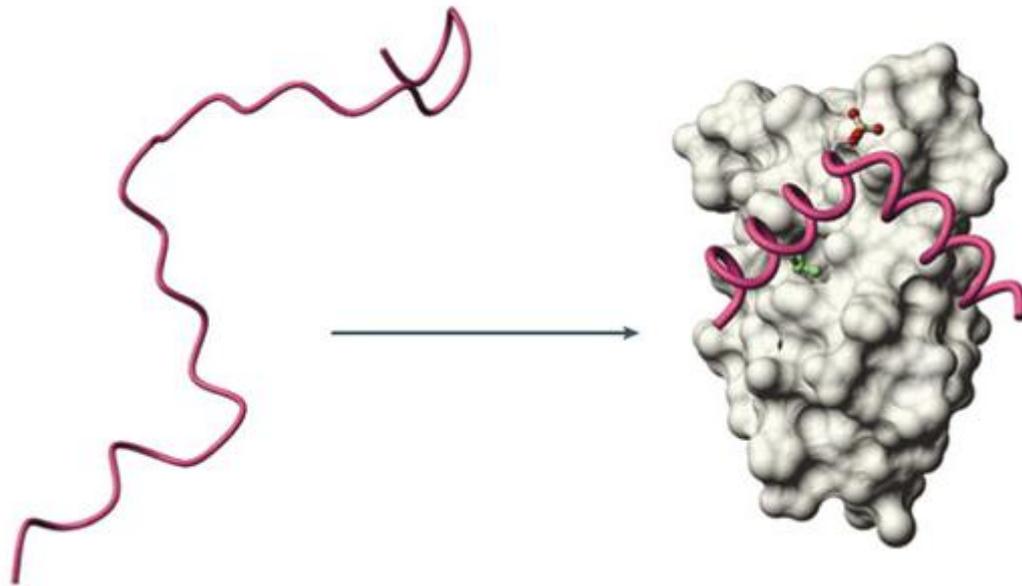
Nos últimos anos, diversos trabalhos têm se focado numa classe bastante peculiar de proteínas conhecidas como IDPs (do inglês, *intrinsic disordered proteins*). Estas proteínas vêm desafiando o paradigma quase dogmático da “estrutura-função”, arraigado nos livros-texto de biologia básica e de bioquímica, o qual propõe que a estrutura é um pré-requisito para a função de uma determinada proteína (Tompa, 2012). Diferentemente da visão comum de que as proteínas têm uma única estrutura 3D estável com posições relativamente fixas dos seus átomos e ângulos de Ramachandran, as IDPs possuem fragmentos altamente dinâmicos na sua estrutura em condições fisiológicas que não permitem a obtenção de um único conjunto de coordenadas e ângulos de Ramachandran para descrever as suas propriedades conformacionais (Dunker, 2013).

A alta flexibilidade encontrada em algumas porções das IDPs deve-se, principalmente, da sequência incomum de aminoácidos destas proteínas. Sabe-se que resíduos de aminoácidos carregados e de aminoácidos polares são bastante prevalentes nas IDPs, como a alanina, arginina, glicina, ácido glutâmico, metionina, serina, glutamina, lisina e prolina, os quais são denominados resíduos promotores de desordem. Por outro lado, as IDPs são carentes em aminoácidos apolares e hidrofóbicos, como a cisteína, tirosina, triptofano, fenilalanina, leucina, valina, histidina, treonina e asparagina, conhecidos como resíduos promotores de ordem (Forman-Kay, 2013). A composição de aminoácidos das IDPs também confere a estas proteínas uma alta carga líquida positiva ou negativa, permitindo que estas realizem principalmente interações eletrostáticas com seus parceiros de interação (Forman-Kay, 2013). Dentre os aminoácidos promotores de desordem, a prolina é considerada o resíduo que mais causa desordem na estrutura de uma IDP, apesar do seu caráter não carregado; estima-se que a prevalência deste resíduo em IDPs seja 1,4 maior do que nas proteínas enoveladas (Theillet, 2013). Este alto conteúdo de prolinas está intimamente relacionado com a falta de estrutura nas IDPs, visto que estes resíduos desfavorecem uma estrutura secundária rígida,

mas favorecem um motivo conformacional aberto conhecido como poliprolina II (PP II) (Tompa, 2002).

As IDPs também são conhecidas pela sua capacidade de sofrer um enovelamento induzido (uma transição da desordem para a ordem) por meio de uma interação com um parceiro de interação [Figura 11] (Tompa, 2012, Dyson, 2005). Esta alta flexibilidade permite às IDPs interagirem com múltiplos parceiros e participarem de diversos processos metabólicos dentro da célula, sendo frequentemente usadas como *hubs* (ou nós centrais) de redes de interação de proteínas, motivo pelo qual elas também serem conhecidas como proteínas promíscuas (Uversky, 2011). Visto que os *hubs* são essenciais para o funcionamento correto de redes de interação proteína-proteína, a sua perda, muitas vezes, é letal para o organismo. De fato, muitas IDPs consideradas *hubs* estão envolvidas em processos de patogênese de doenças humanas, como a p53,  $\alpha$ -sinucleína, BRCA1, XPA, p21, p27, receptor de estrogênio, os quais podem interagir com centenas de parceiros dentro da célula (Uversky, 2011). Uma vez que as IDPs estão intimamente relacionadas com algumas doenças, estas proteínas vêm sendo consideradas alvos bastante promissores para o desenvolvimento de fármacos (Tompa, 2012). No entanto, visto que elas não possuem atividade enzimática, sua inibição não pode ser alcançada da forma tradicional que os fármacos atuam, por meio da interação de um ligante ao sítio ativo da enzima. Apesar disso, já foi sugerido que pequenas moléculas interferentes possam ter como alvo IDPs, como verificado para a oncoproteína c-Myc e para a proteína precursora amiloide (APP) na doença de Alzheimer, ligando-se a motivos especiais encontrados nestas proteínas, como os SLiMs (motivos lineares pequenos), os PSEs (elementos estruturais pré-formados) e os MoRFs (funções de reconhecimento molecular) (Tompa, 2012).

**Figura 11.** Enovelamento do domínio fosforilado induzível por quinase (pKID) induzido pela ligação ao domínio de ligação KID (KIX) da proteína de ligação a CREB (CBP).



**Fonte:** Dyson, 2005.

**Nota:** na representação à esquerda é possível observar que o domínio pKID do fator transcricional CREB é desordenado quando está livre em solução. Por outro lado, a representação à direita mostra que este mesmo domínio sofre um processo de enovelamento quando ligado ao domínio KIX da proteína CBP.

As IDPs são encontradas nos três grandes domínios da vida - Eukarya, Bacteria e Archaea -, porém, a sua prevalência em cada grupo varia consideravelmente. Estudos envolvendo a identificação de IDPs por meio de predição *in silico* mostraram que há uma discrepância entre eucariotos e procariotos: 18,9 % das proteínas eucarióticas são preditas por conter regiões de desordem, enquanto somente 5,7 e 3,8 % das proteínas de bactérias e arqueas, respectivamente, possuem estas sequências (Ward, 2004). Além disso, foi observado que o *M. tuberculosis* se destaca entre as eubactérias, possuindo uma alta frequência de regiões de desordem; esse fenômeno pode ser um resultado da alta prevalência dos aminoácidos alanina, glicina, arginina e prolina (Ward, 2004). Visto que algumas IDPs são apontadas como proteínas-chave para o metabolismo de certos organismos, e que o *M. tuberculosis* possui uma alta prevalência destas proteínas, já foi sugerido que algumas IDPs deste patógeno possam servir de alvos para o desenvolvimento de fármacos anti-TB, entre elas a FtsW (Rv2154c), membro da família das SEDS (forma, alongamento, divisão e esporulação); a GlmU (Rv1018c), uma proteína bifuncional que contém um domínio uridil-transferase e outro acetil-transferase; e a Obg (Rv2440c), a qual pertence à família das proteínas de ligação ao GTP (Anurag, 2009). Por fim, como sugerido pelos autores deste estudo, mais investiga-

ções devem ser realizadas a fim de se identificar novas IDPs no *M. tuberculosis* e de se entender seus respectivos papéis no metabolismo da micobactéria.

## 2 JUSTIFICATIVA

Estudos anteriores demonstraram que a MDP2 é capaz de interagir com diferentes moléculas de DNA (Werlang, Schneider *et al.*, 2009; Sharadamma, Harshavardhana *et al.*, 2010), entretanto, pouco se sabe sobre as suas características estruturais e o seu papel na biologia da micobactéria. A geração de uma linhagem nocaute para esta proteína em *M. tuberculosis* poderá auxiliar na elucidação da função desta proteína dentro do bacilo, bem como confirmar a não-essencialidade do gene *hns* constatada pela mutagênese mediada por transposon (Sasseti, Boyd *et al.*, 2003), uma vez que esta metodologia serve principalmente como uma ferramenta de rastreamento, limitada a uma condição de cultivo. A linhagem nocaute poderá ser utilizada também em um modelo de latência *in vitro* (Deb, Lee *et al.*, 2009), visto que já foi demonstrado que a MDP2 é super-expressa durante períodos de restrição de nutrientes (Betts, Lukey *et al.*, 2002). Por meio deste experimento, será possível verificar a importância desta proteína para a sobrevivência do bacilo durante o estado de latência. Além disso, caso esta proteína participe na modulação da expressão de diferentes genes, como fazem as NAPs, será possível constatar diferenças na expressão gênica entre a linhagem selvagem e a linhagem nocaute para a MDP2.

Aliado aos estudos de nocaute gênico, a validação da MDP2 de *M. tuberculosis* como uma proteína intrinsecamente desordenada poderá evidenciar sua participação nas redes de interações de proteínas dentro da micobactéria. Apesar de ser predita como não-essencial para a sobrevivência do bacilo, a MDP2 pode atuar como uma proteína-chave para o funcionamento de diversos processos celulares, uma vez que já foi verificado que esta proteína atua tanto na descompactação do nucleóide micobacteriano quanto na modulação de diferentes genes relacionados ao metabolismo de lipídeos. Futuramente, estes estudos poderão conduzir ao desenvolvimento de uma cepa atenuada de *M. tuberculosis*, caso a MDP2 se mostre importante para o processo de infecção do bacilo dentro dos macrófagos e para o desenvolvimento da doença dentro do hospedeiro.

Além disso, apesar da anotação do genoma do *M. tuberculosis* (H37Rv) indicar a MDP2 como sendo uma proteína pertencente ao grupo das H-NS, algumas evidências, como a baixa identidade de sequências e a incapacidade de complementar células nocautes para o gene *hns* em *E. coli* (Werlang, Schneider *et al.*, 2009, Ghosh, 2013), apontam que esta proteína não representa, de fato, uma H-NS. Contudo, até o momento, nenhum estudo comparou a

estrutura terciária destas duas proteínas, a fim de se identificar uma possível conservação de estrutura. Como dito anteriormente, algumas proteínas apresentam baixa identidade na sua sequência primária, mas compartilham similaridades na sua estrutura terciária, fazendo com que ambas desempenhem funções similares (Sharadamma, Harshavardhana *et al.*, 2010). Diante destas informações, torna-se importante elucidar a estrutura tridimensional da MDP2, para que seja possível estabelecer relações estruturais e funcionais desta proteína com as NAPs já caracterizadas.

### 3 OBJETIVOS

#### 3.1 Objetivo Geral

Validar a proteína MDP2 de *Mycobacterium tuberculosis* H37Rv como uma proteína intrinsecamente desordenada e obter uma cepa mutante para o gene *hns*.

#### 3.2 Objetivos Específicos

- › Determinar a essencialidade do gene *hns* de *M. tuberculosis* *in vivo* por meio de no-  
caute gênico.
- › Verificar *in silico* o grau de desordem da MDP2 de *M. tuberculosis* por meio de predi-  
tor de desordem.
- › Verificar o comportamento da MDP2 recombinante diante de técnicas indiretas utili-  
zadas para a validação de proteínas intrinsecamente desordenadas.
- › Obter a massa intacta e o mapeamento peptídico da MDP2 homogênea por espec-  
trometria de massa.
- › Determinar a estrutura tridimensional da proteína MDP2 recombinante de *M. tuber-  
culosis*.

## CAPÍTULO 2

### ARTIGO CIENTÍFICO

The Mycobacterial DNA-binding Protein 2 (MDP2) from *Mycobacterium tuberculosis* H37Rv is an intrinsically disordered protein.

Bruno L. Abbadi, Zilpa Adriana Sánchez Quitian, Paula Eichler, Paulo C. Patta, Luiz A. Basso, Diógenes S. Santos, Cristiano V. Bizarro.

Artigo submetido em 28 de fevereiro de 2014 ao periódico *Journal of Biological Chemistry* (JBC).

The Mycobacterial DNA-binding Protein 2 (MDP2) from *Mycobacterium tuberculosis* H37Rv is an intrinsically disordered protein\*

**Bruno L. Abbad**<sup>1,2</sup>, **Zilpa A. Sánchez Quitian**<sup>1,2</sup>, **Paula Eichler**<sup>1</sup>, **Paulo C. Patta**<sup>1,2</sup>, **Luiz A. Basso**<sup>1,2</sup>, **Diógenes S. Santos**<sup>1,2</sup>, **Cristiano V. Bizarro**<sup>1,2</sup>

<sup>1</sup>Instituto Nacional de Ciência e Tecnologia em Tuberculose (INCT-TB), Centro de Pesquisas em Biologia Molecular e Funcional (CPBMF); <sup>2</sup>Programa de Pós-Graduação em Biologia Celular e Molecular, Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Porto Alegre, RS, Brazil

Running title: *MDP2 of M. tuberculosis is an IDP*

To whom correspondence should be addressed: Cristiano V. Bizarro, Centro de Pesquisas em Biologia Molecular e Funcional (CPBMF), Instituto Nacional de Ciência e Tecnologia em Tuberculose, Pontifícia Universidade Católica do Rio Grande do Sul, Av. Ipiranga, 6681/prédio 92A, TECNOPUC, Porto Alegre, RS, Brazil, tel./fax: (51) 3320-3629; e-mail: cristiano.bizarro@puers.br.

**Keywords:** Anomalous Migration; Crosslinking; EMSA; Intrinsically Disordered Proteins; Gel Filtration; Heat Resistance; Limited Proteolysis; *Mycobacterium tuberculosis*; Mass Spectrometry; Protein Purification

**Background:** MDP2 is involved in mycobacterial nucleoid organization and anchors genomic DNA to the plasma membrane.

**Results:** MDP2 migrates anomalously in SDS-PAGE, is resistant to aggregation induced by heat or acid, is sensitive to proteolysis and self-associates.

**Conclusion:** MDP2 is an intrinsically disordered protein.

**Significance:** Protein disorder may be central to nucleoid organization activity of DNA-binding proteins.

## ABSTRACT

**The mycobacterial DNA-binding protein 2 (MDP2) of *Mycobacterium tuberculosis* is a small, basic protein known to bind DNA in a non-specific manner and to anchor this molecule to the plasma membrane. Motivated by a structural disorder prediction performed *in silico*, we have used a complementary set of techniques to characterize the structure of this protein, such as heat and chemical stability, gel filtration, limited proteolysis and electrophoretic mobility. Here, we show that MDP2 is an intrinsically disordered protein (IDP) since it (1) was predicted *in silico* to be an IDP; (2) resisted to denaturation induced by boiling temperature and to the chemical trichloroacetic acid (TCA); (3) migrated anomalously on SDS-PAGE, appearing to be 1.4-fold higher than expected; and (4) was**

**highly sensitive to proteolysis by proteinase K. Combined with previous results related to MDP2 from other works, we speculate that this protein uses its highly disorder N-terminal region to bind to DNA through its repeating KAAK and PAKK sequences in a non-specific manner.**

Intrinsically disordered proteins (IDPs) constitute a novel class of proteins well-known to have challenged the dogmatic “structure-function paradigm” still found in basic biology and biochemistry textbooks (1). These proteins are mainly recognized by their highly dynamic fragments, which come from their unusual amino acid composition (2). Some of them lack completely ordered structures under physiological conditions, with no globular structures as expected, while others have some disordered fragments along its sequence (2,3). In fact, IDPs have multiple faces, since they can display different degrees of intrinsic disorder: they may be more or less compact, have larger or smaller amounts of flexible secondary structures and keep smaller or larger numbers of tertiary structures (4).

Many terms have been used to nominate these proteins in the last decade (*e.g.* flexible, mobile, partially folded, natively denatured, intrinsically unstructured, natively unfolded), which could cause some misunderstandings between researchers; however, a recent study supported

the term IDPs as a common description of those biologically active proteins lacking one 3D-structure, in an attempt to universalize the term (2). Nevertheless, an opposite viewpoint supported by Janin and Sternberg suggested that most IDPs - if not all - are in fact “proteins waiting for partners (PWPs), arguing that disordered is just an artifact of current methods of protein expression, since they may not fold correctly in the absence of their interaction partner (5).

The intrinsic disorder found in IDPs essentially comes from their uncommon amino acid composition, characterized by an enrichment of charged and polar residues (alanine, arginine, glycine, glutamic acid, methionine, serine, glutamine, lysine and proline), also recognized as disorder-promoting residues, and a scarcity of non-polar and hydrophobic residues (cysteine, tryptophan, tyrosine, phenylalanine, leucine, valine, histidine, threonine, asparagine), known as ordered-promoting residues (6-8). Proline is recognized as the most disorder-promoting residue, and it is 1.4-times more frequent in IDPs than on surfaces of folded proteins (8). This distinct amino acid profile has been widely exploited by different computational algorithms used by over 50 protein disorder predictors nowadays, in order to determine *in silico* if a protein might be intrinsically disordered (1,7).

*In silico* prediction has also been used to investigate the presence of disorder in several archaea, eubacteria and eukaryote genomes (3). These studies have shown that IDPs are more common among eukaryotes - about 19.0 % of their proteins are predicted to be disordered - in comparison to eubacteria and archaea - 5.7 and 3.8 % respectively (3,9). Surprisingly, *Mycobacterium tuberculosis*, the deadly pathogen that causes human tuberculosis (10), stands out among the eubacteria, having an atypical high disorder frequency in its genome (3). Moved by this data, a study conducted by Anurag and Dash has highlighted the potential of IDPs from *M. tuberculosis* as candidates for drug targets, since these proteins may alternate between disorder-to-order states, which are considered a new dimension to drug target development (11).

Many methods have been used to validate a protein of interest as an IDP *in vitro*. These methods range from indirect tools, which allow the visualization of the disordered behavior of the protein under some conditions, such as heat (12-18) and chemical-stability (16,18), sodium dodecyl sulfate-polyacrylamide gel electrophore-

sis (SDS-PAGE) (19-22), and proteolytic sensitivity (17,19,21,23-25); hydrodynamic techniques, such as size-exclusion chromatography (22,23,25-27), dynamic light scattering (DLS) (23), and small-angle X-ray scattering (SAXS) (23); up to more direct ways of identifying the lack of 3-D structure, including X-ray crystallography, far-UV circular dichroism (CD) (17,19,23,25-27), and NMR spectroscopy (17,23,28,29, 4). Here, we have used the disorder predictor PONDR® and a complementary set of direct and indirect experimental techniques in order to characterize the mycobacterial DNA-binding protein 2 (MDP2) of *M. tuberculosis* as an IDP.

MDP2 is a small, basic protein known to bind to DNA in a nonspecific manner, with a slight preference for curved sequences (rich in A-T nucleotides) (30,31). This protein was first classified as a histone-like nucleoid-structuring (H-NS) protein from *M. tuberculosis* (Rv3852) by the TubercuList database (32), despite the low sequence identity between the H-NS of *E. coli* and the MDP2 of *M. tuberculosis* (< 10%) (30,31). Moreover, MDP2 was incapable to complement an *E. coli* strain lacking the H-NS function (30,33); oddly enough, another protein of *M. tuberculosis* called Lsr2 performs a functional homologous role to H-NS protein of *E. coli*, even though they do not share a similar sequence or tertiary structure (34). Ghosh and colleagues also reported that the MDP2 may play an important role in the nucleoid organization, by anchoring this structure to the plasma membrane, and that it may be involved in the lipid metabolism, by modulating the expression of different genes (33).

