Raphael Santa Rosa Sayegh

**Flexibilidade Conformacional do Domínio
Catalítico da Fosfatase Cdc25B**

Versão da tese corrigida

São Paulo

Data do depósito na SPG:

27/02/2016

Raphael Santa Rosa Sayegh

Flexibilidade Conformacional do Domínio Catalítico da Fosfatase Cdc25B

Flexibilidade conformacional do domínio catalítico da fosfatase Cdc25B. Tese apresentada ao Instituto de Química da Universidade de São Paulo para a obtenção do título de Doutor em Ciências

Instituto de Química – Universidade de São Paulo
Programa de Pós-Graduação em Ciências Biológicas (Bioquímica)

Orientador: Guilherme Menegon Arantes
Coorientadores: Roberto Kopke Salinas e Sandro Roberto Marana

São Paulo
2016

Ficha Catalográfica

Elaborada pela Divisão de Bibliotecas e Documentação do Conjunto das Químicas da USP.

Sayegh, Raphael Santa Rosa
S274f

Flexibilidade Conformacional do Domínio Catalítico da Fosfatase Cdc25B / Raphael
Santa Rosa Sayegh. – São Paulo, 2016.
129 p.

Tese (doutorado) – Instituto de Química da Universidade de São Paulo. Departamento
de Bioquímica.

Orientador: Arantes, Guilherme Menegon

Coorientador: Salinas, Roberto Kopke

Coorientador: Marana, Sandro Roberto

1. Biofísica I. T. II. Arantes, Guilherme Menegon, orientador. III. Salinas, Roberto
Kopke, coorientador. IV. Marana, Sandro Roberto, coorientador.

574.191 CDD

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE QUÍMICA

“Flexibilidade conformacional do domínio catalítico da fosfatase Cdc25B”

RAPHAEL SANTA ROSA SAYEGH

Tese de Doutorado submetida ao Instituto de Química da Universidade de São Paulo como parte dos requisitos necessários à obtenção do grau de Doutor em Ciências no Programa de Ciências Biológicas (Bioquímica) - Área de Concentração: Bioquímica.

**Prof. Dr. Guilherme Menegon Arantes
(Orientador e Presidente)**

Aprovado (a) por:

**Prof. Dr. Shaker Chuck Farah
IQ - USP**

**Prof. Dr. Leandro Martinez
IQ - UNICAMP**

**Prof. Dr. Fábio Ceneviva Lacerda Almeida
UFRJ**

**SÃO PAULO
14 de março de 2016**

Aos meus pais, pelo apoio e carinho

Agradecimentos

- Ao meu orientador, Prof. Guilherme Menegon Arantes, por seu empenho dedicado ao desenvolvimento desta tese e à minha formação acadêmica. Sob sua influência compreendi que a verdadeira ciência é fruto de uma análise sistemática e exaustiva dos fenômenos observados e o quão prazerosa ela pode ser. Nossa convivência dentro e fora do ambiente de trabalho certamente também contribuiu para minha formação pessoal;
- Ao meu coorientador, Prof. Roberto Kopke Salinas, pela contribuição decisiva para realização deste trabalho e principalmente pelos seus ensinamentos sobre ressonância magnética nuclear. Seu empenho tem sido estimulante para que eu aprofunde meus conhecimentos sobre essa técnica;
- Ao meu coorientador, Prof. Sandro Roberto Marana, pela enorme colaboração para a concretização deste trabalho, por abrir as portas do seu laboratório e pelas discussões sobre cinética enzimática. Sua maneira de enxergar os resultados e idealizar novos experimentos foram marcantes;
- Ao Dr. Fábio Kendi Tamaki, por sua colaboração direta nas etapas iniciais deste estudo e por me auxiliar no laboratório a conduzir a etapa de biologia molecular do projeto;
- Aos meus colegas de laboratório, Ariane, Jota, Murilo e Vane, pelo ótimo convívio, pelas discussões e pela disposição em me auxiliar com os obstáculos que surgiram ao longo desses anos;
- À Profa. Denize Cristina Favaro, pela sua ajuda na implementação dos experimentos de ressonância magnética nuclear e na operação do espectrômetro;
- Ao Prof. Shaker Chuck Farah, por possibilitar o acesso a equipamentos do seu laboratório;
- À minha mãe, Marilene, por ter sempre me apoiado e também se preocupado com minhas decisões pessoais e profissionais;
- Ao meu pai, Luiz, pelos valores ensinados;
- Às minhas irmãs, Carolina e Leticia, pelo carinho e os ótimos momentos que passamos juntos;

- À minha namorada, Beatriz, pelo apoio, companheirismo e, muitas vezes, paciência. Momentos ao seu lado são inspiradores e alegram minha vida. Se isso é manifestação do amor, eu te amo;
- Aos Mentecaptos, meu amigos de vida, pela parceria inabalável;
- À família Bio pela amizade intimista e pelo convívio mais prazeroso que eu consigo imaginar;
- À Central Analítica do Instituto de Química da USP por disponibilizar o uso de equipamentos;
- A todos os funcionários e professores do Instituto de Química da USP que direta ou indiretamente contribuíram para este trabalho e minha formação;
- Ao CNPq pela concessão de bolsa;
- À FAPESP pela concessão de bolsa (2012/00543-1) e financiamento do projeto.

“(..) este brave new world é obra de uma sociedade secreta de astrônomos, biólogos, engenheiros, metafísicos, poetas, químicos, algebristas, moralistas, pintores, geômetras (...) dirigidos por um obscuro homem de gênio. Sobram indivíduos que dominam essas diversas disciplinas, mas não os capazes de invenção e menos ainda os capazes de subordinar a invenção a um rigoroso plano sistemático.”

Jorge Luis Borges

Resumo

Sayegh, R.S.R. **Flexibilidade Conformacional do Domínio Catalítico da Fosfatase Cdc25B**. 2016. 129 p. Tese (doutorado) – Programa de Pós-Graduação em Ciências Biológicas (Bioquímica). Instituto de Química, Universidade de São Paulo, São Paulo.

A fosfatase Cdc25B atua na progressão do ciclo celular através da ativação de complexos Cdk/Ciclina. Atualmente, nos modelos estruturais propostos do domínio catalítico da Cdc25B não estão incluídos os últimos 16 resíduos da região C-terminal. Este segmento tem importante papel no reconhecimento do substrato proteico e pode estar envolvido na complexação de pequenas moléculas com a Cdc25B. Assim, o principal objetivo desta tese foi avaliar a flexibilidade conformacional do domínio catalítico *completo* da Cdc25B em solução através de simulações computacionais e por medidas experimentais de ressonância magnética nuclear (RMN). A similaridade entre as estruturas cristalográficas e em solução foi confirmada pela previsão de ângulos diedrais φ/ψ da cadeia principal a partir dos deslocamentos químicos (CS) e pela concordância entre os acoplamentos dipolares residuais (RDC) medidos e calculados a partir da geometria cristalina. Medidas de parâmetros de relaxação de ^{15}N e RDC evidenciaram a presença de desordem conformacional na região C-terminal, em acordo com a ausência de densidade eletrônica desse segmento no experimento de difração de raios-X. Através da comparação entre CS experimentais e previstos de simulações de dinâmica molecular (DM) longas (total de 6 μs de duração) foram apontados artefatos de cristalização, possíveis erros nos campos de força usados nas simulações, falhas na composição do sistema simulado e estados conformacionais populados pela Cdc25B em solução distintos da geometria cristalográfica. De maneira geral, os CS previstos a partir das simulações para a flutuação estrutural dos resíduos da região C-terminal desordenada estão em acordo com os valores experimentais, sugerindo que os estados conformacionais deste segmento foram razoavelmente bem amostrados nas simulações. Em particular, verificou-se que o contato tipo cátion- π entre as cadeias laterais dos resíduos 550W do C-terminal desordenado e 482R do núcleo proteico, ausente na estrutura cristalográfica, pode ser importante em solução. A formação desse contato na simulação de DM também está de acordo com medidas experimentais de perturbação de deslocamentos químicos (CSP) entre construções completa e truncada do domínio catalítico da Cdc25B. Assim, através do uso conjunto de simulações computacionais e medidas experimentais foi possível obter uma representação mais completa e realista da flexibilidade conformacional do domínio catalítico da Cdc25B em solução, incluindo a determinação de possíveis contatos intramoleculares entre a região C-terminal desordenada e o núcleo proteico. Essas informações poderão ser usadas na construção de um *ensemble* conformacional da Cdc25B.

Palavras-chave: ressonância magnética nuclear; RMN; simulação computacional; dinâmica molecular; parâmetros de ordem; acoplamento dipolar residual; desordem conformacional; contatos intramoleculares; trajetória do deslocamento químico.

Abstract

Sayegh, R.S.R. **Conformational Flexibility of the Catalytic Domain of Cdc25B Phosphatase**. 2016. 129 p. PhD Thesis – Graduate Program in Biochemistry. Instituto de Química, Universidade de São Paulo, São Paulo.

Cdc25B phosphatase acts on the progression of cell cycle through the activation of Cdk/Cyclin complexes. Currently, the proposed structural models of Cdc25B catalytic domain lack the last 16 residues from the C-terminal region. This segment is important for protein substrate recognition and might be involved in small molecule binding to Cdc25B. Thus, the main goal of this thesis was to evaluate the conformational flexibility of the *complete* catalytic domain from Cdc25B through computer simulations and experimental nuclear magnetic resonance (NMR) measurements. Similarity between crystal and in solution structures was confirmed by the prediction of backbone φ/ψ dihedral angles from chemical shifts (CS) and by the agreement between observed and back-calculated residual dipolar couplings (RDC). ^{15}N relaxation and RDC measurements pointed to the conformational disorder of the C-terminal region, in agreement with the X-ray diffraction experiment where this segment showed no electronic density. Comparison between experimental and predicted CS from long molecular dynamics (MD) simulations (6 μs total running time) pointed to the presence of crystallographic artifacts, possible deficiencies in simulation force fields, inaccurate composition of the simulated system and conformational states visited by Cdc25B in solution that were not observed in the crystallographic geometry. Generally, CS predicted from simulations for the structural fluctuation of the disordered C-terminal region were in agreement with experimental values, suggesting that the simulations sampled the conformational states populated by this segment reasonably well. In particular, a cation- π contact not observed in the crystal structure between side chains of residue 550W from the disordered C-terminal tail and residue 482R from the protein core might be important in solution. This contact is also in agreement with experimental chemical shift perturbations (CSP) measured between complete and truncated constructs of Cdc25B catalytic domain. Therefore, the joint use of computer simulations and experimental measurements allowed the achievement of a more complete and realistic representation of the conformational flexibility of the Cdc25B catalytic domain in solution, including intramolecular contacts between the disordered C-terminal region and the protein core. This information might be used to obtain a conformational ensemble of Cdc25B.

Keywords: nuclear magnetic resonance; NMR; computer simulation; molecular dynamics; order parameter; residual dipolar coupling; conformational disorder; intramolecular contacts; chemical shift trajectory.

Lista de ilustrações

Figura 1 – Escalas de tempo típicas de movimentações presentes em proteínas. . .	27
Figura 2 – Pontos de checagem que atuam em diferentes etapas do ciclo celular. . .	30
Figura 3 – Regulação da atividade enzimática dos complexos Cdk/Ciclina.	31
Figura 4 – Sobreposição das estruturas cristalográfica da Cdc25A (magenta; PDB: 1C25) e Cdc25B (verde; PDB: 1QB0) determinadas por difração de raios-X.	33
Figura 5 – Estruturas representativas do <i>ensemble</i> conformacional da Cdc25B obtido por simulação LBMC.	35
Figura 6 – Desestruturação da hélice C-terminal da estrutura cristalográfica durante simulação de DM.	36
Figura 7 – Taxas de relaxação da magnetização transversal (R_2) dos <i>spins</i> de ^{15}N medidas para construções do domínio catalítico da Cdc25B selvagem e mutante (C473S).	37
Figura 8 – Mapa do plasmídeo pET-28a com detalhes da região de clonagem. . . .	42
Figura 9 – Sequência primária da região N-terminal das construções <i>C1</i> e <i>C2</i> . . .	43
Figura 10 – Sequência primária das construções <i>C1</i> e <i>C2</i>	45
Figura 11 – Alinhamento entre o sequenciamento das construções <i>C1</i> , <i>C2</i> e o cDNA da Cdc25B.	55
Figura 12 – Dupla digestão dos plasmídeos pET-28a_ <i>C1</i> e pET-28a_ <i>C2</i> com as enzimas de restrição NdeI e BamHI.	55
Figura 13 – Perfis de expressão de proteínas das bactérias transformadas com os vetores pET-28a_ <i>C1</i> e pET-28a_ <i>C2</i>	56
Figura 14 – Perfil proteico das etapas do primeiro passo de purificação da construção <i>C2</i> em coluna Ni-NTA agarose.	56
Figura 15 – Composição das amostras obtidas após as etapas de purificação da construção <i>C1</i>	57

Figura 16 – Regiões do espectro tridimensional HNCA utilizadas no assinalamento da cadeia principal da α -hélice C-terminal (536F–546K) da construção <i>C1</i>	58
Figura 17 – Espectro de ^1H – ^{15}N HSQC da construção <i>C1</i>	59
Figura 18 – Assinalamento do espectro de ^1H – ^{15}N HSQC construção <i>C1</i>	60
Figura 19 – Espectros sobrepostos de ^1H – ^{15}N HSQC adquiridos durante a titulação da construção <i>C1</i> com NaPi.	61
Figura 20 – Perturbações do deslocamento químico da construção <i>C1</i> obtidas por titulação com NaPi.	62
Figura 21 – Regiões do espectro tridimensional HN(CA)CB onde são apresentados os picos de correlação $^1\text{H} \times ^{13}\text{C}^\beta$ das cisteínas 426, 473 e 484.	62
Figura 22 – Comparação entre ângulos diedrais φ e ψ extraídos da estrutura cristalográfica e previstos pelo programa TALOS-N para a construção <i>C1</i>	63
Figura 23 – Comparação entre a estruturas secundárias cristalográfica e previstas por TALOS-N ⁵⁹ ou CSI ⁸³	63
Figura 24 – Populações de estruturas secundárias e flexibilidade conformacional da construção <i>C1</i> estimadas pelo método $\delta 2\text{D}$ ⁸⁶	66
Figura 25 – Medidas de relaxação dos <i>spins</i> de ^{15}N da cadeia principal da construção <i>C1</i>	68
Figura 26 – Parâmetros do formalismo de Lipari-Szabo ajustados a partir das medidas de relaxação (R_1 , R_2 e NOE) da construção <i>C1</i>	69
Figura 27 – Acoplamentos dipolar residual dos grupos amida da cadeia principal (D_{NH}) da construção <i>C1</i>	70
Figura 28 – Sobreposição dos espectros de ^1H – ^{15}N HSQC das construções <i>C1</i> (laranja) e <i>C2</i> (azul).	71
Figura 29 – Contatos entre a região C-terminal e o núcleo da construção <i>C1</i>	72
Figura 30 – Contatos entre diferentes moléculas na unidade assimétrica observados na estrutura cristalográfica da Cdc25B.	76