Making use of the disorder predictor PONDR® we first verified that MDP2 had almost 86 % of disordered residues in its sequence, which is a strong evidence that the protein is an IDP. Motivated by this result, we next addressed some indirect tools to validate *in vitro* this protein as an IDP. The results showed that MDP2 (1) resists to denaturant temperatures; (2) resists to chemical denaturation induced by trichloroacetic acid (TCA); (3) has an anomalous migration on SDS-PAGE, with an apparent molecular weight 1.4-fold higher than its real value; and (4) is highly sensitive to proteolysis by proteinase K. Together, our results strongly suggest that MDP2 belongs to the intriguing class of intrinsically disordered proteins.

## EXPERIMENTAL PROCEDURES

*In silico Intrinsic Disordered Prediction* - *In silico* intrinsic disorder prediction of MDP2 was performed with the entire amino acid sequence taken from NCBI (MDP2, NP\_218369). The disorder predictor used was PONDR® VL-XT, which considers a residue as disordered if its value exceeds or matches a threshold of 0.5 (35-37). The mean net charge  $\langle R \rangle$  and the mean hydrophobicity  $\langle H \rangle$  were also calculated using the PONDR® server. The border between intrinsically disordered and ordered proteins is described by the formula (38):

$$\langle R \rangle = 2.785 \langle H \rangle - 1.151 \quad (\text{Eq. 1})$$

A cumulative distribution function (CDF) curve was created by plotting predicted disorder scores against their cumulative frequency as described elsewhere (39). PONDR® predictor is available at <http://www.pondr.com/index.html>. Access to PONDR® was provided by Molecular Kinetics, Inc. (6201 La Pas Trail Suite 160, Indianapolis, IN 46268; phone: 317-280-8737; e-mail: [main@molecularkinetics.com](mailto:main@molecularkinetics.com)). The amino acid composition analysis was performed using the ProtParam tool of the ExpPASy portal available at <http://web.expasy.org/protparam>.

*Gene Amplification and Cloning* - Gene amplification was performed as described previously (30). Briefly, the *hns* gene (Rv3852) was PCR amplified on thermal cycler (TC-412 - Analítica) from total genomic DNA of *M. tuberculosis* H37Rv strain (40) using two synthetic oligonucleotide primers containing *Nde*I (sense primer: 5' AACATATGCCAGACCCGCAGGA TCGACCC3') and *Hind*III (antisense primer: 5'GTAA GCTTTCAGCGCGGCGCAGTTGCC3') restriction sites (underlined). Dimethyl sulfoxide (DMSO) was added to the reaction at a final concentration of 10 %. The PCR amplicon was first cloned into the pCR™-Blunt vector (Invitrogen) and then subcloned into the pET-23a(+) expression vector (Novagen), using *E. coli* DH10B electrocompetent cells. The integrity of the constructs was confirmed in all cases by digestion reactions with the respective restriction enzymes (New England Biolabs) and automated DNA sequencing (ABI-PRISM 3100 Genetic Analyzer - Applied Biosystems).

*Heterologous Expression of the MDP2* - *M. tuberculosis* MDP2 was expressed using *E. coli*

C41 (DE3) as host cells. Cells were electroporated and transformant colonies containing the recombinant plasmid pET-23a(+):*hns* were selected on Lysogeny Broth (LB) agar plates supplemented with 50  $\mu\text{g mL}^{-1}$  of ampicillin. Single colonies were picked up and grown overnight (at 37°C and 180 rpm) in 50 mL of LB medium supplemented with 50  $\mu\text{g mL}^{-1}$  of ampicillin. Aliquots of 10 mL from the grown culture were then inoculated into 500 mL of LB medium also containing 50  $\mu\text{g mL}^{-1}$  of ampicillin. These cultures were grown at 37 °C by shaking at 180 rpm up to an optic density ( $\text{OD}_{600 \text{ nm}}$ ) of 0.4-0.6. At this growth stage, we induced MDP2 expression by adding isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM. After 6 hours of growth, cells were harvested by centrifugation (11,800 g – himac CR 21G high-speed refrigerated centrifuge) for 30 min at 4 °C and then stored at -20 °C.

*MDP2 Purification* - Protein purification was performed by Fast Performance Liquid Chromatography (FPLC) on Äkta Purifier System (GE HealthCare) at 4 °C. Frozen cells (6 g wet weight) were suspended in 30 mL of buffer A (Tris HCl 50 mM NaCl 0.3 M pH 7.9) and disrupted by sonication with the ultrasonic processor Vibra-Cell VCX (14 pulses of 10 s at 60 % of amplitude, with intervals of 1 min on ice). Supernatant was clarified by centrifugation (38,900 g for 30 min in all cases). In order to remove all nucleic acids from the solution and to perform a first purification step of MDP2, a titration with polyethyleneimine (PEI) (Sigma-Aldrich) was carried out as described previously (41). For this step, 5 % (w/v) PEI stock solution was added to aliquots of 1 ml of the supernatant to give final PEI concentrations of 0 to 0.5 % (w/v) in 0.05 % (w/v) intervals. Tubes were rapidly mixed, incubated for 5 min at 0 °C, mixed again and centrifuged for 5 min at 4 °C. Supernatants were assayed for the absence of the MDP2 by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis. After picking out the final PEI concentration that precipitates both nucleic acids and MDP2 (0.15 % (w/v)), the NaCl concentration required to elute MDP2 from the pellet was determined. Identical pellets were resuspended for 10 min in 1.5 mL of Tris HCl 50 mM pH 7.9 containing varied concentrations of NaCl (from 0 to 1 M) and then centrifuged. The lowest concentration of NaCl required to release MDP2 from the pellet into the

supernatant was found to be 0.6 M by SDS-PAGE analysis (not shown). After eluting the protein from pellet with buffer B (Tris HCl 50 mM NaCl 0.6 M pH 7.9), free PEI was removed by adding 60 % of saturated ammonium sulfate to the supernatant, stirring for 30 min and centrifuging. The resulting pellet was resuspended with buffer C (Hepes 50 mM NaCl 100 mM pH 7.9) and centrifuged. The supernatant was then loaded on the SP Sepharose Fast Flow cation exchange column (GE Healthcare) pre-equilibrated with buffer C. The adsorbed material was eluted by a linear gradient (0 to 100 %) of Hepes 50 mM NaCl 1 M pH 7.9 (buffer D) at a 1 mL min<sup>-1</sup> flow rate. Fractions containing MDP2, confirmed by 15 % SDS-PAGE, were pooled, concentrated and loaded on the Sephacryl S-100 size exclusion column (GE Healthcare) previously equilibrated with buffer C; proteins were then eluted in isocratic conditions. Lastly, fractions containing MDP2 were loaded on the Mono-S cation exchange column (GE Healthcare) pre-equilibrated with buffer C. Proteins were eluted by a linear gradient (0 to 100 %) of buffer D. Fractions containing homogeneous MDP2 were subsequently pooled and concentrated (Centricon 10,000 MWCO – Millipore); through the concentration step, MDP2's buffer was changed to buffer C. Protein concentration was assessed by the Qubit® 2.0 Fluorometer (Invitrogen) according to the manufacturer's protocol.

*Electrophoretic Mobility Shift Assay (EMSA)* – To confirm the DNA-binding ability of MDP2, a pair of complementary oligonucleotides were annealed (5'-CAACTTCCCTAACCCAGCTGCGATCCGTT-3' and 5'-AACGGATCGCAGCTGGGTTAGGGAAGTTGG-3'), resulting in a double-stranded DNA (dsDNA) with 30 bp as reported before (33). The dsDNA (300 ng) was then incubated with increasing concentrations of protein (0-2 µg) at room temperature for 15 minutes in the binding buffer (10 mM Hepes, pH 7.5, 80 mM NaCl, 1 mM DTT, 1 mM EDTA, 5 % glycerol (v/v)) in a 25 µL reaction volume. DNA-protein complexes were electrophoresed on 1.5 % agarose gel (w/v) in TAE buffer (40 mM Tris-acetate, 1 mM EDTA, pH 8.0) for 1 h at 50 v. The gel was stained with a 3x solution of GelRed™ (Biotium) and the bands were visualized by the ChemiDoc™ MP (Bio-Rad).

*Behavior of MDP2 in SDS-PAGE* - The electrophoretic behavior of MDP2 was analyzed by SDS-PAGE under denaturant conditions. Laemmli gels were prepared containing a 4% stacking gel (pH 6.8) and a 15% running gel (pH 8.8) using a solution of 1:29 of N,N'-methylenebisacrylamide to acrylamide. The polymerization process was carried out by the addition of ammonium persulfate (APS – Sigma-Aldrich) and N,N,N',N'-tetramethylethylenediamine (TEMED - Sigma Aldrich). After heating the protein at 100 °C for 10 min in the presence of the loading buffer (62.5 mM Tris-HCl pH 6.8, 2.5 % SDS, 0.002 % bromophenol blue, 0.7135 M (5 %) β-mercaptoethanol, 10 % glycerol), the sample was loaded onto the gel. Unstained Protein Molecular Weight Marker (Thermo Scientific) was used as the protein standard. Vertical electrophoresis was carried out at 60 v for the stacking step, and subsequently at 160 v for almost an hour using the Power PAC 3000 (Bio-Rad). The gel was stained with Coomassie Brilliant Blue R-250 (Sigma-Aldrich).

*MDP2 Identification by Mass Spectrometry* - Identification of MDP2 was accomplished by LC-MS/MS peptide mapping experiments. In gel protein digestion with trypsin (Sigma) was performed with aid of 0.1 % (w/v) Rapigest SF surfactant (Waters Corp.) as recommended by its manufacturer. The band expected to correspond to MDP2 was excised from the gel and chopped into a tube where the whole trypsinization procedure went on. In solution digestion was performed after protein purification and followed a protocol adapted from Klammer and MacCoss (42). LC-MS/MS experiments were performed in a nanoLC Ultra system (nanoLC Ultra 1D plus, Eksigent) connected to a hybrid mass spectrometer LTQ-XL/LTQ Orbitrap Discovery with a Finnigan™ nanospray ionization source (all by Thermo Electron Corporation). Tryptic peptides were desalted on a 150 µm i.d., 5 cm length pre-column and separated on a 150 µm i.d., 15 cm length analytical column, both packed in-house with Kinetex 2.6 µm C18 core-shell particles (Phenomenex, Inc.) by a slurry packing procedure (43). A 150 min gradient of ACN (mobile phase B) was used to elute peptides at a flow rate of 200 nL min<sup>-1</sup> (2 % B for 10 min, 2–98 % B in 60 min, 98 % B for 15 min, 98-2% in 10 min, and 2 % for 55 min); mobile phase A consisted of 0.1% formic acid in water. Spectra of eluting

peptides were acquired in positive ion mode in a data-dependent fashion. MS survey scans at 400–2000  $m/z$  range with resolution of 30,000 (at  $m/z$  400) were followed by MS/MS spectra of the five most intense ions from each survey scan. MS/MS fragmentation was performed using collision-induced dissociation with activation  $Q$  of 0.250, activation time of 30.0 ms, 35% of normalized collision energy, and 1.0 Da isolation width. Dynamic exclusion of 30 s was applied. LC-MS/MS data were compared with theoretical MS/MS spectra obtained from *in silico* tryptic digests of the MTB H37Rv proteome (<ftp://ftp.ncbi.nih.gov/genomes>) with software Thermo Proteome Discoverer 1.3 and SEQUEST algorithm (Thermo Fisher Scientific Inc). Special parameters set were: two missed trypsin cleavage sites, precursor tolerance of 10 ppm, fragment tolerance of 0.8 Da, static carbamidomethylation on cysteines, and dynamic oxidation on methionine residues. Analyses were restricted to matches with a Xcorr score > 2.0 for doubly charged ions and Xcorr score > 2.5 for triply charged ions, and decoy database was used. For MDP2 molecular mass determination, purified samples were desalted with a reverse chromatography phase resin (POROS R2-50, Applied Biosystems) packed in 200  $\mu$ l pipette tips with glass fiber. Eluted samples were dried by SpeedVac concentrator (Thermo Scientific) and reconstituted in ACN 50 %/FA 0.1 % and directly injected in a static mode into an IonMax electrospray ion source (Thermo Electron Corporation). Electrospray parameters were set as follows: positive mode; 4.5 kV at source; 5 arbitrary units (0–100 range) of sheath gas, and 6 of auxiliary gas flows; 250 °C and 48 V at the capillary ion transfer tube, and 170 V at the tube lens. Full spectra (650 – 2000  $m/z$  range) were acquired during 23 min in profile mode at a nominal resolution  $r = 30,000$  at  $m/z$  400 using FT automatic gain control target value of 1,000,000 charges. Spectra average of 1000 scans was processed with Xtract for charge state deconvolution and isotopic molecular mass determination.

*MDP2 Quaternary Structure Assessment by Gel Filtration and In vitro Crosslinking* - Analytical FPLC gel filtration chromatography was performed using LMW and HMW Gel Filtration Calibration Kits combined (GE Healthcare) with a 1.0 cm x 30 cm High Resolution Superdex 200 column (GE Healthcare). Ribonuclease A (13.7 kDa), chymotrypsinogen A (25.6 kDa), co-

albumin (44 kDa), aldolase (158 kDa), catalase (232 kDa), and ferritin (440 kDa) were used as standard proteins. Blue Dextran 2000 was used to determine the void volume ( $V_0$ ) of the column. 360  $\mu$ g of homogeneous MDP2 was loaded into the column previously equilibrated with Hepes 50 mM NaCl 0.2 M pH 7.9 at a flow rate of 0.4 mL  $\text{min}^{-1}$ . Protein elution was monitored at 215, 254 and 280 nm. The partition coefficient ( $K_{av}$ ) was calculated as:

$$K_{av} = (V_e - V_0)/(V_c - V_0) \quad (\text{Eq. 2})$$

where  $V_e$  is the sample elution volume, and  $V_c$  is the geometric column volume.  $K_{av}$  values were plotted against the logarithm of the molecular weights of the samples to generate a calibration curve.

Crosslinking tests were performed as described previously (44), using a VDX plate (Hampton Research) with 24 wells. The bottom of the wells were filled with 40  $\mu$ L of 25 % (v/v) glutaraldehyde (Sigma-Aldrich) acidified with 1  $\mu$ L of 5 N HCl. A 15  $\mu$ L drop of protein (0.5 mg  $\text{mL}^{-1}$  in buffer C) was placed onto the coverslip, which was then used to seal the well, allowing the protein solution to form a hanging drop. The support was placed in a water bath at 30 °C, and samples were collected at different time intervals to be subsequently analyzed by 15 % SDS-PAGE.

*Heat Treatment* - The resistance to heat-induced aggregation was analyzed by 15 % SDS-PAGE. An aliquot of 0.25 mg  $\text{mL}^{-1}$  of MDP2 was heated at 100 °C for 10 min and quickly placed on ice for 5 minutes. Heat-induced aggregates were pelleted by 15 min centrifugation at 16,200 g and 4 °C, while heat-resistant proteins remained in the supernatant and were loaded on the gel. Non-treated control samples were kept at room temperature during the heat treatment step. The following globular proteins were used as positive controls for the method: cytidine deaminase (CDA) and 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) from *M. tuberculosis* (45,46), guanosine monophosphate reductase (GMPR) from *E. coli* (47), and human purine nucleoside phosphorylase (PNP) (48). All incubations were performed in buffer C.

*Chemical denaturation by TCA* - Chemical treatment with trichloroacetic acid (TCA) (Sigma-Aldrich) was performed as described previously (16). An aliquot of 0.2 mg  $\text{mL}^{-1}$  of MDP2

was subjected to acid treatment by the addition of TCA to the final concentration of 1, 3 and 5 %. After incubation on ice for 15 min, samples were centrifuged at 16,200 g for 15 min at 4 °C. The above-mentioned globular proteins were also used as positive controls for the chemical treatment experiment. In order to detect proteins resistant to chemical denaturation, the supernatant obtained after the previous centrifugation step was analyzed by 15% SDS-PAGE.

*Protease Sensitivity Assay* – To verify MDP2 sensitiveness to proteolysis, a protease sensitivity assay was performed with proteinase K (broad specificity) (Gibco®) as described by Kovacs and colleagues (17). 24 µM samples of MDP2 were incubated with proteinase K at ratios of 1:10, 1:250, 1:600, 1:900, 1:1,500, 1:2,000 and 1:10,000 (proteinase K:MDP2) in separate reactions for 5 min at 25 °C. The reactions were rapidly stopped by adding protein loading buffer and then heated at 100 °C for 10 min. Supernatants were analyzed by 15 % SDS-PAGE. CDA and bovine serum albumin (BSA) (Sigma-Aldrich) were used as controls. Quantitative analysis was performed by the ChemiDoc™ MP (Bio-Rad).

## RESULTS

*MDP2 is predicted to be highly disordered* - Based on PONDR® VL-XT prediction, MDP2 has 115 disorder-promoting residues in its primary structure, corresponding to almost 86 % of its entire structure (134 residues). These residues are distributed in three main disordered regions: an IDR of 76 residues (M1-L76; 95% of disorder), which is the longest disordered segment covering from the N- terminal end until the central part of the protein; and two IDRs covering the C-terminal end, one with 31 residues (A87-A117; 85 % of disorder) and another with 8 residues (L127-R134; 85 % of disorder) (Figure 1A). Since the scores of the majority of the residues were above 0.5, which is the threshold distinguishing ordered from disordered regions in a protein, this prediction strongly suggest that MDP2 is mostly disordered.

The amino acid compositional analysis performed by ProtParam is shown in Table 1. A compositional bias was observed in each of the three disordered domains. While the N-terminal disorder region is enriched with Ala (19 %), Lys (18 %) and Pro (20 %), the second IDR is com-

posed mainly by Ala (22 %), Pro (13 %), Ser (13 %) and Val (13 %), and the C-terminal part is enriched with Arg (50 %). Most of these amino acids are well recognized to be disorder-promoting (Ala, Pro, Lys, Ser, Arg) in contrast to order-promoting residues (Cys, Trp, Tyr, Ile, Phe, Leu, His, Thr, Asn) (6,8,19), which contribute to the highly disordered character predicted for MDP2.

PONDR® also provides another two binary prediction tools used to discriminate ordered from disordered proteins: the cumulative distribution function (CDF) and the charge-hydrophathy plot (C-H). As shown in Figure 1B, MDP2's CDF curve resides on the lower part of the graph, a characteristic of fully disordered proteins (49). Based on C-H plot result, however, it is not so clear if the protein is indeed fully disordered, since its boundary hydrophathy value is upon the borderline between ordered and disordered proteins (Figure 1C); yet, it is possible to notice that the mean net charge of MDP2 (0.1045) is higher than the majority of ordered proteins, another sign of disordering (49). This same behavior on the C-H plot was previously observed for the intrinsically disordered N-terminal region of AtREM1.3 protein (19).

*Recombinant MDP2 migrates anomalously in SDS-PAGE* - The identity of the pET-23a(+):*hns* sequence was confirmed by automated DNA sequencing, and no change was found in the nucleotide sequence (not shown). The recombinant MDP2 expressed in *E. coli* C41 (DE3) cells was purified to homogeneity by a three-step chromatographic protocol, and its homogeneity was assessed by 15 % SDS-PAGE. As shown in Figure 6B (lane 1), MDP2 showed a characteristic low mobility in SDS-PAGE: although the predicted molecular mass of MDP2 is 13.68 kDa (without the N-terminal methionine), it has migrated anomalously as a protein with almost 1.4-fold higher molecular mass (~19 kDa). This was the first indirect evidence that indicated the presence of ID in the MDP2 structure. The identity of this protein as MDP2 was confirmed by LC-MS/MS peptide mapping experiments using as samples both the anomalous band, excised from the gel, and also the purified protein. We obtained over 70 % of protein sequence coverage. Intact protein analysis was performed to determine the exact molecular mass of recombinant MDP2. Average spectra from 1000 scans were detected with isotopic resolu-

tion with charge states spanning from 13+ to 21+ (Figure 3A). We obtained FTMS spectra with isotopic resolution (see Figure 3B for charge state 18+ as an example) and found that recombinant MDP2 does not contain the N-terminal methionine. Deconvolution of FTMS spectra led us to the monoisotopic molecular mass of 13,683.6142 Da (Figure 3C), only 0.05 Da (accuracy of 3.9 ppm) different from the theoretical mass expected with the N-terminal methionine removed.