Figura 31 – Gráfico de Ramachandran gerado pelo programa TALOS-N com a previsão dos ângulos diedrais de 473C.	77
Figura 32 – Estruturas iniciais usadas nas simulações <i>DM1</i> e <i>DM2/DM3</i> da Cdc25B.	85
Figura 33 – Desvio médio quadrático (RMSD) das coordenadas de N da cadeia principal das trajetórias de dinâmica molecular <i>DM1</i> , <i>DM2</i> e <i>DM3</i> da Cdc25B.	87
Figura 34 – Flutuação média (RMSF) de N da cadeia principal dos resíduos da Cdc25B.	87
Figura 35 – Mudança conformacional da cadeia principal dos segmentos HCT e LCT observada na simulação <i>DM3</i>	89
Figura 36 – Descomplexação do íon PO_4^{3-} do P-loop durante as trajetórias de DM.	90
Figura 37 – Comparação entre os deslocamentos químicos experimentais e previstos da estrutura cristalográfica ou calculados a partir da simulação <i>DM1+2+3</i> .	93
Figura 38 – Segmentos do domínio catalítico da Cdc25B cujos deslocamentos químicos experimentais estão em melhor acordo com a geometria cristalográfica ($\Gamma < 0$; verde) ou as geometrias amostradas nas simulações de DM ($\Gamma > 0$; vermelho).	94
Figura 39 – Desestabilização do sítio P2 de ligação a PO_4^{3-} observado nas simulações de DM.	96
Figura 40 – Desestruturação do segmento 446–454 nas simulações de DM.	97
Figura 41 – Trajetórias dos deslocamentos químicos de $^{13}\text{C}^\beta$ do resíduo 395H e ^1H do resíduo 396Q previstos da simulação <i>DM1+2+3</i>	98
Figura 42 – Contatos cristalográficos do segmento 395–397 observados no experimento de difração de raios-X.	98
Figura 43 – Trajetória dos deslocamentos químicos de ^1H do resíduo 549S previstos da simulação <i>DM1+2+3</i>	100
Figura 44 – Estados conformacionais populados pelo segmento CTD na trajetória <i>DM1+2+3</i>	101

Figura 45 – Desvios médios quadráticos (RMSD^{cml}) entre médias cumulativas dos deslocamentos químicos previstas ao longo das trajetórias de DM ($\delta_{\text{DM}}^{\text{cml}}$) e os δ_{Exp} de $^{13}\text{C}^{\alpha}$ e $^{13}\text{C}^{\beta}$ dos resíduos do CTD* (551–566) e do núcleo (387–546).	103
Figura 46 – Desvios entre δ_{DM} da simulação <i>DM1+2+3</i> e δ_{Exp} dos resíduos do CTD* res. 551–565 da Cdc25B.	104
Figura 47 – Trajetória dos deslocamentos químicos de $^{13}\text{C}^{\alpha}$ e $^{13}\text{C}^{\beta}$ do resíduo 556R previstos dinâmica molecular <i>DM1+2+3</i>	105
Figura 48 – Matriz simétrica da distância média de C^{α} obtida da simulação <i>DM1+2+3</i> .	106
Figura 49 – Perturbações entre os deslocamentos químicos dos grupos amida das construções <i>C1</i> e <i>C2</i> medidas experimentalmente (vermelho) e calculadas a partir de estruturas obtidas da simulação <i>DM1+2+3</i> (preto). . .	106
Figura 50 – Trajetória dos deslocamentos químicos de $^{13}\text{C}^{\alpha}$, $^{13}\text{C}^{\beta}$, $^{13}\text{C}'$ do resíduo 482R previstos pelo programa SPARTA+ da simulação <i>DM1+2+3</i> . . .	107
Figura 51 – Contato tipo cátion- π formado entre as cadeias laterais de 482R e 550W durante as simulações de DM.	107

Lista de tabelas

Tabela 1 – <i>Primers</i> diretos (FW) e reversos (RV) utilizados na clonagem das construções <i>C1</i> e <i>C2</i> da Cdc25B.	39
Tabela 2 – Ciclos de temperatura da reação de PCR para amplificação do cDNA das construções <i>C1</i> e <i>C2</i>	40
Tabela 3 – Modelos para o ajuste dos parâmetros dinâmicos da equação (2.15) do formalismo de Lipari-Szabo.	54
Tabela 4 – Sinais esperados, observados e assinalados no espectro de $^1\text{H}-^{15}\text{N}$ HSQC da construção <i>C1</i> (Figura 17).	58
Tabela 5 – Deslocamentos químicos de ^1H e ^{15}N da cadeia principal do segmento C-terminal 549–566 observados e calculados para regiões desordenadas.	65
Tabela 6 – Parâmetros do tensor de difusão rotacional experimentais (RMN) e calculados (hidrodinâmicos).	67
Tabela 7 – Trajetórias de dinâmica molecular de 2 μs do domínio catalítico da Cdc25B.	82
Tabela 8 – Desvio médio quadrático (RMSD) entre os deslocamentos químicos experimentais e calculados da estrutura cristalográfica ou trajetórias de DM da Cdc25B.	91
Tabela 9 – Deslocamentos químicos da Cdc25B em significativo maior acordo com os valores previstos da estrutura cristalográfica (PDB) ou com os valores previstos das trajetórias de dinâmica molecular (DM).	92
Tabela 10 – Desvio médio padrão (RMSD) dos deslocamentos químicos do segmento CTD* (551-565) calculados das trajetórias de DM da Cdc25B.	101

Lista de abreviaturas e siglas

BMER	β -mercaptoetanol
Cdk	quinase dependente de ciclina
CS/ δ	deslocamento químico
CSI	índice do deslocamento químico
CSP	perturbação do deslocamento químico
CTD	C-terminal desordenado (resíduos 547–566)
δ_{DM}	deslocamentos químicos previstos das simulações de dinâmica molecular
δ_{Exp}	deslocamentos químicos experimentais
δ_{PDB}	deslocamentos químicos previstos da estrutura cristalográfica
EDTA	ácido etilenodiamino tetra-acético
FID	decaimento livre da indução
Hb	hemoglobina
HCT	hélice C-terminal (resíduos 534–546)
IDP	proteína intrinsecamente desordenada
IDRP	região proteica intrinsecamente desordenada
LB	Luria Broth
LCT	<i>linker</i> C-terminal (resíduos 525–533)
NTD	N-terminal desordenado (resíduos 366–396)
PCR	reação em cadeia da polimerase

PEG	polietilenoglicol
Pi	fosfato inorgânico (PO_4^{3-})
PMSF	fenilmetilsulfonilflúor
PTP	proteína tirosina fosfatase
$R_{1\rho}$	taxa de relaxação da magnetização longitudinal no referencial girante
R_1	taxa de relaxação da magnetização longitudinal
R_2	taxa de relaxação da magnetização transversal
RCI	índice de desordem conformacional
RDC	acoplamento dipolar residual
RMN	ressonância magnética nuclear
RMSD	desvio quadrático médio
RMSF	flutuação quadrática média
T_1	tempo de relaxação da magnetização longitudinal
T_2	tempo de relaxação da magnetização transversal
τ_c	tempo de correlação rotacional global

Sumário

	Sumário	23
1	INTRODUÇÃO	25
1.1	Flexibilidade conformacional de proteínas	25
1.2	Cdc25 e o ciclo celular	29
1.3	Descrições estruturais da Cdc25B	32
1.4	Objetivos	38
2	MEDIDAS ESPECTROSCÓPICAS DE RMN	39
2.1	Metodologia	39
2.1.1	Clonagem das construções <i>C1</i> e <i>C2</i>	39
2.1.2	Expressão e purificação	43
2.1.3	RMN	45
2.1.3.1	Assinalamento da cadeia principal	45
2.1.3.2	Previsão de ângulos diedrais φ/ψ da cadeia principal	47
2.1.3.3	Previsão de desordem conformacional	48
2.1.3.4	Acoplamento dipolar residual	48
2.1.3.5	Taxas de relaxação R_1 e R_2 e NOE heteronuclear $\{^1\text{H}\}-^{15}\text{N}$	50
2.1.3.6	Cálculo dos parâmetros dinâmicos pelo formalismo de Lipari-Szabo	51
2.2	Resultados	54
2.2.1	Clonagem, expressão e purificação	54
2.2.2	Assinalamento da cadeia principal	57
2.2.3	Flexibilidade conformacional	64
2.2.4	Contatos entre a região C-terminal e o núcleo proteico	70
2.3	Discussão	72
3	MODELAGEM MOLECULAR DO DOMÍNIO CATALÍTICO	79
3.1	Metodologia	79

3.1.1	Simulações de dinâmica molecular com potenciais clássicos	79
3.1.2	Previsão dos deslocamentos químicos	83
3.2	Resultados e discussão	85
3.2.1	Simulações da construção <i>C1</i>	85
3.2.2	Comparação dos modelos estruturais a partir de medidas experimentais . .	89
3.2.2.1	Concordância com a estrutura cristalográfica	94
3.2.2.2	Concordância com as trajetórias simuladas	97
3.2.3	Deslocamentos químicos previstos da região C-terminal desordenada	99
3.2.4	Contatos entre a região C-terminal e o núcleo proteico	103
4	CONCLUSÕES	109
	REFERÊNCIAS	111
	ANEXOS	121
	ANEXO A – MECANISMOS DE REGULAÇÃO ALOSTÉRICA NA	
	HEMOGLOBINA	123
	ANEXO B – MEDIDAS DE RELAXAÇÃO E PARÂMETROS DI-	
	NÂMICOS DA CDC25B	125
	ANEXO C – COMPARAÇÃO ENTRE OS DESLOCAMENTOS QUÍ-	
	MICOS PREVISTOS E EXPERIMENTAIS	129

1 Introdução

1.1 Flexibilidade conformacional de proteínas

Proteínas desempenham papéis fundamentais e diversificados na manutenção de um organismo. Elas podem atuar como componentes estruturais, transportadores, catalisadores de reações químicas, anticorpos, reguladores da expressão gênica, etc. Sua estrutura tridimensional é determinada pela sequência de aminoácidos. A relação entre a estrutura tridimensional e função de proteínas tornou-se o paradigma central da biologia estrutural, reforçado pelo grande número de estruturas tridimensionais de proteínas resolvidas principalmente por difração de raios-X¹.

O reconhecimento de uma proteína por um ligante é mediado pela complementariedade estrutural. Existem dois mecanismos consagrados que descrevem o reconhecimento proteína-ligante: o modelo “chave-fechadura”, proposto por Fischer^{2, 3} em 1894 e o modelo de “ajuste induzido”, proposto por Koshland⁴ em 1958. No modelo chave-fechadura, as geometrias do receptor e do ligante no estado ligado ou livre em solução são essencialmente as mesmas. Já no modelo de ajuste induzido, mudanças conformacionais podem ocorrer em ambas as moléculas após a complexação a fim de tornar suas geometrias mais compatíveis. Em ambos casos, supõe-se que as proteínas apresentam uma única conformação estável. Esses dois mecanismos de reconhecimento molecular foram utilizados na elaboração de modelos para explicar a cooperatividade observada experimentalmente na ligação entre O₂ e a hemoglobina (Hb) (Anexo A).

Mais recentemente, consagrou-se o termo “seleção conformacional” como um mecanismo de reconhecimento molecular⁵. Aqui, assume-se que um substrato é capaz de reconhecer um determinado estado da proteína, causando uma perturbação na superfície de energia livre que altera sua distribuição conformacional. Esse mecanismo está intimamente relacionado ao modelo MWC proposto para explicar a cooperatividade alostérica da ligação entre O₂ e Hb (Anexo A). Torna-se implícito, portanto, que as proteínas visitam diferentes estados conformacionais. Neste contexto, o modelo chave-fechadura pode ser visto como

um caso específico onde o ligante reconhece o estado de menor energia livre da proteína.

De fato, proteínas são caracterizadas como macromoléculas flexíveis em solução pois apresentam movimentações em diferentes escalas de tempo, incluindo desde vibrações das ligações covalentes até mudanças na orientação entre grandes domínios^{6, 7} (Figura 1). Assim como sua estrutura tridimensional, as movimentações de uma proteína em solução também são importantes para sua função^{8, 9, 6}. Essas movimentações podem ser racionalizadas por uma superfície de energia livre¹⁰ composta por diferentes estados conformacionais em equilíbrio dinâmico e separados por barreiras energéticas (Figura 1). A altura das barreiras ($\Delta G^\ddagger$) determina a frequência (k) com que ocorrem as transições entre os estados (cinética):

$$k = A \exp\left(\frac{-\Delta G^\ddagger}{k_B T}\right),$$

onde k_B é a constante de Boltzmann, T é a temperatura e A é um fator pré-exponencial que está relacionado à fricção com o solvente decorrente da transição. Já a diferença de energia livre entre dois estados A e B (ΔG^{AB}) estabelece uma relação entre suas populações P_A e P_B ¹¹ (termodinâmica):

$$\frac{P_A}{P_B} = \exp\left(\frac{\Delta G^{AB}}{k_B T}\right).$$

Dessa maneira, proteínas em solução podem ser vistas como um conjunto de estruturas que seguem uma distribuição conformacional relacionada à energia térmica ($k_B T$) do sistema.

Enquanto a função de diversas proteínas está intimamente associada ao seu enovelamento espacial, uma considerável fração do proteoma de um organismo é formada por segmentos ou por proteínas completas que não apresentam uma estrutura tridimensional bem definida, mas que ainda assim são funcionais nas condições fisiológicas. Estas são classificadas como regiões intrinsecamente desordenadas de proteínas (IDRP) ou proteínas intrinsecamente desordenadas (IDP)^{12, 13}. Aproximadamente 44% das proteínas humanas apresentam IDRPs com mais de 30 resíduos¹⁴. A baixa complexidade da estrutura primária e a predominância de resíduos com cadeias laterais polares ou carregadas em relação a resíduos com cadeias laterais hidrofóbicas, que comumente formam o núcleo de uma proteína globular, são características de segmentos desordenados¹⁵.

Figura 1 – Escalas de tempo típicas de movimentações presentes em proteínas. A taxa de transição (k) entre dois estados hipotéticos **A** e **B** é proporcional à energia de ativação ($\Delta G^\ddagger$). A população relativa dos estados (P_A, P_B) é proporcional à diferença da energia livre entre os estados (ΔG^{AB}).

Essa desordem estrutural confere uma plasticidade conformacional que está relacionada com a promiscuidade no reconhecimento molecular, uma vez que diversas IDP e IDRPs são capazes de formar complexos com mais de uma proteína, muitas vezes através de uma transição desordem-ordem¹⁶. Neste contexto, um mesmo segmento ou proteína pode adotar diferentes estruturas na formação de complexos com diferentes parceiros. Esse comportamento é refletido no papel central desempenhado por IDPs em redes de interação proteica. Especula-se também que a maior quantidade de segmentos desordenados no proteoma de organismos eucariontes em comparação aos procariontes está relacionada à maior complexidade dos primeiros¹⁵. Além disso, observa-se um importante papel de IDPs na sinalização celular¹⁷, inclusive no controle da divisão celular¹⁸.

Por exemplo, a p27 é uma IDP capaz de se ligar e inibir a atividade de diferentes complexos Cdk/Ciclina que atuam no ciclo celular¹⁹. No seu estado livre, p27 apresenta características estruturais de uma IDP, como grande flexibilidade da cadeia principal e

deslocamentos químicos (CS) do sinal espectroscópico de ressonância magnética nuclear (RMN) de $^{13}\text{C}^\alpha$ típicos de segmentos que apresentam desordem conformacional²⁰. Medidas de calorimetria de titulação isotérmica revelaram que a ligação da p27 com Cdk2/Ciclina A é entropicamente desfavorável²⁰. Este resultado é condizente com uma estruturação de p27 após a complexação, conforme observado na estrutura cristalográfica do complexo p27-Cdk2/Ciclina A²¹.

Além de alvos proteicos, IDRs podem participar da complexação de pequenos compostos orgânicos. Foi demonstrado por cristalografia de raios-X e experimentos de RMN que a ligação de uma piperidona no segmento N-terminal desordenado da MDM2, um regulador negativo da proteína repressora tumoral p53, provoca sua estruturação. Esta região da MDM2 também é importante na complexação com p53. Assim, a ligação da piperidona interfere na formação do complexo MDM2-p53, podendo contribuir para a supressão tumoral através do restabelecimento da atividade de p53²². Também foi verificado que a região C-terminal desordenada da proteína tirosina fosfatase (PTP) 1B, um importante alvo para o desenvolvimento de drogas contra diabetes e obesidade, está envolvida no reconhecimento de um pequeno inibidor alostérico²³.