*MDP2 binds dsDNA* – The DNA-binding ability of MDP2 was confirmed by EMSA on agarose gel. DNA-protein complexes were formed upon incubation of MDP2 with the dsDNA (30 bp) (Figure 2). With increasing protein amount (1 - 2  $\mu\text{g}$ ), larger complexes were visualized, which eventually failed to enter the gel.

*MDP2 resists to aggregation induced by heat and chemical denaturants* - The resistance of MDP2 to heat was assessed by the capability of the protein to remain in solution after 10 min of boiling temperature, as evidenced by SDS-PAGE (Figure 4A). Globular proteins used as controls, such as CDA, GMPR, EPSP and human PNP, on the other hand, disappeared from the supernatant, assembling an insoluble pellet. This result, however, does not allow the distinction of the MDP2 in one of the four groups of heat-resistant proteins proposed by Kim and colleagues (12), since this experiment does not show any structural alteration induced by the heat treatment.

Precipitation induced by a chemical denaturant was also assessed using TCA in different final concentrations (1, 3 and 5 %). Whereas MDP2 remained in solution in all TCA concentrations tested, the globular proteins used as controls denatured and precipitated even at 1% TCA, assembling an insoluble pellet (Figure 4B).

*MDP2 is highly sensitive to proteolytic cleavage* - Proteolytic sensitivity test was performed with proteinase K, a broad range protease. The quantitative analysis of the 15 % SDS-PAGE using the ChemiDoc™ MP showed that MDP2 has segments in its structure highly sensitive to protease cleavage. With only 5 min of reaction, at a ratio of 1:2,000, about 50 % of the MDP2 had already been cleaved by the protease (Figure 5, lane 4); the same digestion pattern for CDA and BSA was only perceived at a ratio of 1:600 (Figure 5, lane 6); in other words: it was

necessary 3-fold less proteinase K to obtain 50% of MDP2 cleavage in comparison to the globular proteins. While CDA and BSA showed drags of cleavage, without the formation of intermediate structures, MDP2 exhibited the formation of well-defined intermediate bands on gel. From the ratio of 1:10,000 up to 1:600 two intermediate bands were formed, both with molecular weights with approximately 16.3 and 15.4 kDa (Figure 5, lanes 2-6); the 16.3 kDa band was also observed in the ratios 1:250 and 1:10 (lanes 7-8). These observations suggest that, although MDP2 has highly sensitive segments to protease cleavage, another diagnostic of IDRs, the protein may also have some resistant motifs, which failed to be cleaved.

*MDP2 self-associates in solution* - Data related to the quaternary structure of the MDP2 was obtained by two different approaches. First, the elution profile of the protein, obtained by gel filtration using a Superdex 200 column, revealed a single peak eluted just before the elution of the Blue Dextran 2000 used to obtain the void volume ( $V_0$ ) of the column (Figure 6A). This same result was observed in all three tests performed. The identity of this peak as MDP2 was confirmed on 15 % SDS-PAGE (not shown). This result suggests that MDP2 forms large aggregates in solution due to a high propensity to self-aggregation.

The second experiment was performed using chemical crosslinking with glutaraldehyde. A drop of purified MDP2 (15  $\mu\text{L}$  of 0.5  $\text{mg mL}^{-1}$  MDP2 in buffer C) was submitted to the presence of glutaraldehyde 25 % in a crystallization support. Aliquots were taken as a function of time and analyzed on 15 % SDS-PAGE. The control without glutaraldehyde revealed a unique band corresponding to monomeric MDP2 abnormally migrated (~19 kDa) as previously observed. Crosslinked species appeared on the top of the gel in all the aliquots taken; these species almost failed to enter in the 15 % SDS-PAGE (Figure 6B), suggesting the formation of a cluster of MDP2. This result is in accordance with the elution profile obtained by gel filtration, since both experiments point to a pronounced self-association propensity by MDP2.

## DISCUSSION

Despite the fact that the *M. tuberculosis* is one of the most widely studied pathogen so far, the bacterium metabolism remains a puzzle and a

challenge for the scientific community (11). Tuberculosis is still a deadly disease, and the emergence of MDR and XDR strains in the past decades only worsened the current situation (10,11). Therefore, a better understating of the pathogenic process, from the infection to the disease progress, may lead to the development of effective tools against this striking pathogen (50).

In the last decade, an intricate class of proteins called IDPs started to be considered prime targets for drug development (1), once these proteins are deeply related to some disease-associated proteins, such as  $\alpha$ -synuclein (27), p53 (51),  $\tau$  protein (52), and cystic fibrosis transmembrane conductance regulator (53). Some IDPs are also crucial for the action of some pathogens, such as the adenovirus early region 1A (E1A) oncoprotein, which changes the host cells by simultaneously recruiting CBP/p300 and phosphorylated retinoblastoma protein (pRb) into a ternary complex (pRb:E1A:CBP) that is capable of overriding the cell-cycle checkpoints of the host (1,54).

The *M. tuberculosis* also incorporated the adaptive potential of IDPs in its metabolism: amongst the eubacteria, the bacillus has an anomalous high disorder frequency in its proteome (9.1 %), which may be a result of its high G-C content and raised propensity to enrichment with the amino acids Ala, Gly, Pro, Arg and Trp (3). The intrinsically disorder ubiquitin-like protein (Pup) found in the mycobacteria, for instance, tags proteins destined for degradation; pupylated proteins are then recognized by the proteasome-assisted ATPase (Mpa) (1,55). Anurag and Dash (11) also announced other IDPs found in the *M. tuberculosis* proteome using a screening method. The authors suggested that those IDPs that were (1) essential, (2) absent in humans and their gut and mouth microbiome, and (3) known to interact with other proteins, would be potential drug targets. Of 222 proteins predicted to be intrinsically disordered by DISpro and IUPred predictors, only 13 were considered intrinsically disordered essential proteins (IDEPs) by the authors (11).

Here, we experimentally demonstrated that the MDP2 of *M. tuberculosis* behaves as an IDP by gathering several indirect evidences of disorder proposed by Tompa (28). These experiments were motivated by an *in silico* IDP prediction using PONDR®, which suggested that the protein have two long disordered regions (LDRs) –

defined as having 30 or more consecutive disordered residues - in the N-terminal region and in the middle portion of the protein. These regions are highly abundant among eukaryotes, but quite rare in prokaryotes (3,7,9).

Accordingly to the TubercuList database (32), the MDP2 is not considered essential for the survival of the bacteria, as suggested by the Himar1-based transposon mutagenesis in H37Rv strain (56), therefore it may not be a good target for drug design at first glance.

Recent studies about the MDP2 structure and function have already been conducted by some groups. Ghosh and colleagues (33) have proposed that the MDP2 may play an important role in lipid biosynthesis by repressing *groEL1* and *fasII* operon genes. In the same study, it was confirmed that the protein binds to DNA in a sequence-independent manner, as previously observed by two other works (30,31), and that it also anchors the nucleoid to the plasma membrane (33). It is worth noting that the sequences of tetrapeptides of the MDP2 suggested to bind to the DNA (KAAK and PAKK) are composed mainly by disorder-promoting amino acids (30,31); nucleic acid binding (3), as well as metal ion binding, heme binding, and interaction with other proteins and membrane bilayers, are some of the most major functions carried out by IDPs (6). Even though no work has suggested that MDP2 may be intrinsically disordered, its binding promiscuity already observed could be pointed out as a sign of ID, since adaptability in binding is expected for this class of proteins (1,4,57).

Despite not having presented data, Ghosh and colleagues also reported that MDP2 showed an anomalous mobility of ~21 kDa in 15 % SDS-PAGE (33), which is in agreement with our observation of a band with ~19 kDa that was unambiguously identified by mass spectrometry as MDP2 of *M. tuberculosis* (Figure 3). IDPs tend to bind less SDS than globular proteins due to their unusual amino acid composition, commonly displaying an apparent molecular weight 1.2-1.8 times higher than the real one in SDS-PAGE experiments (28). This property has been reported for many IDPs, such as ERD10 and ERD14 (17), AtREM1.3 remorin protein (19) and TCP8 transcription factor (22) of *Arabidopsis thaliana*, XPA of *Xenopus laevis* (21),  $\delta$ Ag of hepatitis delta virus (HDV) (20), human  $\alpha$ -synuclein (27). To demonstrate that the recombinant MDP2 retained its DNA-binding ability, we also per-

formed EMSA that showed the protein does bind to dsDNA (Figure 2), confirming previous results from Werlang and Ghosh (30,33).

Another striking feature of IDPs is their stability under high temperatures and strong denaturants. Since these proteins have no hydrophobic core, they are not expected to lose solubility and aggregates at elevated temperatures. Drawing on this unusual feature, many IDPs have been purified by a heat-treatment step along the purification protocol (14). Heat-treatment has also been used by IDPs enrichment for proteome studies (12,13,16) and for the characterization of presumable IDPs (17,27).

Here we have compared the behavior of MDP2 and four globular proteins (CDA, GMPR, PNP and EPSPS) under an extreme condition of temperature. Our results pointed out that the MDP2 showed a behavior expected for an IDP, since it remained soluble under this condition, while all globular proteins tested precipitated (Figure 4A). We have also showed the stability of MDP2 in the presence of the strong chemical denaturant trichloroacetic acid (TCA). The use of acids, such as PCA (perchloric acid) and TCA, for IDP's enrichment are widely spread (16,18). Globular proteins tend to switch to a random coil state at low pH, due to the imbalances of charge caused by the protonation of negatively charge side chains; in this state, the protein exposes its hydrophobic groups, leading to an aggregation and subsequently precipitation (16). This behavior was observed for the structured proteins that we have tested (Figure 4B). On the other hand, IDPs, as well as MDP2, already have exposed segments of its structure with high percentages of charged residues, and consequently they are not affected by a low pH environment; oddly enough, decreasing the pH may actually shift the conformation of some IDPs toward a more compact state (29).

Protease digestion is another useful tool to have insights into protein structure and flexibility (28,58). IDPs are known to have high flexible regions inside their structure that facilitates their recognition by proteases, making these proteins extremely sensitive to proteolysis (24,28). With this in mind, we hypothesized that MDP2 would be more sensitive to cleavage by proteinase K in comparison to BSA and CDA, both known to be structured. According to our results, MDP2 was indeed highly sensitive to proteolysis; however, the protein was not as sensitive as other IDPs (17). There are at least two explanations for this

result. First, it is possible that some segments of the MDP2 are not so flexible than we were expecting, since some intermediate bands resistant to cleavage were formed (Figure 5). Second, some segments of the protein are not accessible to the protease due to its self-association in solution explained in the following.

In the last set of experiments, we asked whether the protein was capable of self-assembling in multiple oligomers or not. It was already reported that MDP2 forms higher-order oligomers in the presence of DNA fragments (30,33). Our results of gel filtration showed that the protein eluted in the void volume of the Superdex 200 column (Figure 6A). Weinreb and colleagues (27) also observed this phenomenon with the NACP, a protein implicated in Alzheimer's disease and learning, and proposed two possible explanations: either (1) the protein was an oligomer or (2) it has an elongated shape in comparison to structured proteins used as standards. To answer this question for the MDP2, we performed a crosslinking reaction with glutaraldehyde. Crosslinked MDP2 almost failed to enter on 15 % SDS-PAGE, seeming to form large clusters (Figure 6B); this strongly suggests that the protein failed to go through the resin's pores due to its self-association and subsequently formation of large structures. In fact, these two features, together with sequence-independent DNA-binding, are expected for nucleoid-associated proteins (59), the group of proteins in which MDP2 may belong.

Altogether our results point out that the MDP2 from *M. tuberculosis* is an IDP capable of self-association in solution, as summarized in Table 2, supporting our previous disorder prediction using the PONDR®. Further studies will be needed to address how the protein uses its intrinsic disorder to interact with the DNA, and to verify if it alternates between disorder-to-order states in the presence of its interaction partner. It is also important to identify the type of aggregation formed by MDP2 in solution and to understand how the protein assembles these large structures inside the cell.

## REFERENCES

1. Tompa, P. (September 16, 2012) Intrinsically disordered proteins: a 10-year recap. *Trends Biochem Sci.* 10.1016/j.tibs.2012.08.004
2. Dunker, A. K., *et al.* (2013) What's in a name? Why these proteins are intrinsically disordered. *Intrinsically Disordered Proteins* **1**, 10.4161/idp.24157
3. Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F., and Jones, D. T. (2004) Prediction and Functional Analysis of Native Disorder in Proteins from the Three Kingdoms of Life. *J. Mol. Biol.* **337**, 635-645
4. Uversky, V. N. (2013) Digested disorder: Quarterly intrinsic disorder digest (January/February/March). *Intrinsic Disordered Proteins* 10.4161/idp.25496
5. Janin, J., and Sternberg, M. J. E. (2013) Protein flexibility, not disorder, is intrinsic to molecular recognition. *F1000 Biol Rep.* **5**, 10.3410/B5-2
6. Uversky, V. N. (2011) Intrinsically disordered proteins from A to Z. *Int J Biochem Cell Biol* **43**, 1090-1103
7. Forman-Kay, D. J., and Mittag, T. (September 13, 2013) From Sequence and Forces to Structure, Function, and Evolution of Intrinsically Disordered Proteins. *Structure* **21**, 10.1016/j.str.2013.08.001
8. Theillet, F., *et al.* (2013) The alphabet of intrinsic disorder. I: Act like a Pro: On the abundance and roles of proline residues in intrinsically disordered proteins. *Intrinsically Disordered Proteins* 10.4161/idp.24360
9. Peng, Z., Mizianty, M. J., and Kurgan, L. (2013) Genome-scale prediction of proteins with long intrinsically disordered regions. *Proteins* **82**, 145-158.
10. Lawn, S. D., and Zumla, A. I. (2011) Tuberculosis. *The Lancet* **378**, 57-72.
11. Anurag, M., and Dash, D. (2009) Unraveling the potential of intrinsically disordered proteins as drug targets: application to *Mycobacterium tuberculosis*. *Mol BioSyst.* **5**, 1752-1757
12. Kim, T. D., Ryu, H. J., Cho, H. I., Yang, C., and Kim, J. (2000). Thermal Behavior of Proteins: Heat-Resistant Proteins and Their Heat-Induced Secondary Structural Changes. *Biochemistry* **39**, 14839-14846
13. Tsetkov, P., Myers, N., Moscovitz, O., Sharon, M., Prilusky, J., and Shaul, Y. (2012) Thermo-resistant intrinsically disordered proteins are efficient 20S proteasome substrates. *Mol BioSyst.* **8**, 368-373
14. Livernois, A. M., Hnatchuk, D. J., Findlater, E. R., and Graether, S. P. (2009) Obtaining highly purified intrinsically disordered by boiling lysis and single step ion exchange. *Anal Biochem* **392**, 70-76
15. Csizmók, V., Dosztányi, Z., Simon, I., and Tompa, Peter. (2007) Towards Proteomic Approaches for the Identification of Structural Disorder. *Curr Protein Pept Sci.* **8**, 173-179
16. Cortese, M. S., Baird, J. P., Uversky, V. N., and Dunker, A. K. (2005). Uncovering the Unfoldome: Enriching Cell Extracts for Unstructured Proteins by Acid Treatment. *Journal of Proteome Research* **4**, 1610-1618
17. Kovacs, D., Kalmar, E., Torok, Z., and Tompa, P. (2008) Chaperone Activity of ERD10 and ERD14, Two Disordered Stress-Related Plant Proteins. *Plant Physiol.* **147**, 381-390
18. Campos, F., Guillén, G., Reyes, J. L., and Covarrubias, A. A. (2011) A general method of protein purification for recombinant unstructured non-acidic proteins. *Protein Expr Purif.* **80**, 47-51
19. Marín, M., Thallmair, V., and Ott, T. (2012) The intrinsically Disordered N-terminal Region of AtREM1.3 Remorin Protein Mediates Protein-Protein Interactions. *The Journal of Biological Chemistry* **47**, 39982-39991
20. Alves, C., Cheng, H., Roder, H., and Taylor, J. (2010) Intrinsic disorder and oligomerization of the hepatitis delta virus antigen. *Virology* **407**, 333-340
21. Iakoucheva, L. M., *et al.* (2001) Identification of intrinsic order and disorder in the DNA repair protein XPA. *Protein Science* **10**, 560-571

22. Valsecchi, I., *et al.* (2013) The intrinsically disordered C-terminal region of *Arabidopsis thaliana* TCP8 transcription factor acts both as a transactivation and self-assembly domain. *Mol BioSyst.* **9**, 2282-2295
23. Paz, A., *et al.* (2008) Biophysical Characterization of the Unstructured Cytoplasmic Domain of the Human Neuronal Adhesion Protein Neuroligin 3. *Biophysical Journal* **95**, 1928-1944.
24. Fontana, A., de Laureto, P. P., Spolaore, B., and Frare, E. (2012) Identifying disordered regions in proteins by limited proteolysis. *Methods Mol Biol.* **896**, 297-318.
25. Haaning, S., *et al.* (2008) An Unusual Intrinsically Disordered Protein from The Model Legume Lotus japonicus Stabilizes Proteins in vivo. *J Biol Chem.* **283**, 31142-31152
26. Dasgupta, I., Sanglas, L., Enghild, J., and Lindberg, I. (2012) The Neuroendocrine Protein 7B2 Is Intrinsically Disordered. *Biochemistry* **51**, 7456-7464
27. Weinreb, P. H., Zhen, W., Poon, A. W., Conway, K. A., and Lansburym P. T. Jr. (1996). NACP, A Protein Implicated in Alzheimer's Disease and Learning, Is Natively Unfolded. *Biochemistry* **35**, 13709-13715
28. Tompa, P. (2002) Intrinsically unstructured proteins. *TRENDS Biochem Sci.* **27**, 527-533
29. Tompa, P. (2010) Structure and function of intrinsically disordered proteins, Chapman & Hall/CRC, Taylor & Francis Group, NW
30. Werlang, I. C. R., Schneider, C. Z., Mendonça, J. D., Palma, M. S., Basso, L. A., and Santos, D. S. (2009) Identification of Rv3852 as a nucleoid-associated protein in *Mycobacterium tuberculosis*. *Microbiology* **155**, 2652-2663
31. Sharadamma, N., Harshavardhana, Y., Singh, P., Muniwappa, K. (2010) *Mycobacterium tuberculosis* nucleoid-associated DNA-binding protein H-NS binds with high-affinity to the Holliday junction and inhibits strand exchange promoted by RecA protein. *Nucleic Acids Res.* **38**, 3555-3569
32. Lew, J. M., Kapopoulou, A., Jones, L. M., and Cole, S. T. (2011) TubercuList - 10 years after. *Tuberculosis* **91**, 1-7.
33. Ghosh, S., Indi, S. S., and Nagajara, V. (2013) Regulation of Lipid Biosynthesis, Sliding Motility, and Biofilm Formation by a Membrane-Anchored Nucleoid-Associated Protein of *Mycobacterium tuberculosis*. *J Bacteriol.* **195**, 1769-1778
34. Gordon, B. R. G., Imperial, R., Wang, L., Navarre, W. W., and Liu, Jun. (2008) Lsr2 of *Mycobacterium tuberculosis* Represents a Novel Class of H-NS-Like Proteins. *J Bacteriol.* **190**, 7052-7059
35. Li, X., Romero, P., Rani, M., Dunker, A. K., and Obradovic, Z. (1999) Predicting protein disordered for N-, C- and internal regions. *Genome Informatics* **10**, 30-40
36. Romero, P., Obradovic, Z., Li, X., Garner, E., Brown, C., and Dunker, A. K. (2001) Sequence complexity of disordered protein. *Struct Funct. Gen.* **42**, 38-48
37. Romero, P., Obradovic, Z., and Dunker, A. K. (1997) Sequence data analysis for long disordered regions prediction in the calcineurin family. *Genome Informatics* **8**, 110-124
38. Uversky, V. N., Gillespie, J. R., and Fink, A. L. (2000) Why Are "Natively Unfolded" Proteins Unstructured Under Physiologic Conditions? *PROTEINS: Structure, Function, and Genetics* **41**, 415-427
39. Xue, B., Oldfield, C. J., Dunker, A. K., and Uversky, V. N. (2009) CDF It All: Consensus Prediction of Intrinsically Disordered Proteins Based on Various Cumulative Distribution Functions. *FEBS Lett.* **9**, 1469-1474
40. Cole, S. T., *et al.* (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**, 537-544
41. Burgess, R. R. (1991) Use of Polyethyleneimine in Purification of DNA-Binding Proteins. *Methods in Enzymology* **208**, 3-10
42. Klammer, A.A. and MacCoss, M.J. (2006) Effects of modified digestion schemes on the identification of proteins from complex mixtures. *J Proteom Res.* **5**, 695-700
43. Moritz, R.L. (2007) Configuration, column construction, and column packing for a capillary liquid chromatography system. *CSH Protoc.*, 10.1101/pdb.prot4578.