Em comparação com uma proteína ou segmento proteico enovelado, a superfície de energia livre de uma IDP ou IDR é mais plana, caracterizando um *ensemble* de conformações rapidamente interconversíveis. Detalhar a estrutura e dinâmica conformacional de IDP ou IDR a partir de medidas experimentais é uma tarefa árdua, já que as medidas refletem médias temporais e populacionais das conformações visitadas. Contudo, o uso de medidas experimentais tanto para guiar como para validar simulações computacionais tem se mostrado uma abordagem eficiente para avaliar o comportamento de segmentos desordenados em solução^{24, 25, 26}. Dentre as técnicas experimentais, medidas espectroscópicas de RMN são capazes de fornecer informações globais e locais sobre a estrutura e dinâmica conformacional. Essas propriedades também são exploradas simultaneamente por simulações de dinâmica molecular (DM).

Através da comparação entre propriedades calculadas a partir de trajetórias de DM e medidas experimentais de RMN do fator de transcrição de levedura GCN4, foram obtidas evidências a respeito do mecanismo molecular para o reconhecimento entre o segmento

N-terminal desordenado de GCN4 e a molécula de DNA²⁷. Este segmento adota uma conformação em α -hélice no complexo DNA-proteína. Verificou-se que a alta helicidade residual nos resíduos centrais da interface DNA-proteína observada numa das trajetórias de GCN4 na ausência de DNA estão em melhor acordo com os CS experimentais e medidas de relaxação. A existência de uma hélice pré-formada em solução pode diminuir a penalidade entrópica associada a uma transição desordem-ordem observada na ligação do DNA e sugere um mecanismo de seleção conformacional para o reconhecimento molecular. Esse reconhecimento inicial poderia desencadear, através do mecanismo de ajuste induzido, a estruturação do restante do segmento desordenado de GCN4. Assim, foi proposto que a complexação entre o fator de transcrição GCN4 e o DNA é mediado por um mecanismo híbrido de seleção conformacional e ajuste induzido²⁷.

Considerando-se ainda que a caracterização dos estados conformacionais populados por IDPs ou IDRs livres em solução pode auxiliar na compreensão dos mecanismos de reconhecimento molecular mediados por regiões desordenadas.

1.2 Cdc25 e o ciclo celular

O ciclo celular é caracterizado por uma sequência ordenada de eventos onde ocorre a duplicação do material genético nuclear e sua divisão equacional entre duas células-filhas. Esse processo pode ser dividido em quatro fases sequenciais: *G1*, *S* (síntese do DNA) e *G2*, chamadas conjuntamente de intérfase, e *M* (mitose). Nas fases *G1* e *G2* ocorre o crescimento celular e aumento da expressão de proteínas necessárias para a duplicação dos cromossomos durante a fase *S* e sua segregação na fase *M*, respectivamente. Em seres unicelulares, como bactérias e leveduras, a divisão celular é responsável pela formação de novos organismos completos. Em espécies multicelulares, um organismo funcional é gerado por eventos consecutivos de divisão celular. No indivíduo adulto, esse processo também é responsável pela reposição de células mortas e regeneração de tecidos²⁸.

Nas células eucarióticas, a progressão do ciclo celular pode ser controlada por uma rede complexa de proteínas regulatórias que atuam em pontos de checagem. Os três principais ocorrem nas transições *G1/S* e *G2/M* da intérfase e metáfase/anáfase

da fase *M* (Figura 2). No ponto de checagem *G1/S* ocorre a verificação das condições intra e extracelular para o comprometimento das células com as fases *S* e *M* do ciclo. No ponto de checagem *G2/M* há mecanismos que verificam se as condições externas ainda estão favoráveis, se todo o material genético foi replicado e se eventuais danos ao DNA foram reparados. Durante o ponto de checagem da transição metáfase/anáfase, a ligação dos cromossomos aos fusos mitóticos é verificada para que o início da separação das cromátides irmãs proceda durante a anáfase (Figura 2). Dessa maneira, os pontos de checagem possuem mecanismos que avaliam as condições celulares e podem bloquear a progressão do ciclo até que eventuais problemas possam ser corrigidos ou, caso contrário, direcionar a célula para apoptose²⁸.

Figura 2 – Pontos de checagem que atuam em diferentes etapas do ciclo celular.

O cerne do controle do ciclo celular ocorre através da regulação da atividade enzimática de quinases dependente de ciclina (Cdk). A atividade das Cdk oscila durante a progressão do ciclo, alterando o padrão de fosforilação de alvos intracelulares envolvidos na iniciação ou regulação dos principais eventos, como a duplicação e a segregação dos cromossomos. O aumento da atividade Cdk no ponto de checagem *G2/M*, por exemplo, promove a fosforilação e ativação de proteínas relacionadas ao controle da condensação dos cromossomos e quebra do envoltório nuclear. Essa oscilação na atividade das Cdk é resultado de alterações nos níveis de expressão de ciclinas envolvidas na formação de complexos heterodiméricos Cdk/Ciclina.

O sítio ativo da Cdk na sua forma descomplexada encontra-se parcialmente ocluído.

Mudanças conformacionais que expõem o sítio de ligação a ATP ocorrem após a ligação da ciclina, aumentando a atividade quinase da Cdk. A ativação completa de complexos Cdk/Ciclina ocorre após a fosforilação de um resíduo de treonina no T-loop, localizado próximo ao sítio ativo, por uma quinase ativadora de Cdk (CAK) ²⁹ (Figura 3).

A atividade de complexos Cdk/Ciclina em etapas específicas do ciclo celular em mamíferos também está sujeita a um segundo mecanismo de regulação baseado no estado de fosforilação de resíduos consecutivos de tirosina e treonina localizados na região N-terminal da Cdk. A fosforilação desse segundo sítio de regulação é catalisada pela quinase Wee1 e inativa o complexo Cdk/Ciclina. Por outro lado, a desfosforilação, catalisada por fosfatases Cdc25, ativa o complexo Cdk/Ciclina^{30, 31, 32} (Figura 3). Dessa maneira, a atividade das fosfatases Cdc25 contribui para a progressão do ciclo celular pois estão envolvidas na ativação de complexos Cdk/Ciclina em etapas específicas.

Figura 3 – Regulação da atividade enzimática dos complexos Cdk/Ciclina. A completa ativação do complexo ocorre através da fosforilação de uma treonina no sítio ativo da Cdk. Resíduos consecutivos de tirosina e treonina localizados no N-terminal da Cdk estão sujeitos a fosforilação e compreendem um segundo mecanismo de controle da atividade enzimática do complexo Cdk/Ciclina.

Em humanos, estão presentes três isoformas da Cdc25 (Cdc25A, Cdc25B e Cdc25C) conservadas no domínio C-terminal catalítico e mais diversificadas no domínio regulatório N-terminal, que ainda apresenta variantes de *splicing*. O domínio regulatório tem sido relacionado ao controle da atividade, localização intracelular e controle do balanço entre síntese e degradação, mas seu papel específico em cada isoforma ainda não é conhecido³³.

Danos no DNA provocados por agentes mutagênicos ou durante a replicação do material genético podem promover a inativação ou degradação das Cdc25s e a consequente retenção do ciclo celular. Desta forma, as fosfatases Cdc25s desempenham um papel fundamental nos pontos de checagem do ciclo celular em resposta à instabilidade genômica^{33, 34}. Em particular, a Cdc25A participa da ativação de complexos Cdk2/Ciclina

E e Cdk2/Ciclina A na transição *G1/S* do ciclo celular, mas também possui importante papel na ativação do complexo Cdk1/Ciclina B para promover a transição *G2/M*. As Cdc25B e Cdc25C atuam principalmente na passagem *G2/M*, ativando o complexo Cdk1/Ciclina B³⁵. No entanto, também foi verificado que as três isoformas podem apresentar funções redundantes e atuar como reguladores em diferentes etapas do ciclo celular³².

O descontrole do ciclo celular é uma característica marcante das neoplasias e provoca um rápido aumento no número de células de um tumor. De fato, as Cdc25A e Cdc25B apresentam-se superexpressas em diversos tipos de tumores primários humanos, contribuindo para a contínua ativação de complexos Cdk/Ciclina e progressão do ciclo celular. Essa superexpressão está correlacionada com a agressividade tumoral e piores prognósticos da doença³⁵. Dessa maneira, os complexos Cdk/Ciclina e as fosfatases Cdc25 tornam-se importantes alvos para o desenvolvimento de novas drogas para o tratamento do câncer. Neste sentido, o estudo da estrutura conformacional desses alvos e a avaliação da ligação de pequenos inibidores podem fornecer dados importantes para o desenvolvimento racional dessas drogas.

A possibilidade de inibição da Cdc25B motivou estudos de *screenings* computacionais e experimentais de pequenas moléculas que inibem sua atividade³⁶. Alguns desses trabalhos caracterizaram inibidores competitivos^{37, 38, 39} ou que interferem na ligação entre a Cdc25B e um dos seus substratos naturais⁴⁰, o complexo Cdk2-pTpY/Ciclina A.

1.3 Descrições estruturais da Cdc25B

As Cdc25s pertencem à família das fosfatases duais pois estão envolvidas na desfosforilação de resíduos de Tyr e Thr/Ser. O domínio catalítico das três isoformas apresenta grande similaridade na sequência de aminoácidos (~60% em 200 resíduos) e possuem o motivo conservado His-Cys-(X)₅-Arg com a cisteína catalítica, que também é característico do sítio ativo de outras PTPs. A sequência (X)₅-Arg forma um sítio de ligação a fosfato (*P-loop*)⁴¹, onde X é um aminoácido qualquer, exceto prolina.

As estruturas cristalográficas do domínio catalítico da Cdc25A⁴² (resíduos 336–523; PDB: 1C25) e da Cdc25B⁴³ (res. 356–566; PDB: 1QB0) determinadas por difração de

raios-X mostraram-se muito similares (Figura 4): um domínio α/β formado por uma folha- β central composta por cinco fitas- β empacotadas por cinco α -hélices. Contudo, a porção C-terminal apresenta diferenças notáveis entre as isoformas. Na Cdc25A, os resíduos 488–496 estão numa conformação mais estendida e distante do núcleo proteico, associando-se com outra molécula na unidade assimétrica do cristal. A densidade eletrônica dos últimos 28 resíduos (496–523) não foi determinada, sugerindo desordem ou grande flexibilidade conformacional. Na Cdc25B, esta região corresponde ao segmento terminal 530–566. Contatos entre os resíduos 530–550 na Cdc25B e o núcleo proteico formam uma superfície que estende-se do sítio ativo até uma região que coordena um íon Cl^- no cristal. Os resíduos 534–546 na Cdc25B formam uma α -hélice que se localiza próxima ao núcleo proteico e que não é observada na Cdc25A. A densidade eletrônica dos últimos 16 resíduos (551–566) na Cdc25B não foi determinada, apontando para sua desordem conformacional.

Figura 4 – Sobreposição das estruturas cristalográficas da Cdc25A (magenta; PDB: 1C25) e Cdc25B (verde; PDB: 1QB0) determinadas por difração de raios-X. Acima estão representadas as construções utilizadas para a obtenção dos cristais e os respectivos segmentos cuja estrutura cristalográfica foi resolvida. À direita estão representados os sítios P1 e P2 de ligação a SO_4^{2-} presentes na Cdc25B. O íon Cl^- observado na estrutura cristalográfica da Cdc25B está representado por uma esfera amarela.

Diferentemente de outras PTP onde o sítio ativo encontra-se numa cavidade

profunda que possibilita somente a desfosforilação de resíduo de pTyr (e não pThr ou pSer), o sítio ativo da Cdc25B é consideravelmente raso. Esta característica está relacionada à sua natureza dual, uma vez que tanto resíduos de pTyr como pThr/pSer podem ser acomodados no sítio ativo⁴¹.

A hidrólise de resíduos de pTyr, pThr e pSer pelas Cdc25s é realizada em duas reações consecutivas que já foram intensamente estudadas^{44, 41, 45}. Na Cdc25B, a primeira etapa envolve o ataque nucleofílico da cisteína catalítica (473C) com a transferência de próton do ácido geral (474E) para o grupo de saída e a formação da fosfo-enzima (Reação 1.1). Na segunda etapa, a enzima livre é regenerada através da hidrólise da fosfo-enzima intermediária (Reação 1.2). Dessa maneira, dois resíduos catalíticos possuem importante papel no sítio ativo da Cdc25B: a cisteína 473 com o grupo tiolato desprotonado e o ácido glutâmico 474 com o grupo carboxila da cadeia lateral protonado.

Na estrutura cristalográfica da Cdc25B foram observados dois sítios (P1 e P2) de ligação a sulfato. No sítio P1 (P-loop), correspondente ao sítio ativo formado pelo motivo His-Cys-(X)₅-Arg (472–479), a complexação com SO_4^{2-} ocorre através de ligações de hidrogênio com a cadeia principal do segmento 474–478 e com a cadeia lateral de 479R. No sítio P2, as cadeias laterais dos resíduos 488R e 492R estão envolvidas na formação complexo (Figura 4). Sítios de ligação a SO_4^{2-} ou outro oxiânion não foram observados na Cdc25A⁴².

Mamonov *et al.*⁴⁶ utilizaram simulações de Monte Carlo baseadas em biblioteca de rotâmeros (LBMC) para a construção de um *ensemble* conformacional do domínio catalítico da Cdc25B que apresentou grande similaridade com a estrutura cristalográfica (Figura 5). Contudo, a hélice C-terminal nesse *ensemble* não encontra-se enovelada, sugerindo que este segmento da Cdc25B em solução estaria desordenado ou em equilíbrio com uma α -hélice. Observou-se também um número expressivo de estruturas onde a ponte dissulfeto entre as cisteínas 473 (catalítica) e 426 está formada, consistente o experimento de difração

de raios-X (PDB: 1CWR) onde esta ponte também foi observada⁴³. A formação desta ponte pode evitar a oxidação irreversível da cisteína catalítica na presença de espécies reativas de oxigênio e a consequente inativação permanente da enzima. Além disso, este pode ser um mecanismo de controle da atividade enzimática já que a Cdc25B oxidada (ponte dissulfeto formada) pode ser reativada pela via celular da tioredoxina redutase⁴⁷.

Figura 5 – Estruturas representativas do *ensemble* conformacional da Cdc25B obtido por simulação LBMC⁴⁶. A estrutura cristalográfica está representada em verde.

Em acordo com o *ensemble* gerado por simulação LBMC⁴⁶ e com a estrutura cristalográfica da Cdc25A⁴², dados de simulações computacionais por DM e métodos de previsão de desordem conformacional também sugeriram a instabilidade da hélice C-terminal da Cdc25B em solução⁴⁸. Esta simulação de DM com apenas 60 ns em solvente explícito foi iniciada das coordenadas cristalográficas da proteína. O núcleo proteico manteve-se estável ao longo de toda a simulação. No entanto, após 40 ns foi observado um desenovelamento parcial da hélice C-terminal, causando um aumento expressivo no desvio quadrático médio (RMSD) das coordenadas dessa região em relação à estrutura inicial (cristalográfica), e na flutuação quadrática média (RMSF) desse segmento em comparação com o núcleo proteico (Figura 6).

Devido à proximidade com o sítio ativo da Cdc25B, a influência da flexibilidade conformacional da porção C-terminal na afinidade de inibidores também foi avaliada

Figura 6 – Desestruturação da hélice C-terminal da estrutura cristalográfica durante simulação de DM⁴⁸. Observou-se um aumento no RMSD das coordenadas da porção C-terminal após 40 ns de simulação e valores mais elevados de RMSF dessa região em relação ao núcleo proteico.

por métodos computacionais⁴⁸. Modelos estruturais da Cdc25B foram representados por três conjuntos de estruturas com variados níveis de flexibilidade. Um modelo rígido foi representado pela estrutura cristalográfica, um modelo parcialmente flexível foi composto por estruturas obtidas da simulação de DM e um modelo totalmente flexível foi representado pelo *ensemble* conformacional gerado por simulação LBMC⁴⁶. Verificou-se que o modo de ligação dos inibidores é influenciado pela flexibilidade da cadeia principal da região C-terminal e pelo aparecimento de cavidades ausentes no modelo rígido. Dessa maneira, foi sugerido⁴⁸ que o modelo estrutural cristalográfico da Cdc25B poderia ser inadequado para modelar complexos com pequenas moléculas visto que a flexibilidade conformacional da proteína deve ser considerada na formação de complexos proteína-ligante (Seção 1.1).