44. Fadoulglou, V. E., Kokkinidis, M., and Glykos, N. M. (2008) Determination of protein oligomerization state: Two approaches based on glutaraldehyde crosslinking. *Analytical Biochemistry* **373**, 404-406
45. Sánchez-Quitian, Z. A., *et al.* (2011) Crystal structure determination and dynamic studies of *Mycobacterium tuberculosis* Cytidine deaminase in complex with products. *Archives of Biochemistry and Biophysics* **509**, 108-115
46. Oliveira, J. S., Mendes, M. A., Palma, M. S., Basso, L. A., and Santos, D. S. (2003) One-step purification of 5-enolpyruvylshikimate-3-phosphate synthase enzyme from *Mycobacterium tuberculosis*. *Protein Expression and Purification* **28**, 287-292
47. Martinelli, L. K. B., Ducati, R. G., Rosado, L. A., Breda, A., Selbach, B. P., Santos, D. S., and Basso, L. A. (2011). Recombinant *Escherichia coli* GMP reductase: kinetic, catalytic and chemical mechanisms, and thermodynamics of enzyme-ligand binary complex formation. *Mol. BioSyst.* **7**, 1289-1305
48. de Azevedo, W. F. Jr., *et al.* (2003). Crystal structure of human purine nucleoside phosphorylase at 2.3 Å resolution. *Biochem. Biophys. Res. Commu.* **3**, 545-552
49. Oldfield, C. J., Cheng, Y., Cortese, M. S., Brown, C. J., Uversky, V. N., and Dunker, A. K. (2005) Comparing and Combining Predictors of Mostly Disordered Proteins. *Biochemistry* **44**, 1989-2000
50. Delogu, G., Sali, M., and Fadda, G. (2013) The Biology of *Mycobacterium tuberculosis* Infection. *Mediterranean Journal of Hematology and Infectious Diseases* 10.4084/MJHID.2013.070
51. Wells, M. *et al.* (2008). Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain. *PNAS* **105**, 5762-5767
52. Mukrasch, M. *et al.* (February, 2009) Structural Polymorphism of 441-Residue Tau at Single Residue Resolution. *PLoS Biology* 10.1371/journal.pbio.1000034
53. Baker, J. M. R. *et al.* (2007). CFTR regulatory region interacts with NBD1 predominantly via multiple transient helices. *Nature Structural & Molecular Biology* **14**, 738-745.
54. Ferreón, J. C., Martínez-Yamout, M. A., Dyson, H. J., and Wright, E. (2009) Structural basis for subversion of cellular control mechanisms by the adenoviral E1A oncoprotein. *PNAS* **106**, 13260-13265.
55. Pearce, M. J., *et al.* (2008) Ubiquitin-like protein involved in the proteasome pathway of *Mycobacterium tuberculosis*. *Science* **322**, 1104-1107
56. Griffin, J. E., Gawronski, J. D., DeJesus, M. A., Ioerger, T. R., Akerley, B. J., and Sasseti, C. M. (September 29, 2011) High-Resolution Phenotypic Profiling Defines Genes Essential for Mycobacterial Growth and Cholesterol Catabolism. *PLoS Pathogens* 10.1371/journal.ppat.1002251
57. Oldfield, C. J., Meng, J., Yang, M. Q., Uversky, V. N., and Dunker, A. K. (2008) Flexible nets: disorder and induced fit in the associations of p53 and 14-3-3 with their partners. *BCM Genomics* 10.1186/1471-2164-9-S1-S1
58. Dunker, A. K. *et al.* (2001) Intrinsically disordered protein. *J Mol Graph Model* **19**, 26-59
59. Dame, R. T. (2005) The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. *Molecular Microbiology* **56**, 858-870

*Acknowledgments* – We thank Dr. Leonardo K. B Martinelli and Rafael A. Caceres for providing homogeneous GMPR and human PNP, respectively; and Dr. Leonardo A. Rosado for supplying us with good literature.

## FOOTNOTES

\*This work was supported by funds awarded by Decit/SCTIE/MS-MCT-CNPq-FNDCT-CAPES to National Institute of Science and Technology on Tuberculosis (INCT-TB) to D.S.S. and L.A.B. L.A.B. and D.S.S. also acknowledge financial support awarded by FAPERGS-CNPq-PRONEX-2009. L.A.B. (CNPq, 520182/99-5) and D.S.S. (CNPq, 304051/1975-06) are Research Career Awardees of the National Research Council of Brazil (CNPq). Z.A.S.Q. acknowledges a PhD scholarship awarded by CAPES. B.L.A. and P.C.P. acknowledges a master scholarship awarded by CAPES and CNPq, respectively, and P.E. acknowledges a post-doctoral fellow of CAPES.

<sup>1</sup>To whom correspondence may be addressed: Centro de Pesquisas em Biologia Molecular e Funcional (CPBMF), Instituto Nacional de Ciência e Tecnologia em Tuberculose, Pontifícia Universidade Católica do Rio Grande do Sul, Av. Ipiranga, 6681/prédio 92A, TECNOPUC, Porto Alegre, RS, Brazil, tel./fax: (51) 3320-3629; e-mail: cristiano.bizarro@pucls.br.

<sup>2</sup>The abbreviations used are: MDP2, mycobacterial DNA-binding protein 2; IDPs, intrinsically disordered proteins; TCA, trichloroacetic acid; CDA, cytidine deaminase; EPSPS, 5-enolpyruvylshikimate-3-phosphate synthase; GMPR, guanosine monophosphate reductase; PNP, human purine nucleoside phosphorylase; IDRs, intrinsically disordered regions; CDF, cumulative distribution function; C-H plot, charge-hydrophathy plot.

## FIGURE LEGENDS

**FIGURE 1.** Disorder prediction for MDP2 of *M. tuberculosis*. (A) PONDR® VL-XT prediction curve of the disorder in MDP2. Each residue was plotted against its disorder-prediction scores. Residues that show values above 0.5 are considered disordered. The thicker black line on the threshold indicates a region of 40 or more consecutive disordered residues. (B) CDF curve of the MDP2. The optimal separation line, represented as seven points, was previously estimated for a large number of ordered and disordered proteins (39). Ordered proteins have a CDF curve in the upper part of the graph, whereas disordered ones reside in the lower part. (C) Charge-hydrophathy plot. Wholly disordered proteins are mainly located on the left side of the order-disorder boundary (white squares), whereas wholly ordered are on the right side (black triangles) (49). MDP2 is located upon the borderline between ordered and disordered proteins (black circle).

**FIGURE 2.** EMSA for the analysis of DNA-MDP2 interaction. C = dsDNA (300 ng) in the absence of MDP2. Lanes 1 - 6 = dsDNA (300 ng) incubated with increasing amounts of MDP2 (0.1 - 2.0 µg, respectively). Black triangle represents increasing concentration of MDP2 toward the right side.

**FIGURE 3.** Determination of MDP2 monoisotopic molecular mass by mass spectrometry analysis. (A) ESI-FTMS spectra of MDP2 spanning from charge state 13+ to 21+. (B) Isotopic envelope of charge state 18+. (C) Deconvoluted spectra of MDP2 resulted in a peak corresponding to a monoisotopic molecular mass of 13,683.6142 Da. This value matches the theoretical monoisotopic mass for MDP2 (13,683.6682 Da) with 3.9 parts-per-million mass accuracy (3.9 ppm).

**FIGURE 4.** Resistance to aggregation induced by heat and chemical denaturants. (A) Heat treatment of MDP2, CDA, GMPR, PNP and EPSPS. Proteins were (+) and were not (-) incubated at 100 °C for 10 min. Samples were sedimented with centrifugation and the supernatants were analyzed on 15 % SDS-PAGE. (B) Chemical treatment with TCA. Proteins were treated with different final concentrations of TCA (1, 3 and 5 %). Supernatants were analyzed on 15 % SDS-PAGE. Lane 1 = samples not treated. Lanes 2-4 = treatment with 1, 3 and 5 % of TCA respectively. Black triangle represents increasing concentration of TCA toward the right side.

**FIGURE 5.** Limited proteolysis with proteinase K. Globular proteins (BSA and CDA) and MDP2 were and were not treated with increasing molar ratios of the protease. Samples were kept at 25 °C in a thermal cycler for 5 min. The reaction was stopped by adding protein loading buffer and heating at 99 °C for 10 min. Proteolysis was analyzed on 15 % SDS-PAGE. Lanes 1 = proteins not treated. Lanes 2 – 8 = reactions with 1:10,000, 1:2,000, 1:1,500, 1:900, 1:600, 1:250 and 1:10 of proteinase K, respectively. Black triangle represents increasing concentration of proteinase K toward the right side.

**FIGURE 6.** Oligomeric structure of MDP2 assessed by gel filtration and *in vitro* crosslinking with 25 % glutaraldehyde. (A) Chromatogram resulted from size-exclusion column showing the elution of the Blue Dextran 2000 ( $V_0$ ) and the MDP2 almost at the same volume. Elution was monitored at 280, 254 and 215 nm. Since MDP2 does not absorb at 280 and 254 nm due to a lack of aromatic residues, we only show the absorbance at 215 nm (solid line). Blue Dextran 2000, on the other hand, absorbs at 280 nm and 254 nm (dashed line). (B) 15 % SDS-PAGE showing the crosslinking test. MDP2 were (+) and were not crosslinked with glutaraldehyde (-). Aliquots were taken as a function of time. M = BenchMark™ Protein Ladder. Lane 1 = control without glutaraldehyde. Lanes 2-7 = crosslinked MDP2 collected in different periods of time (10, 20, 30, 40, 50 and 60 minutes respectively).

Table 1. Amino acid composition performed by ProtParam tool.

<sup>a</sup> Amino acid	<sup>b</sup> Amount	<sup>c</sup> Frequency
Pro (P)	19	14.2 %
Glu (E)	4	3.0 %
Ser (S)	10	7.5 %
Gln (Q)	6	4.5 %
Lys (K)	16	11.9 %
Ala (A)	30	22.4 %
Gly (G)	3	2.2 %
Asp (D)	6	4.5 %
Thr (T)	6	4.5 %
Arg (R)	8	6.0 %
Met (M)	1	0.7 %
Asn (N)	3	2.2 %
Val (V)	6	4.5 %
His (H)	1	0.7 %
Leu (L)	12	9.0 %
Phe (F)	0	0.0 %
Tyr (Y)	0	0.0 %
Ile (I)	3	2.2 %
Trp (W)	0	0.0 %
Cys (C)	0	0.0 %

<sup>a</sup>Residues are arranged according to their increasing intrinsic disorder propensity. <sup>b</sup>Amount of each residue in the MDP2 sequence. <sup>c</sup>Relative frequency of each residue.

Table 2. Summary of the tools used for MDP2 validation as an IDP and the results obtained respectively.

<b>Tools used for IDP validation</b>	<b>Behavior of MDP2</b>
PONDR® analysis (% of disorder)	86 %
Anomalous electrophoretic mobility ( $Mw_{app}/Mw_{real}$ )	1.4-fold
Resistance to heat (+/-)	+
Resistance to chemical denaturation (+/-)	+
Elution profile during gel filtration ( $Mw_{app}$ )	> 440 kDa (aggregates)
Glutaraldehyde crosslinking ( $Mw_{app}$ )	> 220 kDa (aggregates)
Enhanced proteolytic sensitivity (+/-)	+

Figure 1

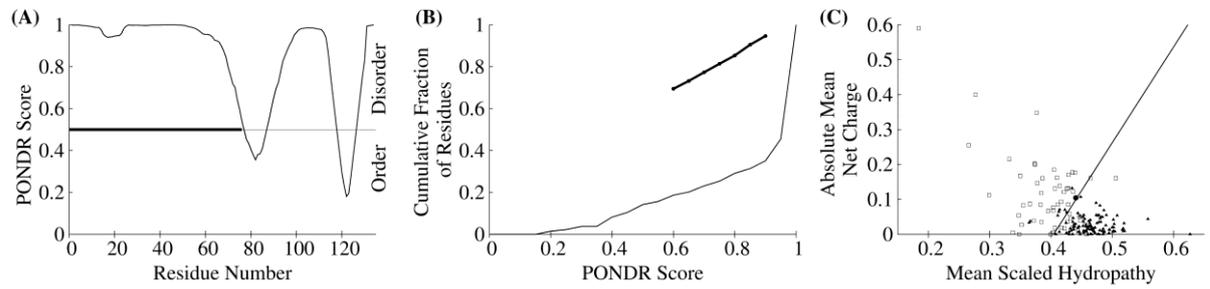


Figure 2

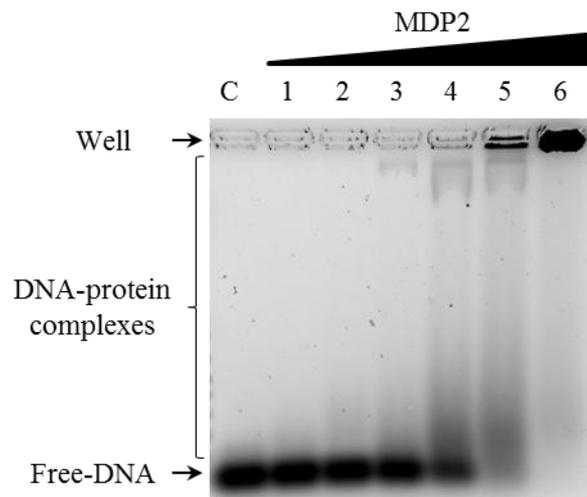
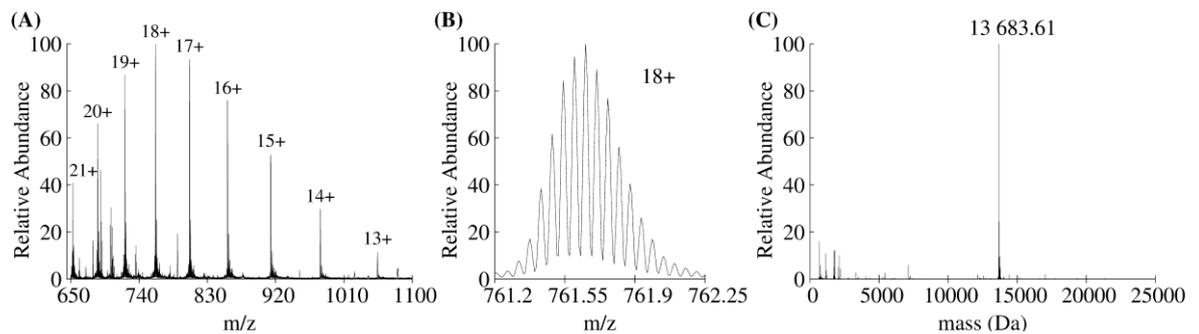


Figure 3





## **CAPÍTULO 3**

### **RESULTADOS PRELIMINARES E PERSPECTIVAS**

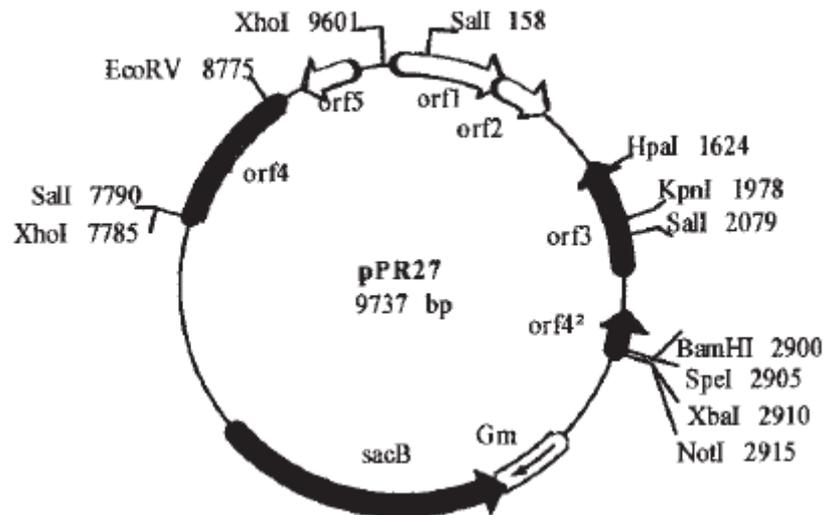
### **Nocaute do gene *hns* de *M. tuberculosis***

Além dos estudos relatados no capítulo 2, este trabalho também se propôs a realizar o nocaute do gene *hns* de *Mycobacterium tuberculosis* H37Rv, responsável por codificar a proteína MDP2 (Rv3852, NP\_218369.1), a fim de se confirmar a não essencialidade do gene *hns* para a micobactéria já observada pela mutagênese mediada por transposon (Sassetti, Boyd *et al.*, 2003). A técnica utilizada para tal fim baseia-se na mutagênese sítio-direcionada por substituição gênica ou troca alélica, a qual troca um alelo selvagem de interesse por um alelo mutante (geralmente truncado por um gene que confere resistência a um determinado antibiótico) por meio de um evento de recombinação homóloga (Parish, 2001). Este tipo de experimento serve como um passo inicial para a caracterização funcional de proteínas e de fatores de virulência presentes no *M. tuberculosis*. Apesar da micobactéria possuir as ferramentas necessárias para ocorrer o processo de recombinação homóloga, este evento ocorre a uma frequência bastante baixa ( $< 10^{-3}$  eventos/célula), dificultando a obtenção de mutantes no *M. tuberculosis*. Além disso, sabe-se que a maioria das células transformadas realizam eventos de recombinação ilegítima (80 a 100 %), dificultando ainda mais a obtenção de um mutante (Parish, 2001).

Para que se possa selecionar um mutante que sofreu um evento de duplo *cross-over* (DCO), ou seja, que realizou a troca do gene selvagem pelo gene mutante, sem inserir o restante do plasmídeo, faz-se necessário utilizar certas ferramentas genéticas que compensem a baixa eficiência de transformação do *M. tuberculosis* e que permitam uma seleção eficiente dos mutantes que sofreram o evento de DCO (Parish, 2001). Para tanto, certos marcadores contra-seletivos têm sido empregados no corpo dos vetores plasmidiais usados no evento de mutagênese, os quais garantem a seleção das colônias que integraram o alelo mutante do gene e que perderam o corpo do vetor plasmidial que carrega o marcador contra-seletivo. Um exemplo de marcador comumente empregado nesta técnica consiste na expressão do gene *sacB*, o qual é letal à micobactéria na presença de sacarose no meio de cultura (Parish, 2001). Outra estratégia bastante empregada em conjunto com o marcador contra-seletivo *sacB* baseia-se no uso de vetores replicativos termosensíveis, os quais têm sua replicação inibida em uma determinada temperatura. Um exemplo deste tipo de vetor é chamado de pPR27:*xyIE* (Figura 12), o qual será utilizado nos experimentos de mutagênese deste trabalho: além de possuir o gene *sacB*, o vetor pPR27:*xyIE* tem sua replicação inibida à

temperatura de 39 °C. Este mesmo vetor também possui o gene repórter *xylE*, o qual confere às colônias que possuem o plasmídeo íntegro a coloração amarela, garantindo uma seleção eficiente das células transformadas (Pelicic, 1997).