Recentemente, Lund e Cierpicki⁴⁹ avaliaram a estrutura e a flexibilidade conformacional do domínio catalítico da Cdc25B através de medidas de RMN em solução. Foi utilizada uma construção que possuía somente os resíduos observados na estrutura cristalográfica (372–551), contendo uma mutação pontual na cisteína catalítica (C473S) para aumentar a estabilidade da proteína em solução. Dessa maneira, informações estruturais e dinâmicas do segmento formado pelos 15 resíduos do C-terminal (552–566) e do sítio ativo da enzima selvagem não foram avaliadas.

A mutação C473S, segundo os autores, não afetou consideravelmente a estrutura e dinâmica conformacional da Cdc25B⁴⁹ conforme comparação entre os deslocamentos químicos e as medidas de relaxação das proteínas mutante e selvagem. Contudo, notam-se

tempos de relaxação da magnetização transversal de ^{15}N no segmento 525–535 da Cdc25B nativa significativamente mais elevados que a proteína mutante (Figura 7). Esta diferença sugere a presença de movimentações lentas nessa região da proteína nativa que não são observadas no mutante C473S. O segmento 525–535 antecede a hélice C-terminal e forma uma fenda com o sítio ativo da enzima (Figura 4). Assim, a mutação da cisteína catalítica pode ter alterado a dinâmica conformacional dessa região da proteína em solução.

Figura 7 – Taxas de relaxação da magnetização transversal (R_2) dos *spins* de ^{15}N medidas para construções do domínio catalítico da Cdc25B selvagem e mutante (C473S)⁴⁹.

De maneira geral, Lund e Cierpicki⁴⁹ verificaram a grande similaridade entre as estruturas da Cdc25B cristalográfica e em solução⁴⁹, inclusive na porção C-terminal. Os CS e medidas de acoplamentos dipolar residual (RDC) da cadeia principal são consistentes com a presença da α -hélice no C-terminal. Medidas de relaxação também evidenciaram que esse segmento é relativamente rígido em solução para movimentações em diferentes escalas de tempo (ps–ns e μs –ms), em desacordo com as simulações computacionais que apontaram sua instabilidade^{46, 48} (Páginas 34 e 35).

Conclui-se que em todos os modelos estruturais do domínio catalítico da Cdc25B existentes atualmente não estão incluídos os 15 resíduos desordenados da região C-terminal (551–566). Demonstrou-se que essa região também participa do reconhecimento do substrato natural da Cdc25B, o complexo Cdk2-pTpY/Ciclina A bifosforilado. Particularmente, foi verificado através de mutações sítio-dirigidas que os resíduos 556R e 562R são importantes para a formação do complexo⁵⁰, tornando-se mais um exemplo do importante papel desempenhado por regiões terminais desordenadas no reconhecimento molecular⁵¹.

A compreensão da flexibilidade conformacional de proteínas flexíveis é relevante

para o desenvolvimento de inibidores, uma vez que é possível que ocorra o aparecimento de sítios crípticos, ausentes em estruturas cristalinas, e que os inibidores ou parceiros proteicos provoquem ou impeçam rearranjos conformacionais da proteína, bloqueando ou propiciando a formação do seu estado ativo.

A flexibilidade do C-terminal desordenado da Cdc25B pode influenciar a afinidade de pequenos inibidores competitivos devido à proximidade com o sítio ativo. Essa região também participa do reconhecimento do substrato proteico. Dessa maneira, a obtenção de modelos estruturais que incluam a flexibilidade conformacional do C-terminal pode auxiliar na compreensão dos mecanismos de reconhecimento molecular de proteínas e de pequenas moléculas por regiões intrinsecamente desordenadas. Nesse sentido, a fosfatase Cdc25B pode ser considerada um modelo para o estudo da influência da flexibilidade proteica no reconhecimento molecular. Além disso, a obtenção de um *ensemble* conformacional da Cdc25B pode permitir a estimativa da afinidade de pequenos ligantes por métodos computacionais⁵².

1.4 Objetivos

Devido à ausência de modelos estruturais da Cdc25B contendo todos os resíduos da região C-terminal e motivado pelas simulações computacionais que sugeriram a instabilidade da hélice C-terminal da Cdc25B^{46, 48}, o objetivo geral desta tese foi avaliar a estrutura e flexibilidade conformacional do domínio catalítico *completo* em solução através de técnicas computacionais e medidas experimentais de RMN. Torna-se importante salientar que este estudo foi iniciado anteriormente à publicação do trabalho de Lund e Cierpicki⁴⁹ em 2014.

Com o objetivo de se obter um modelo estrutural mais realista do domínio catalítico da Cdc25B completo, também foram avaliados os contatos existentes entre a região C-terminal e o núcleo proteico. Esses contatos podem auxiliar na previsão dos estados conformacionais populados pelo segmento C-terminal desordenado em solução.

2 Medidas espectroscópicas de RMN

2.1 Metodologia

2.1.1 Clonagem das construções *C1* e *C2*

O domínio catalítico da fosfatase Cdc25B de humanos (resíduos 373–566; NCBI GenBank: AAA58416.1) foi clonado a partir do plasmídeo pUC18 contendo o cDNA da proteína completa (domínio regulatório + domínio catalítico) adquirido da empresa OriGene (código SC126594). Foram amplificadas duas construções (*C1* e *C2*) a partir de *primers* diretos (FW) e reversos (RV) contendo sequências adaptadoras e sítios de clivagem para as enzimas de restrição NdeI e BamHI, respectivamente (Tabela 1). As construções *C1* e *C2* referem-se ao domínio catalítico da Cdc25B completo e com 32 resíduos truncados no C-terminal, respectivamente. Desta forma, a construção *C2* não possui a α -hélice C-terminal (534–546) e o segmento 547–566, onde também estão incluídos os resíduos que não apresentaram densidade eletrônica no experimento de difração de raios-X (Figura 4).

Tabela 1 – *Primers* diretos (FW) e reversos (RV) utilizados na clonagem das construções *C1* e *C2* da Cdc25B. Sítios de restrição para NdeI e BamHI estão representadas em azul e o códon de parada de tradução em vermelho.

Construções	Resíduos	Massa (kDa)	FW	RV
<i>C1</i>	373–566	23,7	5'–GGA ATT CCA TAT GGA ATT CCA GAG TGA CCA CCG AGA GCT GAT TGG A–3'	5'–CGG GAT CCC GTC ACT GGT CCT GCA GCC GGC TAC A–3'
<i>C2</i>	373–534	19,8	5'–GGA ATT CCA TAT GGA ATT CCA GAG TGA CCA CCG AGA GCT GAT TGG A–3'	5'–CGG GAT CCC GTC ACT CGT GGT TCA TGG GCC GGT AGT C–3'

A reação de PCR foi realizada em volume final de 50 μ l [1 μ l *primers* FW e RV (10 μ M); 1 μ l do plasmídeo pUC18 contendo o cDNA da Cdc25B (100 ng/ μ l); 1 μ l

dNTPs (10 mM); 1,5 µl MgCl²⁺ (50 mM); 5 µl de tampão para PCR 10×; 1 µl InvitrogenTM PlatinumTM Taq polimerase 2,5 U/µl; 38,5 µl H₂O] com os ciclos de temperatura apresentados na tabela 2.

Tabela 2 – Ciclos de temperatura da reação de PCR para amplificação do cDNA das construções C1 e C2.

Temperatura (°C)	95	95	66	72	72	4
Tempo (min)	4:00	1:00	1:00	1:00	10:00	∞
Repetições	1		35		1	

Os produtos da reação de PCR foram submetidos a eletroforese em gel de agarose 1% (p/v) e corados com brometo de etídio. As bandas referentes aos produtos amplificados foram recortadas do gel e purificadas com o kit “Wizard[®] SV Gel and PCR Clean-Up System” (Promega), seguindo instruções do fabricante. Posteriormente, os produtos da amplificação por PCR foram inseridos no vetor de clonagem pGEM-T Easy (Promega), seguindo instruções do fabricante, com uma relação molar de 3:1 (inserto:vetor).