**Figura 12.** Plasmídeo termosensível pPR27 contendo o marcador contra-seletivo *sacB*.



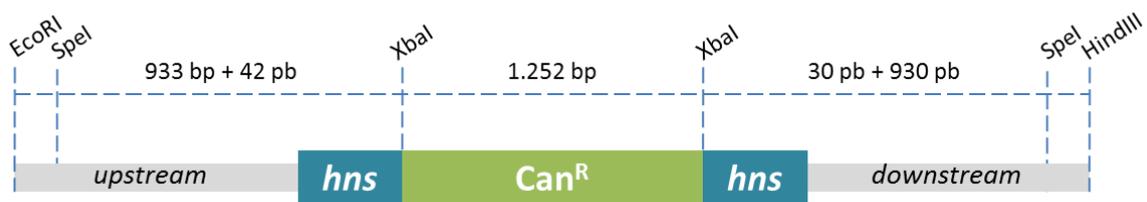
**Fonte:** Parish, 2001.

**Nota:** nesta representação do plasmídeo pPR27 não está representado o marcador seletivo por coloração *xylE* também utilizado neste trabalho. No entanto, esta sequência foi inserida no sítio da enzima BamHI.

O primeiro passo do experimento de mutagênese sítio-dirigida consistiu na construção de um vetor plasmidial recombinante, formado pelo vetor pUC19 (Invitrogen) ligado a duas regiões flanqueadoras *upstream* (933 pb + 42 pb do início da região codificadora) e *downstream* (930 pb + 30 pb do final da região codificadora) ao gene *hns* [Figura 13]. Ambos os fragmentos foram inicialmente amplificados a partir do DNA genômico de *M. tuberculosis* H37Rv por meio de uma reação em cadeia da polimerase (PCR), utilizando um par de oligonucleotídeos iniciadores para cada fragmento. Para o fragmento *upstream* foram utilizados os seguintes iniciadores: 5'-TTTT**GAATTC**ACTAGTCAGCGCGGAATCAAAG-3', contendo os sítios de reconhecimento para as enzimas EcoRI e SpeI (os sítios estão grifados em negrito), e 5'-GGGG**TCTAG**AGTCGCTCGGCTCGCTGTC-3', contendo o sítio para a enzima XbaI; para o fragmento *downstream* foram usados os seguintes iniciadores: 5'-GGGG**TCTAG**ACTGCTGATCCGGCAACTG-3', contendo o sítio para a enzima XbaI, e 5'-TTTT**AAGCTT**ACTAGTATCGACGGACAACAAGTGA-3', contendo os sítios para as enzimas SpeI e HindIII. Para a PCR também foram utilizados desoxirribonucleotídeos (dNTPs), dimetilsul-

fóxido (DMSO) numa concentração final de 10 % e a enzima *Pfu* DNA Polimerase (4G P&D), conhecida pela sua alta fidelidade replicativa. A confirmação da amplificação de ambos os fragmentos foi realizada por eletroforese em gel de agarose 1 % corado em banho de Gel-Red™ [Figura 14].

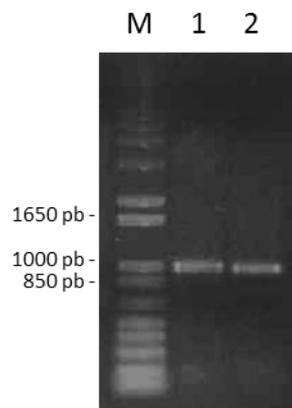
**Figura 13.** Sequência  $\Delta hns:Can^R$  utilizada como sítio para a mutagênese por troca alélica.



**Fonte:** Produção do próprio autor do trabalho.

**Nota:** Representação das sequências presentes na construção  $\Delta hns:Can^R$ , como a região *upstream* ao gene *hns* (933 pb), o início da sequência codificadora do gene *hns* (42 pb), o cassete que confere resistência ao antibiótico canamicina (1.252 pb), o final da sequência codificadora do gene *hns* (30 pb) e a região *downstream* ao gene *hns* (930 pb). Os sítios de reconhecimento para as enzimas de restrição também estão representados.

**Figura 14.** Eletroforese em gel de agarose dos produtos de PCR das regiões flangeadoras *upstream* e *downstream* ao gene *hns*.



**Fonte:** Produção do próprio autor do trabalho.

**Legenda:** M = Marcador de peso molecular de DNA 1 kb Plus. 1-2 = amplicons das regiões flangeadoras *upstream* (975 pb) e *downstream* (960 pb) respectivamente.

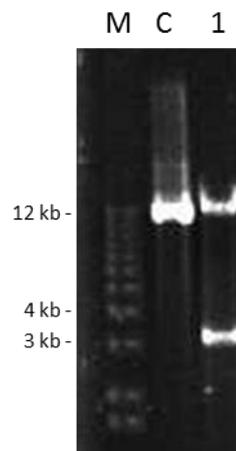
Os próximos passos consistiram em realizar clivagens com enzimas de restrição tanto dos dois fragmentos quanto do vetor pUC19, seguidas de reações de ligação por meio da enzima DNA Ligase (New England Biolabs) e, subsequentemente, clonar os vetores recombinantes em células de *E. coli* DH10B eletrocompetentes. O fragmento *upstream* foi o primeiro a ser ligado ao vetor pUC19 por meio das suas extremidades coesivas formadas pelas enzimas EcoRI e XbaI, seguido pelo fragmento *downstream*, contendo as extremidades coesivas

formadas pelas enzimas XbaI e HindIII. Após a ligação de ambos os fragmentos, foi necessário clivar o vetor recombinante com a enzima XbaI e, posteriormente, “cegar” as extremidades coesivas formadas pela enzima por meio de uma reação com a enzima *Pfu* DNA polimerase na presença de dNTPs. Após a obtenção das extremidades cegas entre os fragmentos *upstream* e *downstream*, adicionou-se um cassete com 1.252 pb que confere resistência ao antibiótico canamicina (Can<sup>R</sup>) entre os dois fragmentos, formando uma sequência denominada  $\Delta hns:Can^R$ . Este cassete foi previamente extraído do vetor pUC4K por meio de uma reação de clivagem com a enzima de restrição HincII, a qual forma extremidades cegas.

O vetor pUC19 tem como função servir de base para a construção da sequência nucleotídica, que servirá de sítio para o evento de recombinação homóloga, e de facilitar a manipulação dos fragmentos de DNA. Porém, para a realização do nocaute gênico no *M. tuberculosis*, faz-se necessário utilizar o vetor pPR27:*xyIE*, o qual é dotado de alguns marcadores seletivos e contra seletivos que facilitam a seleção das colônias mutantes, como: (1) origem de replicação micobacteriana termosensível, a qual garante a replicação do plasmídeo à temperatura de 32 °C, mas não a 39 °C; (2) presença do gene *sacB* de *B. subtilis*, o qual é letal para a micobactéria na presença de sacarose no meio de cultura, servindo de marcador contra seletivo; e (3) presença do gene repórter *xyIE*, o qual pode ser facilmente testado por meio da adição de uma solução de catecol às colônias, as quais têm sua coloração alterada de branco para amarelo na presença do gene (Pelicic, 1997). Além destes marcadores, geralmente adiciona-se um gene que confere resistência a um determinado antibiótico para servir de marcador seletivo das colônias transformadas; neste trabalho, utilizou-se a resistência à canamicina.

Para transferir a sequência construída no vetor pUC19 para o pPR27:*xyIE*, foi necessário clivar tanto o primeiro vetor com a enzima SpeI, a fim de liberar a sequência construída ( $\Delta hns:Can^R$ ), quanto o segundo, para que a reação de ligação pudesse ser realizada. Após esta etapa de ligação, o tamanho da sequência construída ( $\Delta hns:Can^R$ ) foi verificado por clivagem com a enzima SpeI seguida de eletroforese em gel de agarose 1 % [Figura 15] e a sua identidade foi confirmada por sequenciamento automatizado de DNA (ABI-PRISM 3100 Genetic Analyzer - Applied Biosystems) (dados não apresentados).

**Figura 15.** Eletroforese em gel de agarose mostrando a liberação da sequência  $\Delta hns:Can^R$  a partir do vetor pPR27:*xyIE* pela clivagem da enzima SpeI.



**Fonte:** Produção do próprio autor do trabalho.

**Legenda:** M = marcador de peso molecular de DNA 1 kb Plus. C = controle não clivado. 1 = sequência pPR27:*xyIE*: $\Delta hns:Can^R$  clivada com a enzima SpeI.

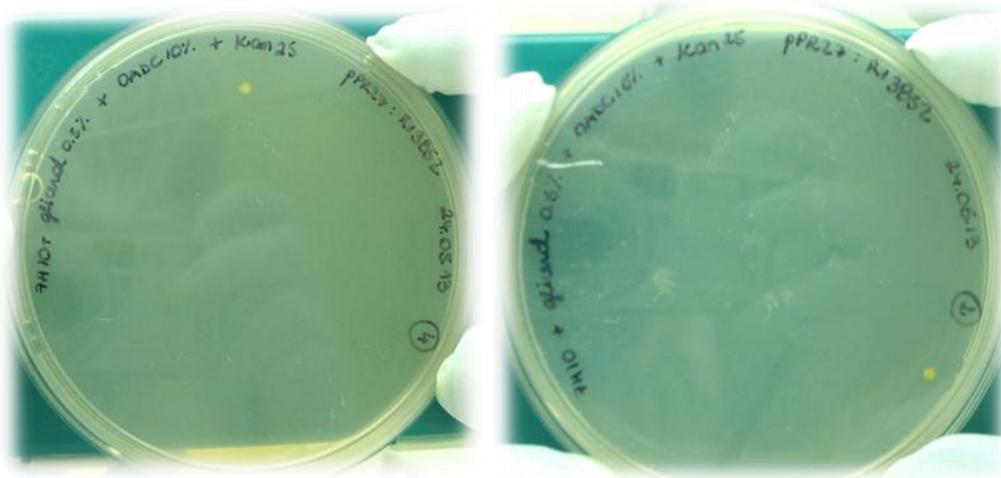
**Nota:** a sequência  $\Delta hns:Can^R$  liberada do pPR27:*xyIE* contém 3.199 pb.

Os próximos passos foram realizados em um laboratório de segurança biológica nível 3 (NB3). O vetor recombinante pPR27:*xyIE*: $\Delta hns:Can^R$  foi inserido em células de *M. tuberculosis* H37Rv eletrocompetentes por meio de eletroporação (Gene Pulser Xcell Microbial System – Bio-Rad). As células transformadas foram rapidamente recuperadas em 5 mL do meio de cultura líquido Middlebrook 7H9 (Sigma-Aldrich), tween-80 0,05 %, OADC 10 % (suplemento contendo ácido oleico, albumina, dextrose e catalase), e incubadas a 32 °C por 48 h. Após este período, as células foram semeadas em meio de cultura sólido Middlebrook 7H10 (Sigma-Aldrich), OADC 10 %, canamicina 25  $\mu\text{g mL}^{-1}$  e incubadas a 32 °C por seis semanas.

Duas colônias apresentaram resistência à canamicina ( $Can^R$ ), e estas foram testadas para a presença do gene *xyIE* por meio da adição de uma solução com 1 % de catecol [Figura 16]. Ambas as colônias apresentaram a coloração amarela ( $Can^R$  e  $xyIE^+$ ) após a adição do catecol, e portanto, foram inoculadas em 5 mL de meio de cultura líquido Middlebrook 7H9, tween-80 0,05 %, OADC 10 %, canamicina 25  $\mu\text{g mL}^{-1}$ , e incubadas a 32 °C sem agitação por seis semanas. Após este período, as culturas foram diluídas (0,  $10^{-1}$ ,  $10^{-2}$ ,  $10^{-3}$ ,  $10^{-4}$ ) utilizando meio de cultura, e 200  $\mu\text{L}$  de cada diluição foi semeado em Middlebrook 7H10, OADC 10 %, canamicina 25  $\mu\text{g mL}^{-1}$ , sacarose 2 %. As placas foram incubadas a 39 °C por três semanas e após o crescimento das colônias, estas foram testadas com 1 % de catecol para verificar a presença do gene *xyIE*. Desta vez, as colônias que apresentassem a coloração branca ( $Can^R$ ,  $xyIE^-$ ,  $Sac^R$ ) indicariam que o evento de recombinação homóloga ocorreu de forma satisfató-

ria e que o corpo do plasmídeo foi perdido: nove colônias brancas e 26 pequenas colônias amarelas foram observadas na placa sem diluição, enquanto que 20 colônias brancas foram observadas na placa com diluição  $10^{-3}$  [Figura 17]. Dez colônias brancas foram então selecionadas e inoculadas em 5 mL de meio de cultura líquido Middlebrook 7H9, tween-80 0,05 %, OADC 10 %, canamicina  $25 \mu\text{g mL}^{-1}$ , a  $37^\circ\text{C}$  sob agitação por três semanas. Por fim, soluções estoque das culturas foram preparadas em 25 % de glicerina e armazenadas a  $-20^\circ\text{C}$  e do volume restante da amostra extraiu-se o DNA genômico.

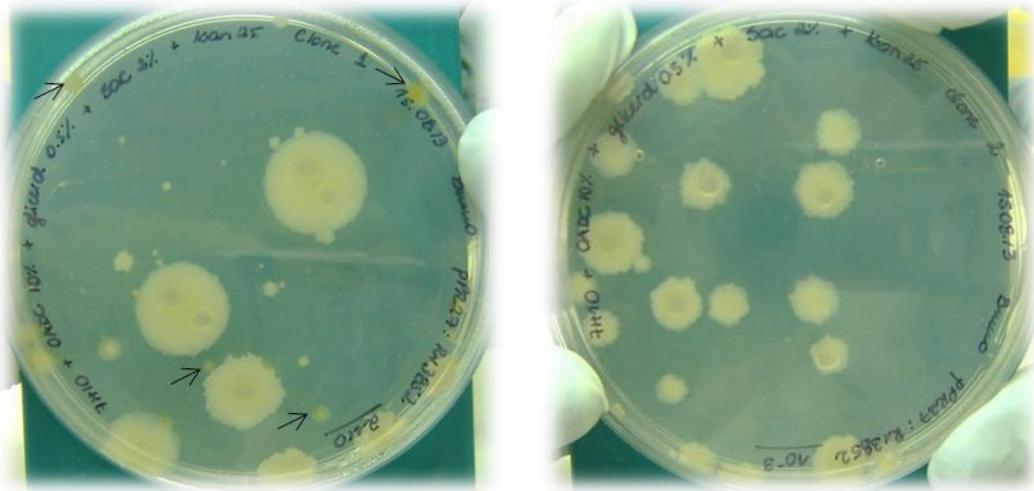
**Figura 16.** Duas colônias amarelas de *M. tuberculosis* transformadas com o vetor pPR27:*xylE*: $\Delta$ *hns*:Can<sup>R</sup> após serem testadas com catecol.



**Fonte:** Produção do próprio autor do trabalho.

**Nota:** colônias cultivadas em Middlebrook 7H10, OADC 10 %, canamicina  $25 \mu\text{g mL}^{-1}$ , a  $32^\circ\text{C}$  por seis semanas. Nesta etapa, as colônias que apresentassem a coloração amarela após serem testadas com 1 % de catecol indicariam a presença do plasmídeo pPR27:*xylE*: $\Delta$ *hns*:Can<sup>R</sup> dentro da micobactéria. Somente duas colônias amarelas foram observadas, as quais foram inoculadas em 5 mL de meio de cultura líquido Middlebrook 7H9, tween-80 0,05 %, OADC 10 %, canamicina  $25 \mu\text{g mL}^{-1}$ , e incubadas a  $32^\circ\text{C}$  sem agitação por seis semanas.

**Figura 17.** Colônias brancas de *M. tuberculosis* após serem testadas com 1% catecol.



**Fonte:** Produção do próprio autor do trabalho.

**Legenda:** as setas indicam a presença de pequenas colônias amarelas após a adição de catecol.

**Nota:** colônias cultivadas em Middlebrook 7H10, OADC 10 %, canamicina  $25 \mu\text{g mL}^{-1}$ , sacarose 2 %, à  $39^\circ\text{C}$  por três semanas. Nesta etapa, as colônias que apresentassem a coloração branca após serem testadas com 1 % de catecol indicariam que houve a integração da sequência  $\Delta hns:Can^R$  ao genoma do *M. tuberculosis*. Portanto, 10 colônias brancas foram selecionadas para crescer em meio líquido Middlebrook 7H9, tween-80 0,05 %, OADC 10 %, canamicina  $25 \mu\text{g mL}^{-1}$ , a  $37^\circ\text{C}$  sob agitação por três semanas. Também foi possível observar a presença de 26 pequenas colônias amarelas (algumas estão indicadas por setas) na placa sem diluição, as quais provavelmente mantiveram o plasmídeo íntegro e não sofreram o processo de recombinação homóloga.

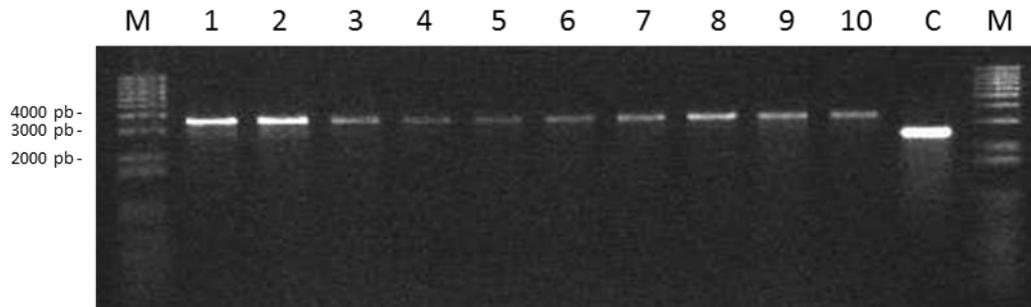
Para a extração do DNA, cada cultura foi centrifugada a 3.000 rpm por 5 min e o *pellet* resultante foi recuperado com 250  $\mu\text{L}$  de uma solução I contendo Tris-HCl 50 mM, sacarose 25 %, EDTA 50 mM, lisozima  $500 \mu\text{g mL}^{-1}$ , pH 8.0. As células recuperadas foram incubadas a  $37^\circ\text{C}$  por 16 h sob agitação em *shaker*. No dia seguinte, adicionou-se às culturas 250  $\mu\text{L}$  de uma solução II contendo Tris-HCl 100 mM, SDS 1 %, proteinase K  $400 \mu\text{g mL}^{-1}$  (GIBCO), pH 8.0, e estas ficaram incubadas por 16 h a  $37^\circ\text{C}$  sob agitação em *shaker*. No terceiro dia, às soluções foi adicionado fenol/clorofórmio/álcool isoamílico (25:24:1 – Sigma Aldrich), e estas ficaram sob agitação à temperatura ambiente por 16 h. No último dia de extração, as amostras contendo as soluções de cada cultura foram centrifugadas a 12.000 rpm. Ao sobrenadante foi adicionado, novamente, 500  $\mu\text{L}$  de fenol/clorofórmio/álcool isoamílico. Os tubos foram agitados em vórtex por 30 segundos e centrifugados a 12.000 rpm. O sobrenadante foi transferido para um novo tubo, no qual foi adicionado cloreto de sódio 3 M e etanol 100 % (Merck). Os tubos ficaram então sob agitação em gangorra por 20 min para que o DNA fosse precipitado e, subsequentemente, foram centrifugados a 12.000 rpm por 30 min à temperatura ambiente. O sobrenadante de cada tubo foi cuidadosamente removido com pipeta e, por fim, o DNA genômico extraído foi recuperado com 60  $\mu\text{L}$  de tampão EB.

Para confirmar a mutagênese por troca alélica do gene *hns*, reações em cadeia da polimerase foram realizadas para as 10 colônias selecionadas. Para isso, dois oligonucleotídeos iniciadores (5'-TAGTCCACGAGATACCGGGA-3' e 5'-GTGACCGACACCATTGAACG-3') foram desenhados para anelarem *upstream* e *downstream* às sequências clonadas no vetor pPR27*xyIE*. O DNA genômico da cepa H37Rv de *M. tuberculosis* também foi utilizado, com o intuito de se ter um controle do experimento. Existem três eventos plausíveis que podem acontecer durante o processo do nocaute gênico (Parish, 2001):

- (1) Duplo *cross-over* (DCO): um único fragmento de 3.473 pb seria observado no gel, que consiste na sequência  $\Delta hns:Can^R$ ; este resultado indicaria que houve a troca do alelo selvagem para o alelo mutante e que o processo de mutagênese por troca alélica foi realizado com sucesso;
- (2) Simples *cross-over* (SCO): caso ocorresse um único evento de *cross-over*, o fragmento resultante seria grande o suficiente para não ser amplificado pela PCR (aproximadamente 15 kb), devido à inserção completa do vetor pPR27:*xyIE* juntamente com a sequência  $\Delta hns:Can^R$  ao locus do gene-alvo;
- (3) Recombinação ilegítima: um único fragmento de 2.554 pb seria observado no gel, correspondente à sequência selvagem do gene *hns* e às regiões flangeadoras *upstream* e *downstream*, uma vez que a integração completa do pPR27:*xyIE*:  $\Delta hns:Can^R$  ocorreu de forma não homóloga em outra parte do genoma micobacteriano;

O resultado das PCRs foi posteriormente analisado em gel de agarose 1 %. Conforme mostrado na figura 18, todos os clones apresentaram um único fragmento entre 4.000 e 5.000 pb, que corresponde à sequência  $\Delta hns:Can^R$  que sofreu o evento de DCO (3.473 pb). Como esperado, a amplificação a partir do genoma da cepa H37Rv de *M. tuberculosis* gerou um fragmento entre 3.000 e 2.000 pb, correspondente à sequência selvagem do gene *hns* e das regiões flangeadoras *upstream* e *downstream* (2.554 pb). De acordo com este resultado, portanto, o gene *hns* não é essencial para a sobrevivência do *M. tuberculosis* H37Rv nas condições utilizadas neste experimento, corroborando os dados obtidos por Griffin e colaboradores (Griffin, 2011). Além disso, pode-se confirmar a obtenção da cepa de *M. tuberculosis* H37Rv nocaute para o gene *hns*, a qual foi nomeada  $\Delta hns$  para futuros estudos.

**Figura 18.** Eletroforese em gel de agarose do produto de PCR dos 10 clones selecionados de *M. tuberculosis* que sofreram o processo de recombinação homóloga.

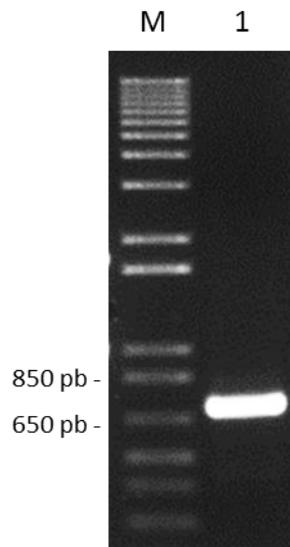


**Fonte:** Produção do próprio autor do trabalho.

**Legenda:** M = marcador de peso molecular de DNA 1 kb Plus. 1-10 = clones  $\Delta hns:Can^R$  (3.476 pb). C = controle do experimento, que corresponde à sequência selvagem do gene *hns* de *M. tuberculosis* H37Rv com as regiões *flanqueadoras upstream* e *downstream* (2.554 pb).

Além da construção da cepa *M. tuberculosis* H37Rv  $\Delta hns$ , uma segunda cepa foi gerada, a qual foi complementada com uma cópia do gene *hns* selvagem. Esta cópia adicional foi integrada em outra região do genoma da micobactéria, antes das células sofrerem o processo de mutagênese por troca alélica para a inativar o gene *hns* selvagem. Primeiramente foi realizada uma PCR com a finalidade de amplificar uma sequência de 687 pb correspondente ao gene *hns* selvagem (405 pb), e às regiões *upstream* e *downstream* ao gene (282 pb ao total). Dois oligonucleotídeos iniciadores foram utilizados para essa reação, ambos contendo sítios de restrição para a enzima *Xba*I (em destaque): 5'-TTTTTCTAGAGACGTGGCCATTCGTCGTA-3' e 5'-TTTTTCTAGATTGGCGGAAGTGTAGGTAC-3'. A reação ocorreu na presença de DMSO 5 % e foi realizada pela enzima *Pfu* DNA Polimerase (4G P&D). A confirmação da amplificação foi realizada em gel de agarose 1 % corado em banho de GelRed™ [Figura 19].

**Figura 19.** Eletroforese em gel de agarose do produto de PCR do gene *hns* selvagem flanqueado pelas regiões *upstream* e *downstream*.



**Fonte:** Produção do próprio autor do trabalho.

**Legenda:** M = marcador de peso molecular de DNA 1 kb Plus. 1 = amplicon de 707 pb.

**Nota:** O amplicon observado no gel possui 707 pb e corresponde ao gene *hns* selvagem (450 pb), às duas sequências flanqueadoras (282 pb para cada sequência) e às sequências extras presentes nos oligonucleotídeos iniciadores (20 pb).

O próximo passo consistiu na clivagem do amplicon e do vetor integrativo pNIP40/b (7.290 pb) derivado do micobacteriófago Ms6, com a enzima XbaI. Vetores derivados de Ms6 integram-se eficientemente na extremidade 3' dos genes que codificam tRNAs<sup>Ala</sup> (sítio *attB*) tanto nas micobactérias de crescimento lento (*M. bovis* BCG e *M. tuberculosis*), quanto nas de crescimento rápido (*M. smegmatis*) (Dos Vultos, 2006; Freitas-Vieira, 1998). O vetor clivado pela enzima foi, em seguida, desfosforilado durante uma hora a 37 °C pela enzima CIP (*alkaline phosphatase, calf intestinal* – New England Biolabs), a fim de impedir a religação das suas extremidades. Após este processo, ambos os fragmentos – amplicon clivado e vetor aberto e desfosforilado – foram ligados pela enzima DNA Ligase (New England Biolabs).

As próximas etapas foram realizadas no laboratório NB3. A inserção do vetor pNIP40/b:*hns* às células eletrocompetentes de *M. tuberculosis* H37Rv ocorreu da mesma forma como descrito para o vetor pPR27:*xyIE*; porém, o antibiótico utilizado para selecionar as colônias transformadas foi a higromicina 50 µg mL<sup>-1</sup>. Após três semanas de crescimento a 37 °C foi possível observar o crescimento de colônias resistentes à higromicina, as quais sofreram o evento de integração do plasmídeo no genoma da micobactéria. Três colônias (clones 1, 2 e 3) foram selecionadas e inoculadas em 5 mL de meio de cultura líquido Middlebrook 7H9, tween-80 0,05 %, OADC 10 %, higromicina 50 µg mL<sup>-1</sup>, e incubadas a 37 °C sob

agitação por três semanas. Após este crescimento inicial, 250  $\mu\text{L}$  do inóculo foi transferido para um segundo frasco contendo 50 mL de meio de cultura líquido Middlebrook 7H9, tween-80 0,05 %, OADC 10 %, higromicina 50  $\mu\text{g mL}^{-1}$ , e incubado a 37 °C sob agitação por três semanas. Células eletrocompetentes foram preparadas a partir deste inóculo, as quais foram, então, transformadas com o vetor pPR27:*xylE*: $\Delta$ *hns*:Can<sup>R</sup>. As colônias mutantes foram novamente selecionadas por meio dos marcadores seletivos e contra-seletivos descritos anteriormente, porém, dois antibióticos foram utilizados nos meios de cultura nesta etapa: canamicina 25  $\mu\text{g mL}^{-1}$  e higromicina 50  $\mu\text{g mL}^{-1}$ . Observou-se o crescimento de colônias dos três clones selecionados em todas as placas (sem diluição, 10<sup>-1</sup>, 10<sup>-2</sup>, 10<sup>-3</sup>, 10<sup>-4</sup>). A placa com a diluição 10<sup>-4</sup> foi então examinada com catecol adicionando-se uma gota do composto: nove colônias amarelas, contendo o vetor pPR27:*xylE*: $\Delta$ *hns*:Can<sup>R</sup>, foram observadas [Figura 20].

**Figura 20.** Colônias amarelas da cepa complementada de *M. tuberculosis* contendo o plasmídeo pPR27:*xylE*: $\Delta$ *hns*:Can<sup>R</sup>.



**Fonte:** Produção do próprio autor do trabalho.

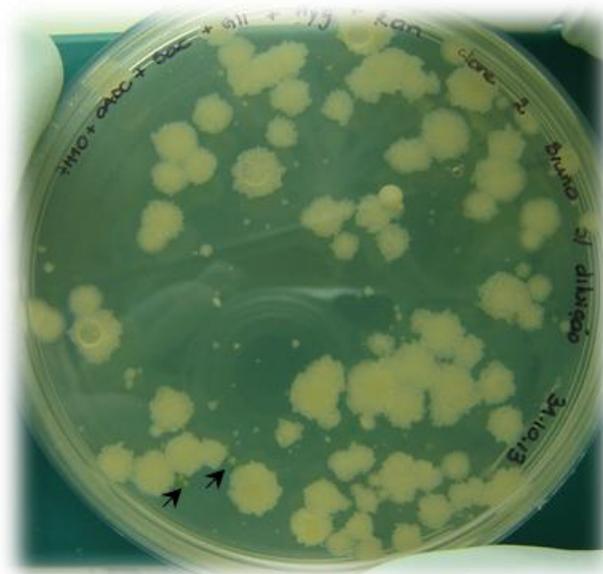
**Legenda:** As setas apontam para as colônias de coloração amarela que indicam a presença do gene *xylE*.

**Nota:** Colônias cultivadas em Middlebrook 7H10, tween-80 0,05 %, OADC 10 %, canamicina 25  $\mu\text{g mL}^{-1}$ , higromicina 50  $\mu\text{g mL}^{-1}$ , a 37 °C por três semanas. Nesta etapa, as colônias que apresentassem a coloração amarela após serem testadas com 1 % de catecol indicariam a presença do plasmídeo pPR27:*xylE*: $\Delta$ *hns*:Can<sup>R</sup>.

Quatro colônias amarelas (clones 1, 2, 3 e 4) foram selecionadas e inoculadas em 5 ml de meio de cultura líquido Middlebrook 7H9, tween-80 0,05 %, OADC 10 %, higromicina 50  $\mu\text{g mL}^{-1}$ , e incubadas a 32 °C sob agitação por seis semanas. Após este período, a cultura foi semeada em placas de meio de cultura sólido Middlebrook 7H10, OADC 10 %, canamicina

25  $\mu\text{g mL}^{-1}$ , higromicina 50  $\mu\text{g mL}^{-1}$ , sacarose 2 %. As placas foram incubadas a 39 °C por três semanas. Diversas colônias foram observadas nas placas (sem diluição,  $10^{-1}$ ,  $10^{-2}$ ,  $10^{-3}$ ,  $10^{-4}$ ), as quais foram então testadas adicionando-se uma gota de catecol. As maiores colônias observadas nas placas mantiveram a coloração branca na presença do catecol, ou seja, estas colônias provavelmente sofreram o processo de recombinação homóloga e perderam o corpo do plasmídeo pPR27xy/E; as colônias menores apresentaram a coloração amarela, indicando que não sofreram o DCO [Figura 21].

**Figura 21.** Colônias brancas de *M. tuberculosis* mutantes para o gene *hns* e complementadas com uma cópia extra do gene que foi integrada em outra região do genoma micobacteriano.



**Fonte:** Produção do próprio autor do trabalho.

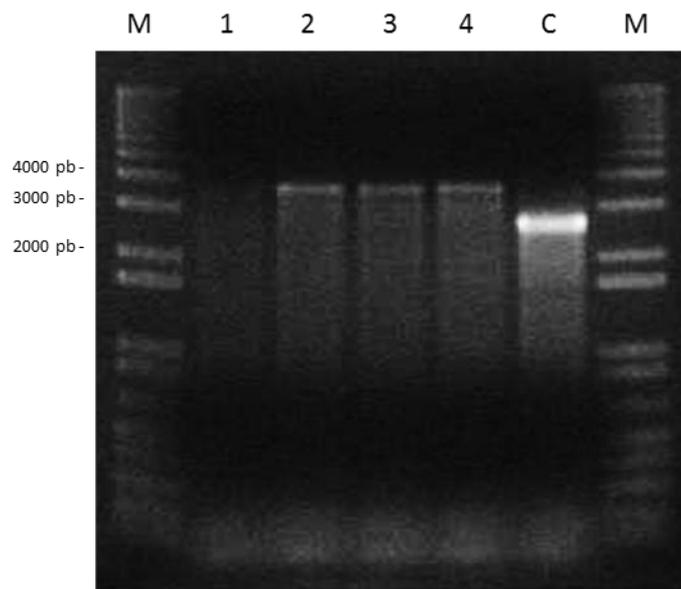
**Legenda:** as setas apontam para colônias amarelas pequenas que não sofreram o processo DCO.

**Nota:** Colônias cultivadas em Middlebrook 7H10, OADC 10 %, canamicina 25  $\mu\text{g mL}^{-1}$ , higromicina 50  $\mu\text{g mL}^{-1}$ , sacarose 2 %, a 37 °C por três semanas. Nesta etapa, as colônias que apresentassem a coloração branca após serem testadas com catecol indicariam que houve o processo de mutagênese por troca alélica de forma satisfatória.

Uma colônia de cada clone foi selecionada para crescer em 5 mL de meio de cultura líquido Middlebrook 7H9, tween-80 0,05 %, OADC 10 %, higromicina 50  $\mu\text{g mL}^{-1}$ , a 37 °C sob agitação por três semanas. Por fim, soluções estoque das culturas foram preparadas em 25 % de glicerina e armazenadas a -20 °C e do volume restante da amostra extraiu-se o DNA genômico. O processo de extração de DNA genômico foi realizado conforme descrito anteriormente.

O DNA genômico dos quatro clones selecionados foi utilizado em PCR para amplificar a sequência  $\Delta hns:Can^R$ , utilizando-se dois oligonucleotídeos iniciadores (5'-TAGTCCACGAGATACCGGGA-3' e 5'-GTGACCGACACCATTGAACG-3'); os amplicons foram então analisados em gel de agarose 1 %. Conforme mostrado na figura 22, três clones (2, 3 e 4) apresentaram um único fragmento entre 4.000 e 5.000 pb, que corresponde à sequência  $\Delta hns:Can^R$  que sofreu o evento de DCO (3.473 pb). Com base neste resultado, pode-se confirmar a obtenção da cepa de *M. tuberculosis* H37Rv mutante para o gene *hns*, porém complementada com uma cópia extra do mesmo gene selvagem que foi integrada no sítio *attB* do genoma da micobactéria; esta cepa foi nomeada  $\Delta hnsC$  para futuros estudos.

**Figura 22.** Eletroforese em gel de agarose do produto de PCR dos quatro clones selecionados de *M. tuberculosis*, complementado com a cópia extra do gene *hns*, e que sofreram o processo de recombinação homóloga.



**Fonte:** Produção do próprio autor do trabalho.

**Legenda:** M = marcador de peso molecular de DNA 1 kb Plus. 1-4 = clones  $\Delta hnsC$  (3.476 pb). C = controle do experimento, que corresponde à sequência selvagem do gene *hns* de *M. tuberculosis* H37Rv com as regiões flangeadoras *upstream* e *downstream* (2.554 pb).

**Nota:** somente no clone 1 não houve amplificação.

A obtenção da cepa  $\Delta hns$  de *M. tuberculosis* é o primeiro passo de um conjunto de experimentos que visam à caracterização funcional da proteína MDP2 *in vivo*, entre eles: (1) a infecção de macrófagos e camundongos para avaliar a capacidade de infecção da cepa mutante em relação à cepa selvagem; (2) taxa de crescimento bacteriano; e também (3) o modelo de latência de Wayne (Wayne, 1996; Patel, 2011). A cepa  $\Delta hnsC$ , por sua vez, será utilizada como controle da técnica de mutagênese por troca alélica, ou seja, esta cepa deve se

comportar da mesma forma que a cepa selvagem de *M. tuberculosis* devido à presença da cópia extra do gene *hns*. Caso seja observada alguma alteração na cepa  $\Delta hnsC$  em relação à cepa selvagem, é possível que o segundo gene presente no mesmo operon do gene *hns* – chamado de *rraA* - tenha sofrido alguma alteração na sua expressão, ocasionada pelo processo de mutagênese por troca alélica; ou que a cópia extra do gene *hns* não seja funcional.

Por meio da obtenção da cepa  $\Delta hns$ , existe a possibilidade de se construir uma cepa atenuada de *M. tuberculosis*, a qual poderia ser utilizada para o desenvolvimento de novos candidatos a vacinas que sejam mais efetivas na imunização de pessoas do que a atual vacina BCG empregada pelos sistemas de saúde.

### Testes de pré-cristalização da MDP2 recombinante

Atualmente existem quase 96.000 estruturas de proteínas depositadas na base de dados *Protein Data Bank* (PDB) (acesso em 06 de fevereiro de 2014), sendo que a maioria destas estruturas foi resolvida pela técnica de cristalografia de raios-X (Berman, 2000; Gall, 2007). Apesar desta enorme quantidade de depósitos no banco de dados, as informações contidas no PDB são, de certa forma, tendenciosas, uma vez que estas favorecem positivamente somente àquelas proteínas que são passíveis de expressão, purificação e cristalização, enquanto que desfavorece proteínas que resistem aos protocolos típicos de cristalização, como as IDPs e as proteínas de membrana (Gall, 2007). Curiosamente, estruturas completas de proteínas ordenadas depositadas no PDB não são abundantes, uma vez que muitas proteínas possuem regiões de desordem na sua sequência, as quais são conhecidas pelo seu comportamento dinâmico que leva à variação da posição dos átomos da proteína e dos ângulos de Ramachandran (Gall, 2007). Como consequência, as IDRs geralmente estão ausentes nos mapas de densidade eletrônica dos experimentos de cristalografia por raios-X, não sendo representadas, portanto, nas estruturas depositadas no PDB (Radivojac, 2003; Gall, 2007). Para facilitar o acesso aos estudos estruturais acerca das IDPs, a base de dados do DisProt foi criada com o intuito de facilitar a gestão e a anotação das informações de novas IDPs, fazendo uma ligação entre a estrutura resolvida e a função da proteína (Sickmeier, 2007). Atualmente existem 694 proteínas e 1.539 regiões desordenadas depositadas no DisProt (acesso em 06 de fevereiro de 2014).

Visto que este trabalho se propôs a realizar a determinação da estrutura cristalográfica da proteína MDP2 de *M. tuberculosis*, um teste de pré-cristalização (PCT™, Hampton Research) foi realizado inicialmente. Este teste tem como objetivo determinar a concentração de proteína apropriada para o *screening* de cristalização por meio de *kits*, como o Crystal Screen™ e Crystal Screen 2™ (Hampton Research), a fim de minimizar e/ou prevenir situações em que um *screen* resulta em uma abundância de precipitados ou soluções translúcidas. Segundo o manual do PCT™, amostras muito concentradas resultam em precipitados amorfos, enquanto que amostras muito diluídas levam a soluções translúcidas; em ambos os casos pode não haver a formação de cristais, impossibilitando a resolução da estrutura da proteína. Após a definição das condições de pré-cristalização e de cristalização da MDP2, a

resolução da estrutura cristalográfica da proteína poderia então ser resolvida pela técnica de raios-X.

Para a realização do PCT™, utilizou-se uma placa VDX de cristalização com 16 poços, nos quais se adicionou 700 µL de cada um dos tampões do *kit* (A1/A2/B1/B2). Em quatro lamínulas separadas adicionou-se 1 µL de MDP2 homogênea (1,4 mg mL<sup>-1</sup> em tampão Hepes 50 mM, NaCl 100 mM pH 7.9), e em cada gota de MDP2 adicionou-se 1 µL de um dos quatro tampões diferentes presentes nos poços. As lamínulas foram então invertidas para deixarem as gotas viradas para a solução de cada poço e seladas com graxa. A reação ocorreu por 30 min, e após este período, cada gota foi analisada com o auxílio de um microscópio óptico. Cada reação foi feita em duplicata.

Utilizando-se como base os resultados apresentados na Tabela 1 do manual do PCT™, a reação da MDP2 com cada reagente resultou em soluções translúcidas, sem qualquer indício de formação de precipitados, mostrando, em um primeiro momento, que a amostra está muito diluída (dados não apresentados). A recomendação do manual para este tipo de situação seria concentrar a amostra para a metade do seu volume original. No entanto, baseando-se nos resultados apresentados no capítulo 2 em relação à MDP2, já era esperado que a proteína não sofresse o processo de cristalização de forma satisfatória, uma vez que as IDPs são difíceis de cristalizar por não assumirem uma estrutura 3-D bem definida; ou, se assumissem, poderiam apresentar conformações não representativas das regiões de desordem e da estrutura quaternária da proteína como um todo (Fong, 2009; Gall, 2007; Tompa, 2009). Estudos envolvendo a proteína básica mielina (MBP), por exemplo, relataram que após 4.600 diferentes condições de cristalização, nenhuma delas resultou na formação de cristais; isso se deve à adoção de uma conformação “*random coil*” por parte da proteína fora do seu ambiente nativo (Tompa, 2009).

Como perspectiva para este experimento, propõe-se que sejam realizados novos testes de pré-cristalização com amostras mais concentradas de proteína, a fim de se verificar se a ausência de cristais se deve à baixa concentração da amostra, e também que sejam feitas tentativas de co-cristalização da MDP2 na presença de DNA dupla-fita, visto que há a possibilidade de a proteína assumir uma estrutura ordenada na presença do seu parceiro de interação (Bukowska, 2013).

## **CAPÍTULO 4**

### **CONSIDERAÇÕES FINAIS**

Estudos envolvendo IDPs têm aumentado consideravelmente nas últimas décadas. Muitos pesquisadores têm se voltado para estas proteínas principalmente (1) pelas características incomuns que estas apresentam em condições fisiológicas, que desafiam até mesmo velhos paradigmas encontrados em livros-texto de biologia básica e bioquímica; (2) pelos seus diversos papéis desempenhados dentro das células; e (3) pela sua importância em algumas doenças humanas (câncer, doenças neurodegenerativas, doenças cardiovasculares, diabetes) e para o processo de patogênese de alguns micro-organismos de importância médica. Diante destes fatos, este presente estudo se propôs primeiramente a validar a proteína MDP2 de *M. tuberculosis* como uma IDP.

Como descrito no capítulo 2, uma análise *in silico* da sequência de aminoácidos da MDP2 foi realizada pelo preditor PONDR®. O resultado da predição mostrou que 86 % da sequência da MDP2 é desordenada, e que a proteína possui três regiões de desordem bem definidas: uma região longa de desordem (LDR, do inglês *long disordered region*) na porção N-terminal contendo 76 resíduos; uma segunda região na porção mediana da proteína contendo 31 resíduos; e uma terceira região na porção C-terminal com apenas 8 resíduos. O servidor PONDR® também permitiu utilizar duas ferramentas que discriminam proteínas ordenadas de desordenadas: a função de distribuição cumulativa (CDF) e o gráfico C-H (do inglês, *charge-hydrophathy plot*). A ferramenta CDF indicou que a MDP2 tem uma característica esperada para proteínas completamente desordenadas, enquanto que o gráfico C-H não apresentou um resultado conclusivo: a proteína permaneceu na linha entre desordem e ordem. Para corroborar os resultados obtidos *in silico*, diversas técnicas experimentais foram utilizadas para acessar o comportamento da MDP2 em determinadas condições *in vitro*.

Técnicas indiretas como mobilidade eletroforética em SDS-PAGE, estabilidade ao calor e a desnaturantes químicos, perfil de eluição por gel filtração e sensibilidade proteolítica, fornecem indícios de que uma determinada proteína possa conter regiões de desordem na sua estrutura. Como mostrado pela técnica de mobilidade eletroforética em SDS-PAGE, a MDP2 apresentou um retardo na sua mobilidade, aparentando uma massa molecular de aproximadamente 19 kDa, enquanto que sua massa molecular teórica calculada é de 13,8 kDa. É esperado que as IDPs se liguem menos ao SDS devido a sua composição incomum de aminoácidos, geralmente apresentando uma massa molecular 1,2-1,8 vezes maior do que o calculado. Para confirmar que a banda visualizada em gel correspondia de fato à proteína MDP2 de *M. tuberculosis*, a técnica de tripsinização em gel foi realizada. Os peptídeos gerados pela

digestão com tripsina foram identificados por nanocromatografia líquida acoplada à espectrometria de massa, confirmando a identidade da banda. A massa intacta da proteína também foi determinada por espectrometria de massa (13,8 kDa), mostrando que a proteína não sofreu nenhuma alteração durante o processo de purificação.

Outra característica bastante marcante das IDPs é a sua resistência a altas temperaturas e a desnaturantes químicos, que se deve ao fato destas proteínas já possuírem segmentos expostos na sua estrutura com uma alta prevalência de resíduos carregados. Para observar este comportamento, a MDP2 foi submetida a 100 °C por 10 min, seguida por uma centrifugação e, por fim o sobrenadante foi analisado em SDS-PAGE. O resultado deste experimento mostrou que a MDP2 não sofreu um processo de agregação e precipitação induzido pela temperatura, diferentemente do que aconteceu com as proteínas ordenadas utilizadas como controle positivo (CDA, GMPR, PNP e EPSPS). Para visualizar o comportamento da proteína diante de um desnaturante químico, o ácido tricloroacético foi utilizado em quatro concentrações diferentes: 0, 1, 3 e 5 %. Como verificado em SDS-PAGE, a MDP2 permaneceu na fração solúvel em todas as concentrações testadas, enquanto que as proteínas ordenadas rapidamente sofreram o processo de precipitação. Ambos os resultados dão fortes indícios de que a proteína se comporta como uma IDP sob estas condições.

A última técnica indireta realizada neste estudo consistiu no teste de sensibilidade proteolítica. Para isso, a MDP2 sofreu o processo de proteólise limitada na presença de proteinase K em diferentes razões. A reação ocorreu por somente cinco minutos a 25 °C. A análise em SDS-PAGE mostrou que a proteína é mais sensível à ação da proteinase K se comparada às proteínas ordenadas utilizadas como controle (BSA e CDA); porém, comparando-se este resultado com os de outras proteínas da literatura, a proteína não apresenta uma alta sensibilidade à proteólise (razões acima de 1:1000). Além disso, foi possível observar a formação de bandas intermediárias no gel, as quais podem ser provenientes de sequências resistentes à clivagem. Existem pelo menos duas explicações possíveis para este resultado. Primeiro, a proteína pode conter segmentos que não estão expostos ou que não são tão flexíveis como o esperado para uma IDP. Segundo, alguns segmentos da MDP2 podem não estar acessíveis à ação da protease devido ao processo de auto agregação realizado pela proteína em solução. Apesar disso, os resultados obtidos pelas técnicas indiretas formam um conjunto de evidências que indicam que a proteína MDP2 de *M. tuberculosis* de fato se comporta como uma IDP.

Além das técnicas indiretas empregadas neste estudo, duas outras ferramentas foram utilizadas para acessar o estado quaternário da MDP2 em solução. Primeiramente foi realizada uma eluição em gel filtração utilizando uma coluna Superdex 200 e um kit de calibração de massa molecular. Os resultados deste experimento mostraram que a proteína eluiu no volume morto da coluna ( $V_0$ ), não sendo possível estabelecer o estado oligomérico da MDP2. Existem duas explicações possíveis para este resultado: (1) a proteína forma oligômeros/polímeros em solução; e (2) a proteína tem uma forma estendida em comparação às proteínas ordenadas usadas como padrões, não permitindo a sua entrada nos poros da resina. Para tentar responder esta questão, outra técnica foi utilizada: *crosslinking* com glutaraldeído. Os resultados obtidos pelo *crosslinking* mostraram que a proteína rapidamente forma grandes aglomerados que quase impedem a sua entrada no SDS-PAGE. Entretanto, nenhuma banda correspondente a monômeros ou oligômeros foi observada. Juntos, estes dados mostram que a proteína tem uma alta propensão a auto agregação em solução. Estudos futuros poderão mostrar se esta característica permanece na presença de um parceiro de interação da proteína, como o DNA dupla-fita, ou se na presença deste parceiro a proteína assume outras conformações. Aliado aos resultados já obtidos sobre esta proteína pode-se especular que a MDP2 utiliza seu longo fragmento N-terminal desordenado para se ligar ao DNA, uma vez que as sequências de tetrapeptídeos (PAKK, KAAK) encontradas nesta porção, as quais são preditas por realizar a ligação ao DNA, são constituídas por resíduos promotores de desordem. Além disso, é possível que a MDP2 atue como uma proteína promíscua dentro da célula, ligando-se a diferentes sequências do DNA e formando grandes aglomerados de proteína dentro do nucleóide micobacteriano, características que não são pertinentes somente às IDPs, mas também às NAPs em geral.

No capítulo 3 foi apresentada a técnica de mutagênese por troca alélica utilizada no no-caute do gene *hns* de *M. tuberculosis*. Para que o processo de recombinação homóloga ocorresse de forma satisfatória, ou seja, que o alelo selvagem do gene *hns* fosse trocado por uma versão truncada por um cassete que confere resistência a canamicina, um plasmídeo recombinante foi construído, o qual continha uma sequência *upstream* e outra *downstream* ao gene *hns* selvagem. O plasmídeo utilizado neste processo é chamado de pPR27:*xylE*, o qual apresenta algumas características fundamentais para a seleção satisfatória das células mutantes: (1) sua replicação é termosensível; (2) contém um marcador por coloração codificado pelo gene *xylE*; e (3) contém um marcador contra-seletivo codificado pelo gene *sacB*. A con-

firmação do nocaute gênico ocorreu por meio de PCR, a qual mostrou que foi possível obter de forma satisfatória 10 colônias mutantes para o gene *hns*. Com base neste resultado, foi possível afirmar que o gene *hns* de *M. tuberculosis* não é essencial para a sobrevivência do bacilo nas condições testadas, uma vez que a sua remoção não ocasiona a morte da micobactéria.

Neste mesmo capítulo foi apresentado também um teste de pré-cristalização da proteína MDP2 homogênea. Para este experimento, um *kit* PCT™ foi utilizado, o qual contém quatro diferentes tampões que induzem o processo de cristalização. Por meio da combinação dos resultados obtidos em cada tampão, pode-se então estimar se a concentração da proteína é satisfatória para o processo de cristalização. O resultado obtido nas quatro reações mostrou que não houve formação de precipitados, ou seja, que a concentração da proteína estava abaixo do normal, ou que a proteína não sofre o processo de cristalização. Visto que foi demonstrado que MDP2 se comporta como uma IDP, já era esperada que esta proteína não sofresse o processo de cristalização, como acontece com diversas outras IDPs relatadas na literatura. Porém, novas tentativas de pré-cristalização serão realizadas na presença de um parceiro de interação (DNA-dupla fita, por exemplo), visto que há casos de IDPs que assumem uma conformação 3D rígida na presença de uma molécula parceira. A obtenção de cristais desta proteína permitirá a sua análise por cristalografia de raios-X, a qual é a via mais direta para a identificação da ausência de estrutura terciária nas IDPs.

Nós esperamos que este trabalho sobre a proteína MDP2 de *M. tuberculosis* possa contribuir para um melhor entendimento das proteínas intrinsecamente desordenadas e também do metabolismo micobacteriano, uma vez que o *M. tuberculosis* ainda é uma grave ameaça global. No entanto, mais estudos deverão ser realizados a fim de se entender como a proteína MDP2 usa sua desordem intrínseca para interagir com o DNA e para verificar se ela alterna entre os estados “desordenado-ordenado” na presença do seu parceiro de interação. Também será importante identificar o tipo de agregado formado pela MDP2 em solução e entender como a proteína forma estas grandes estruturas dentro da célula.

## REFERÊNCIAS

- ALBERTS, B. *et al.* **Biologia Molecular da Célula**. 5. Porto Alegre: Artmed, 2010. 1396.
- ANUCHIN, A. M. *et al.* Histone-like proteins of bacteria (review). **Applied Biochemistry and Microbiology**, v. 47, n. 6, p. 580-585, 2011. Disponível em: < <http://www.springerlink.com/index/10.1134/S0003683811060020> >.
- ANURAG, M., AND DASH, D. Unraveling the potential of intrinsically disordered proteins as drug targets: application to *Mycobacterium tuberculosis*. **Mol BioSyst.**, v. 5, p. 1752-1757, 2009. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19763328> >.
- BERMAN, H. M.; *et al.* The Protein Data Bank. **Nucleic Acids Research**, v. 28, n. 1, p. 235-242, 2000. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/10592235> >.
- BETTS, J. C. *et al.* Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. **Mol Microbiol**, v. 43, n. 3, p. 717-31, 2002.
- BIOSCIENCES, A. **Protein Purification: Handbook** 2001.
- BRADLEY, M. D. *et al.* Effects of Fis on Escherichia coli gene expression during different growth stages. **Microbiology**, England, v. 153, n. 9, p. 2922-40, 2007. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/> >.
- BRENNAN, P. J. Structure, function, and biogenesis of the cell wall of *Mycobacterium tuberculosis*. **Tuberculosis (Edinb)**, England, v. 83, n. 1-3, p. 91-7, 2003. Disponível em: < <http://linkinghub.elsevier.com/retrieve/pii/S1472979202000896> >.
- BRESNICK, E. H. *et al.* The transcriptionally-active MMTV promoter is depleted of histone H1. **Nucleic Acids Research**, v. 20, n. 2, p. 273-278, 1992. Disponível em: < <http://nar.oxfordjournals.org/content/20/2/273.abstract> >.
- BROWNING, D. F.; GRAINGER, D. C.; BUSBY, S. J. Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. **Current opinion in microbiology**, v. 13, n. 6, p. 773-80, 2010. Disponível em: < [http://linkinghub.elsevier.com/retrieve/pii/S1369-5274\(10\)00141-4](http://linkinghub.elsevier.com/retrieve/pii/S1369-5274(10)00141-4) >.
- BUKOWSKA, M. A.; GRÜTTER, M. G. New concepts and aids to facilitate crystallization. **Current Opinion in Structural Biology**, v. 23, p. 409-416, 2013. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/23578532> >.
- CANTÚ, M. D. *et al.* Peptide sequencing using mass spectrometry: a practical guide. **Química Nova**, v. 31, n. 3, 2008. ISSN 0100-4042. Disponível em: < [http://www.scielo.br/scielo.php?pid=S0100-40422008000300034&script=sci\\_arttext](http://www.scielo.br/scielo.php?pid=S0100-40422008000300034&script=sci_arttext) >.
- CHOI, J. S. *et al.* Improved cycle sequencing of GC-rich DNA template. **Exp Mol Med**, v. 31, n. 1, p. 20-4, 1999. ISSN 2092-6413. Disponível em: < [http://www.emm.org/search\\_read.htm?page=20&year=1999&vol=31](http://www.emm.org/search_read.htm?page=20&year=1999&vol=31) >.

CHURCHILL, M. E. A.; TRAVERS, A. A. Protein motifs that recognize structural features of DNA. **Trends in Biochemical Sciences**, v. 16, n. 0, p. 92-97, 1991. ISSN 0968-0004. Disponível em: < <http://www.sciencedirect.com/science/article/pii/0968000491900403> >.

CLINE, J.; BRAMAN, J. C.; HOGREFE, H. H. PCR Fidelity of Pfu DNA Polymerase and Other Thermostable DNA Polymerases. **Nucleic Acids Research**, v. 24, n. 18, p. 3546-3551, 1996. Disponível em: < <http://nar.oxfordjournals.org/content/24/18/3546.abstract> >.

COLE, S. T. *et al.* Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. **Nature**, v. 393, n. 6685, p. 537-44, 1998. Disponível em: < <http://www.nature.com/nature/journal/v393/n6685/full/393537a0.html> >.

CRANE-ROBINSON, C. How do linker histones mediate differential gene expression? **BioEssays**, v. 21, n. 5, p. 367-371, 1999. ISSN 1521-1878. Disponível em: < [http://dx.doi.org/10.1002/\(SICI\)1521-1878\(199905\)21:5<367::AID-BIES2>3.0.CO](http://dx.doi.org/10.1002/(SICI)1521-1878(199905)21:5<367::AID-BIES2>3.0.CO) >.

DAME, R. T.; GOOSEN, N. HU: promoting or counteracting DNA compaction? **FEBS letters**, v. 529, n. 2-3, p. 151-6, 2002. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/12372591> >.

DEB, C. *et al.* A Novel In Vitro Multiple-Stress Dormancy Model for *Mycobacterium tuberculosis* Generates a Lipid-Loaded, Drug-Tolerant, Dormant Pathogen. **PLoS ONE**, v. 4, n. 6, p. e6077, 2009. Disponível em: < <http://www.plosone.org/article/info:doi/10.1371/journal.pone.0006077> >.

DELOGU, G.; BRENNAN, M. J. Functional Domains Present in the Mycobacterial Hemagglutinin, HBHA. **Journal of Bacteriology**, v. 181, n. 24, p. 7464-7469, 1999. Disponível em: < <http://jb.asm.org/content/181/24/7464.abstract> >.

DENG, S.; STEIN, R. A.; HIGGINS, N. P. Organization of supercoil domains and their reorganization by transcription. **Mol, Microbiol**, v. 57, n. 6, p. 1511-1521, 2005. Disponível em: < <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1382059/> >.

DHAVAN, G. M. *et al.* Concerted binding and bending of DNA by Escherichia coli integration host factor. **J Mol Biol**, England, v. 315, n. 5, p. 1027-37, 2002. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0022283601953033> >.

DILLON, S. C.; DORMAN, C. J. Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. **Nature reviews. Microbiology**, v. 8, n. 3, p. 185-95, 2010. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/20140026> >.

DORMAN, C. J. H-NS, the genome sentinel. **Nature reviews. Microbiology**, v. 5, n. 2, p. 157-61, 2007. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/17191074> >.

\_\_\_\_\_. Regulation of gene expression by histone-like proteins in bacteria. **Current Opinion in Genetics & Development**, v. 13, n. 2, p. 179-184, 2003. Disponível em: < <http://linkinghub.elsevier.com/retrieve/pii/S0959437X0300025X> >.

DUCATI, R. G. *et al.* Purine Salvage Pathway in *Mycobacterium tuberculosis*. **Curr Med Chem**, Netherlands, v. 18, n. 9, p. 1258-75, 2011. ISSN 1875-533X. Disponível em: < <http://www.benthamdirect.org/pages/content.php?CMC/2011/00000018/00000009/0002C.SGM> >.

\_\_\_\_\_. The resumption of consumption - A review on tuberculosis. **Mem Inst Oswaldo Cruz**, Rio de Janeiro, v. 101, n. 7, p. 697-714, 2006. ISSN 0074-0276. Disponível em: < <http://dx.doi.org/10.1590/S0074-02762006000700001> >.

DUNKER, A. K., *et al.* What's in a name? Why these proteins are intrinsically disordered. **Intrinsically Disordered Proteins**, v. 1, 2013. 10.4161/idp.24157. Disponível em: < <https://www.landesbioscience.com/journals/idp/article/24157/> >.

DYSON, H. J.; WRIGHT, P. E. Intrinsically Unstructured Proteins and Their Functions. **Nature Reviews**, v. 6, 2005. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/15738986> >.

EDWARDS, N. J. Protein identification from tandem mass spectra by database searching. **Methods Mol Biol**, v. 694, p. 119-38, 2011. ISSN 1940-6029. Disponível em: < <http://www.springerlink.com/content/l15j4185385v4274/?MUD=MP> >.

EGAN, E. S.; FOGEL, M. A.; WALDOR, M. K. MicroReview: Divided genomes: negotiating the cell cycle in prokaryotes with multiple chromosomes. **Molecular Microbiology**, v. 56, n. 5, p. 1129-1138, 2005. ISSN 1365-2958. Disponível em: < <http://dx.doi.org/10.1111/j.1365-2958.2005.04622.x> >.

FOLCO, H. D. *et al.* Histone H1 Is Required for Proper Regulation of Pyruvate Decarboxylase Gene Expression in *Neurospora crassa*. **Eukaryotic Cell**, v. 2, n. 2, p. 341-350, 2003. Disponível em: < <http://ec.asm.org/content/2/2/341.abstract> >.

FONG, J. H.; SHOEMAKER, B. A.; GARBUZYNSKIY, S. O.; LOBANOV, M. Y.; GALZITSKAYA, O. V.; PANCHENKO, A. R. Intrinsic Disorder in Protein Interactions: Insights From a Comprehensive Structural Analysis. **PLoS Computational Biology**, v. 5, i. 3, 2009. Disponível em: < <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2646137/> >.

FORMAN-KAY, D. J., AND MITTAG, T. From Sequence and Forces to Structure, Function, and Evolution of Intrinsically Disordered Proteins. **Structure**, v. 21, 2013. 10.1016/j.str.2013.08.001. Disponível em: < [http://www.cell.com/structure/abstract/S0969-2126\(13\)00292-X](http://www.cell.com/structure/abstract/S0969-2126(13)00292-X) >.

FREITAS-VIEIRA, A.; ANES, E.; MONIZ-PEREIRA, J. The site-specific recombination locus of mycobacteriophage Ms6 determines DNA integration at the tRNA<sup>Ala</sup> gene of *Mycobacterium* spp. **Microbiology**, v. 144, p. 3397-3406, 1998. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/9884232> >.

GABRILOVICH, D. I. *et al.* H1<sup>o</sup> histone and differentiation of dendritic cells. A molecular target for tumor-derived factors. **Journal of Leukocyte Biology**, v. 72, n. 2, p. 285-296, 2002. Disponível em: < <http://www.jleukbio.org/content/72/2/285.abstract> >.

GALL, T. L.; ROMERO, P. R.; CORTESE, M. S.; UVERSKY, V. N.; DUNKER, A. K. Intrinsic Disorder in the Protein Data Bank. **Journal of Biomolecular Structure and Dynamics**, v. 24, n. 4, p. 325-341, 2007. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/17206849> >.

GANDHI, N. R. *et al.* Multidrug-resistant and extensively drug-resistant tuberculosis: a threat to global control of tuberculosis. **The Lancet**, v. 375, 2010. ISSN 9728.

GELLERT, M. *et al.* Novobiocin and coumermycin inhibit DNA supercoiling catalyzed by DNA gyrase. **Proceedings of the National Academy of Sciences**, v. 73, n. 12, p. 4474-4478, 1976. Disponível em: < <http://www.pnas.org/content/73/12/4474.abstract> >.

GORDON, B. R. *et al.* Lsr2 is a nucleoid-associated protein that targets AT-rich sequences and virulence genes in *Mycobacterium tuberculosis*. **Proc Natl Acad Sci U S A**, United States, v. 107, n. 11, p. 5154-9, 2010. ISSN 1091-6490. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/20133735> >.

GORDON, B. R. G. *et al.* Lsr2 of *Mycobacterium* represents a novel class of H-NS-like proteins. **Journal of bacteriology**, v. 190, n. 21, p. 7052-9, 2008. Disponível em: < <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2580683&tool=pmcentrez&rendertype=abstract> >.

GHOSH, S.; INDI, S. S.; NAGAJARA, V. Regulation of Lipid Biosynthesis, Sliding Motility, and Biofilm Formation by a Membrane-Anchored Nucleoid-Associated Protein of *Mycobacterium tuberculosis*. **Journal of Bacteriology**, v. 195, n. 8, p. 1769-1778, 2013. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/23396914> >.

GRIFFIN, J. E.; GAWRONSKI, J. D.; DEJESUS, M. A.; IOERGER, T. R.; AKERLEY, B. J.; SASSETI, C. M. High-Resolution Phenotypic Profiling Defines Genes Essential for Mycobacterial Growth and Cholesterol Catabolism. **PLoS Pathogens**, v. 7, i. 9, 2011. 10.1371/journal.ppat. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/21980284> >.

HAYTER, J. R. *et al.* Proteome Analysis of Intact Proteins in Complex Mixtures. **Molecular & Cellular Proteomics**, v. 2, n. 2, p. 85-95, 2003. Disponível em: < <http://www.mcponline.org/content/2/2/85.abstract> >.

HELLMAN, L. M.; FRIED, M. G. Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. **Nat Protoc**, England, v. 2, n. 8, p. 1849-61, 2007. ISSN 1750-2799. Disponível em: < <http://www.nature.com/nprot/journal/v2/n8/full/nprot.2007.249.html> >.

HENDZEL, M. J. *et al.* The C-terminal Domain Is the Primary Determinant of Histone H1 Binding to Chromatin in Vivo. **Journal of Biological Chemistry**, v. 279, n. 19, p. 20028-20034, 2004. Disponível em: < <http://www.jbc.org/content/279/19/20028.abstract> >.

JEDRUSIK, M. A.; SCHULZE, E. A single histone H1 isoform (H1.1) is essential for chromatin silencing and germline development in *Caenorhabditis elegans*. **Development**, v. 128, n. 7, p. 1069-1080, 2001. Disponível em: < <http://dev.biologists.org/content/128/7/1069.abstract> >.

JIN, D. J.; CABRERA, J. E. Coupling the distribution of RNA polymerase to global gene regulation and the dynamic structure of the bacterial nucleoid in *Escherichia coli*. **J. Struct Biol**, v. 156, n. 2, p. 284-291, 2006. Disponível em: < [http://linkinghub.elsevier.com/retrieve/pii/S1047-8477\(06\)00213-9](http://linkinghub.elsevier.com/retrieve/pii/S1047-8477(06)00213-9) >.

KAMAKAKA, R. T.; THOMAS, J. O. Chromatin structure of transcriptionally competent and repressed genes. **EMBO J**, v. 9, n. 12, p. 3997-4006, 1990. ISSN 0261-4189. Disponível em: < <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC552172/> >.

KASINSKY, H. E. *et al.* Origin of H1 linker histones. **FASEB J**, v. 15, n. 1, p. 34-42, 2001. ISSN 0892-6638. Disponível em: < <http://www.fasebj.org/cgi/pmidlookup?view=long&pmid=11149891> >.

KATSUBE, T. *et al.* Control of Cell Wall Assembly by a Histone-Like Protein in *Mycobacteria*. **Journal of Bacteriology**, v. 189, n. 22, p. 8241-8249, 2007. Disponível em: < <http://jb.asm.org/content/189/22/8241.abstract> >.

KOCH, C. *et al.* The N-terminal part of the E.coli DNA binding protein FIS is essential for stimulating site-specific DNA inversion but is not required for specific DNA binding. **Nucleic Acids Res**, v. 19, n. 21, p. 5915-22, 1991. ISSN 1362-4962. Disponível em: < <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC329047/> >.

KOUL, A. *et al.* The challenge of new drug discovery for tuberculosis. **Nature**, England, v. 469, n. 7331, p. 483-90, 2011. ISSN 1476-4687. Disponível em: < <http://www.nature.com/nature/journal/v469/n7331/full/nature09657.html> >.

LAWN, S. D.; ZUMLA, A. I. Tuberculosis. **The Lancet**, v. 378, p. 16, 2011. ISSN 9785.

LIU, D. *et al.* The essentiality and involvement of *Streptococcus intermedius* histone-like DNA-binding protein in bacterial viability and normal growth. **Molecular Microbiology**, v. 68, n. 5, p. 1268-1282, 2008. ISSN 1365-2958. Disponível em: < <http://dx.doi.org/10.1111/j.1365-2958.2008.06232.x> >.

LUIJSTERBURG, M. S. *et al.* The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: a molecular perspective. **Journal of structural biology**, v. 156, n. 2, p. 262-72, 2006. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/16879983> >.

LUPAS, A.; VAN DYKE, M.; STOCK, J. Predicting coiled coils from protein sequences. **Science**, v. 252, n. 5010, p. 1162-4, 1991. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/2031185> >.

MADIGAN, M. *et al.* **Microbiologia de Brock**. 12. Porto Alegre: Artmed, 2010. 1160.

MDLULI, K.; MA, Z. *Mycobacterium tuberculosis* DNA gyrase as a target for drug discovery. **Infect Disord Drug Targets**, v. 7, n. 2, p. 159-68, 2007. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/17970226> >.

MIRKIN, S. M. DNA Topology: Fundamentals. In: (Ed.). **eLS**, 2001. ISBN 9780470015902.

NYE, M. B. *et al.* *Vibrio cholerae* H-NS Silences Virulence Gene Expression at Multiple Steps in the ToxR Regulatory Cascade. **Journal of Bacteriology**, v. 182, n. 15, p. 4295-4303, 2000. Disponível em: < <http://jb.asm.org/content/182/15/4295.abstract> >.

ONO, S. *et al.* H-NS is a part of a thermally controlled mechanism for bacterial gene regulation. **Biochem J**, v. 391, n. Pt 2, p. 203-13, Oct 15 2005. ISSN 1470-8728.

OWEN-HUGHES, T. A. *et al.* The Chromatin-Associated Protein H-NS Interacts with Curved DNA to Influence DNA Topology and Gene Expression. **Cell**, v. 71, p. 255-265, 1992.

PARISH, T.; STOKER, N. G. ***Mycobacterium tuberculosis* Protocols**. Totowa, NJ: Humana Press Inc., 2001.

PATEL, K.; JHAMB, S. S.; SINGH, P. P. Models of Latent Tuberculosis: Their Salient Features, Limitations, and Development. **Journal of Laboratory Physicians**, v. 3, i. 2, 2011. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/22219558> >.

PELICIC, V. *et al.* Efficient allelic exchange and transposon mutagenesis in *Mycobacterium tuberculosis*. **Proceedings of the National Academy of Sciences**, v. 94, n. 20, p. 10955-10960, 1997. Disponível em: < <http://www.pnas.org/content/94/20/10955.abstract> >.

- PRABHAKAR, S. *et al.* Identification of an immunogenic histone-like protein (HLPMT) of *Mycobacterium tuberculosis*. **Tuber Lung Dis**, v. 79, n. 1, p. 43-53, 1998. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/10645441> >.
- RADIJOVAC, P.; *et al.* Protein flexibility and intrinsic disorder. **Protein Science**, v. 13, p. 71-80, 2004. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/14691223> >.
- ROBINOW, C.; KELLENBERGER, E. The bacterial nucleoid revisited. **Microbiol, Rev**, v. 58, n. 2, p. 211–232, 1994. Disponível em: < <http://www.ncbi.nlm.nih.gov/pmc/articles/pmid/7521510/> >.
- ROUVIERE-YANIV, J.; GROS, F. Characterization of a novel, low-molecular-weight DNA-binding protein from *Escherichia coli*. **Biochemistry**, v. 72, p. 3428-3432, 1975.
- RUSSELL, D. G. Who puts the tubercle in tuberculosis? **Nat Rev Micro**, v. 5, n. 1, p. 39-47, 2007. ISSN 1740-1526. Disponível em: < <http://dx.doi.org/10.1038/nrmicro1538> >.
- SAMBROOK, J.; RUSSEL, D. W. **Molecular cloning: a laboratory manual**. 3. 2001. ISBN 0-87969-576-5.
- SASSETTI, C. M.; BOYD, D. H.; RUBIN, E. J. Genes required for mycobacterial growth defined by high density mutagenesis. **Mol Microbiol**, v. 48, n. 1, p. 77-84, 2003. Disponível em: < <http://onlinelibrary.wiley.com/doi/10.1046/j.1365-2958.2003.03425.x/abstract> >.
- SCHNEIDER, R. *et al.* An architectural role of the *Escherichia coli* chromatin protein FIS in organising DNA. **Nucleic acids research**, v. 29, n. 24, p. 5107-5114, 12/ 2001. Disponível em: < <http://ukpmc.ac.uk/abstract/MED/11812843> >.
- SHARADAMMA, N. *et al.* *Mycobacterium tuberculosis* nucleoid-associated DNA-binding protein H-NS binds with high-affinity to the Holliday junction and inhibits strand exchange promoted by RecA protein. **Nucleic acids research**, v. 38, n. 11, p. 3555-69, 2010. Disponível em: < <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2887947&tool=pmcentrez&rendertype=abstract> >.
- SICKMEIER, M.; *et al.* DisProt: the Database of Disordered Proteins. **Nucleic Acids Research**, v. 35, 2006. 10.1093/nar/gkl893. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/17145717> >.
- SINGH, B. *et al.* Growth, cell division and sporulation in mycobacteria. **Antonie Van Leeuwenhoek**, v. 98, n. 2, p. 165-77, 2010. ISSN 1572-9699. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/20437098> >.
- SONNHAMMER, E. L.; VON HEIJNE, G.; KROGH, A. A hidden Markov model for predicting transmembrane helices in protein sequences. **Proc Int Conf Intell Syst Mol Biol**, v. 6, p. 175-82, 1998. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/9783223> >.
- STAHL, F. W. The Holliday junction on its thirtieth anniversary. **Genetics**, v. 138, n. 2, p. 241-6, 1994. Disponível em: < <http://www.genetics.org/content/138/2/241.full.pdf+html> >.
- STEWART, G. R. *et al.* Dissection of the heat-shock response in *Mycobacterium tuberculosis* using mutants and microarrays. **Microbiology**, v. 148, n. 10, p. 3129-3138, 2002. Disponível em: < <http://mic.sgmjournals.org/content/148/10/3129.abstract> >.
- STOYANOV, A.; ZHUKOV, M.; RIGHETTI, P. G. The Proteome Revisited: Theory and practice of all relevant electrophoretic steps. Elsevier Science, 1 edition, 2001. 410.

SWINGER, K. K. *et al.* Flexible DNA bending in HU-DNA cocrystal structures. **EMBO J**, v. 22, n. 14, p. 3749-60, 2003. Disponível em: < <http://www.nature.com/emboj/journal/v22/n14/full/7595261a.html> >.

SWINGER, K. K.; RICE, P. A. IHF and HU: flexible architects of bent DNA. **Curr Opin Struct Biol**, England, v. 14, n. 1, p. 28-35, 2004. Disponível em: < <http://linkinghub.elsevier.com/retrieve/pii/S0959440X0300188X> >.

SÁNCHEZ-QUITIAN, Z. A. *et al.* Crystal structure determination and dynamic studies of *Mycobacterium tuberculosis* Cytidine deaminase in complex with products. **Archives of Biochemistry and Biophysics**, v. 509, n. 1, p. 108–115, 2011. Disponível em: < <http://dx.doi.org/10.1016/j.abb.2011.01.022> >.

TENDENG, C.; BERTIN, P. N. H-NS in Gram-negative bacteria: a family of multifaceted proteins. **Trends in Microbiology**, v. 11, n. 11, p. 511-518, 2003. Disponível em: < <http://linkinghub.elsevier.com/retrieve/pii/S0966842X03002610> >.

THANBICHLER, M.; WANG, S. C.; SHAPIRO, L. The bacterial nucleoid: a highly organized and dynamic structure. **Journal of cellular biochemistry**, v. 96, n. 3, p. 506-21, 2005. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/15988757> >.

THEILLET, F., *et al.* The alphabet of intrinsic disorder. I.: Act like a Pro: On the abundance and roles of proline residues in intrinsically disordered proteins. **Intrinsically Disordered Proteins**, 2013. 10.4161/idp.24360. Disponível em: < <https://www.landesbioscience.com/journals/idp/article/24360/> >.

TOMPA, P. Intrinsically unstructured proteins. **TRENDS Biochem Sci.**, v. 27, p. 527-533, 2002. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/12368089> >.

\_\_\_\_\_. **Structure and function of intrinsically disordered proteins**. CRC Press, 2009. ISBN 978-1-4200-7892-3.

\_\_\_\_\_. Intrinsically disordered proteins: a 10-year recap. *Trends Biochem Sci.*, 2012. 10.1016/j.tibs.2012.08.004. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/22989858> >.

TUPPER, A. E. *et al.* The chromatin-associated protein H-NS alters DNA topology in vitro. **The EMBO journal**, v. 13, n. 1, p. 258-68, 1994. Disponível em: < <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=394800&tool=pmcentrez&rendertype=abstract> >.

UDWADIA, Z. F. *et al.* Totally Drug-Resistant Tuberculosis in India. **Clinical Infectious Diseases**, v. 54, n. 4, p. 579-581, 2012. Disponível em: < <http://cid.oxfordjournals.org/content/54/4/579.short> >.

UVERSKY, V. N. Intrinsically disordered proteins from A to Z. **Int J Biochem Cell Biol**, v. 43, p. 1090-1103, 2011. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/21501695> >.

VULTOS, T. *et al.* Modification of the mycobacteriophage Ms6 attP core allows the integration of multiple vectors into different tRNA<sup>ala</sup> T-loops in slow- and fast-growing mycobacteria. **BMC Molecular Biology**, v. 7, n. 47, 2006. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/17173678> >.

WARD, J. J., SODHI, J. S., MCGUFFIN, L. J., BUXTON, B. F., AND JONES, D. T. Prediction and Functional Analysis of Native Disorder in Proteins from the Three Kingdoms of Life. **J. Mol. Biol.**, v. 337, p. 635-645, 2004. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/15019783> >.

WAYNE, L. G.; HAYES, L. G. *Mycobacterium tuberculosis* through Two Stages of Nonreplicating Persistence. **Infection and Immunity**, v. 64, n. 6, p. 2062-2069. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/8675308> >.

WERLANG, I. C. R. *et al.* Identification of Rv3852 as a nucleoid-associated protein in *Mycobacterium tuberculosis*. **Microbiology**, v. 155, n. 8, p. 2652-63, 2009. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19477901> >.

WHO. **Global tuberculosis report 2013**. 2013. Disponível em: < [http://www.who.int/tb/publications/global\\_report/en/](http://www.who.int/tb/publications/global_report/en/) >

WOLFE, L. M. *et al.* Proteomic definition of the cell wall of *Mycobacterium tuberculosis*. **J Proteome Res**, v. 9, n. 11, p. 5816-26, 2010. ISSN 1535-3907. Disponível em: < <http://pubs.acs.org/doi/abs/10.1021/pr1005873> >.

WORCEL, A.; BURGI, E. On the structure of the folded chromosome of *Escherichia coli*. **Journal of Molecular Biology**, v. 71, n. 2, p. 127-147, 1972. ISSN 0022-2836. Disponível em: < <http://www.sciencedirect.com/science/article/pii/0022283672903427> >.

**ANEXO A** - Comprovante de submissão do artigo científico ao periódico *Journal of Biological Chemistry*

[JBC/2014/562223 Manuscript Submission](#)

[jbc@asbmb.org](mailto:jbc@asbmb.org)

Para: Cristiano Valim Bizarro

sexta-feira, 28 de de 2014 17:11

MS ID#: JBC/2014/562223

MS TITLE: The Mycobacterial DNA-Binding Protein 2 (MDP2) from *Mycobacterium tuberculosis* H37Rv is an Intrinsically Disordered Protein

Dear Dr. Bizarro:

This is an automatic message acknowledging your online submission to The Journal of Biological Chemistry.

Thank you for your submission.

Sincerely,

Editorial Staff  
The Journal of Biological Chemistry

Sent on: February 28, 2014