Para a amplificação do plasmídeo pGEM contendo os insertos, bactérias competentes *E. coli* XL1-Blue (Stratagene) foram transformadas adicionando-se 4 µl da reação de ligação a uma alíquota de 200 µl das bactérias competentes. Esta preparação foi então mantida por 30 min no gelo seguidos de um choque térmico de 90 s a 42 °C. Após esta etapa, as células foram mantidas no gelo por mais 3 min e posteriormente adicionados 500 µl de meio LB, seguido de incubação a 37 °C por 1 hora sob agitação para a recuperação das bactérias. Alíquotas de 100 µl e 10 µl foram plaqueadas em LB-ágar contendo ampicilina 50 µg/ml, 100 µl IPTG 100 mM e 20 µl X-gal 50 mg/ml. As placas foram incubadas a 37 °C por ~ 16 h para possibilitar o crescimento das colônias.

Algumas colônias brancas foram selecionadas para crescimento *overnight* em 5 ml de meio LB com ampicilina 50 µg/ml. Alíquotas de 1 µl do meio foram então utilizadas numa reação de PCR (Tabela 2) com os primers do inserto [0,5 µl dos *primers* FW e RV (10 µM); 1 µl do meio de cultivo ou 0,2 µl cDNA da Cdc25B (100 ng/µl); 10 µl “PCR Master Mix” (Thermo Scientific); 8 µl H₂O] com o intuito de avaliar a presença do plasmídeo com o inserto de interesse nas colônias selecionadas. A purificação do plasmídeo foi feita com o kit “PureYieldTM Plasmid Miniprep System” (Promega), seguindo instruções do fabricante.

Para avaliar a ocorrência mutações na sequência de nucleotídeos do inserto de interesse, uma reação de PCR para o sequenciamento dos plasmídeos pelo método de Sanger⁵³ foi realizada com os *primers* SP6 (5'-TAT TTA GGT GAC ACT ATA G-3') ou T7 (5'-TAA TAC GAC TCA CTA TAG GG-3') em 15 µl de reação [200 ng do plasmídeo; 3 µl primer (3,3 µM); 3 µl tampão de sequenciamento BigDye® 5×; 2 µl BigDye® v3.1 (Invitrogen); H₂O para completar 15 µl], utilizando-se o seguinte ciclo de temperaturas:

Temperatura (°C)	96	96	50	60	4
Tempo (min)	2:00	0:45	0:30	4:00	∞
Repetições	1		35		1

Os produtos do PCR para sequenciamento foram precipitados adicionando-se 25 µl da solução de precipitação [2,7 ml EtOH; 110 µl CH₃COONa pH 5,2; 110 µl glicogênio 1 mg/ml] aos 15 µl da reação de PCR. Esta reação foi mantida no gelo por 1 h e então centrifugada a 4.000 RPM por 30 min a 4°C. O sobrenadante foi descartado e em seguida dois pulsos de 1.000 RPM foram aplicados com a placa de sequenciamento invertida para a drenagem do restante de solução. Foram então adicionados 50 µl de EtOH 70% gelado e a placa submetida novamente aos passos de centrifugação e drenagem. Após esta etapa, a placa foi colocada no termociclador com a tampa aberta por 1 min a 95°C para secar as amostras e mantida a -20°C. O sequenciamento foi realizado no Instituto de Química da USP em um equipamento ABI PRISM® 3100 Genetic Analyzer (HITACHI).

Os insertos de interesse das construções *C1* e *C2* confirmados pelo sequenciamento de DNA foram retirados do vetor de clonagem pGEM-T Easy e transferidos para o vetor de expressão pET-28a (Figura 8). A linearização do vetor de expressão e a digestão dos insertos nos vetores de clonagem foram realizadas em 30 µl de reação [500 ng do plasmídeo; 0,5 µl de BamHI (10 U/µl); 0,5 µl de NdeI FastDigest™ (Fermentas); 3 µl tampão “Green FastDigest” 10×; H₂O MilliQ para 30 µl], mantidas por 1 h a 37°C. Os produtos de digestão foram purificados por eletroforese em gel de agarose 1% com o kit “Wizard® SV Gel and PCR Clean-Up System”, ressuspensos em 30 µl de H₂O e mantidos a -20°C.

A reação de ligação dos insertos ao vetor de expressão pET-28a foi preparada com 10 µl do inserto digerido, 7 µl do plasmídeo linearizado, 2 µl tampão DNA ligase 10× e 1 µl

Figura 8 – Mapa do plasmídeo pET-28a com detalhes da região de clonagem.

de T4 DNA ligase (Fermentas). Esta solução foi mantida *overnight* a 16°C e então utilizada para transformar bactérias competentes *E. coli* XL-1 Blue (Seção 2.1.1). As bactérias transformadas foram plaqueadas em placas de ágar contendo LB e canamicina 30 µg/ml e mantidas a 37°C por ~ 24 horas. Algumas colônias foram selecionadas e utilizadas para confirmar a presença do inserto de interesse através de PCR (Seção 2.1.1). As colônias que possuíam os plasmídeos pET-28a_C1 ou pET-28a_C2, foram inoculadas em 20 ml de meio líquido LB+canamicina 30 µg/ml, mantidas *overnight* a 37°C e o *pellet* utilizado para a purificação do plasmídeo com o kit “PureYield™ Plasmid Miniprep System” (Promega), seguindo instruções do fabricante.

2.1.2 Expressão e purificação

A expressão das construções *C1* e *C2* da Cdc25B foi realizada na bactéria *E. coli* BL21(GOLD)-DE3 (Stratagene) transformadas (Seção 2.1.1) com os vetores de expressão pET-28a_C1 ou pET-28a_C2 que possuem um gene para resistência a canamicina. No vetor pET-28a, as construções são expressas com uma cauda de His₆ na região N-terminal seguida por um sítio de reconhecimento por trombina que possibilita sua remoção (Figura 9).

Figura 9 – Sequência primária da região N-terminal ($\sim 1,9$ kDa) das construções *C1* e *C2*. Na figura estão destacados a cauda de His₆ e o sítio de reconhecimento por trombina.

Uma alíquota de 4 ml do pré-inóculo em meio LB com canamicina 30 μ g/ml crescido *overnight* a 37°C foi adicionada em 1 l de meio mínimo M9 [6 g/l Na₂HPO₄; 3 g/l KH₂PO₄; 1 g/l NH₄Cl; 0,5 g/l NaCl; 3 mg/l CaCl₂; 2 ml/l MgSO₄ 1 M; 2 g/l glicose] contendo canamicina 30 μ g/ml. O meio foi mantido a 20°C sob agitação (200 RPM) até ser atingida fase exponencial de crescimento (DO_{600nm} $\sim 0,6$). Neste momento foi adicionado IPTG, para induzir a expressão da proteína, a uma concentração final de 0,4 mM e o meio mantido por 18 h a 20°C sob agitação. Ao final do período de indução, o *pellet* de bactérias, coletado por centrifugação (4.000 RPM por 30 min a 4°C), foi ressuspensionado em tampão de lise [Tris 20 mM; NaCl 500 mM; glicerol 10%; BMER 5 mM; PMSF 1 mM; pH 7,4] e mantido a -80°C antes da etapa de purificação. Foram utilizados 10 ml do tampão de lise para ressuspender o *pellet* de bactérias obtidos de 1 l de meio de cultivo.

Para expressão das construções marcadas com ¹³C e/ou ¹⁵N foram adicionados ao meio mínimo ¹³C-glicose e/ou ¹⁵NH₄Cl. A construção *C1*, marcada com ¹³C e ¹⁵N, utilizada na aquisição de espectros de RMN para o assinalamento da cadeia principal foi expressa em meio mínimo preparado com água deuterada (²H₂O), o que promove a deuteração aleatória e parcial da proteína.

O *pellet* de bactérias ressuspensionado em tampão de lise foi submetido a sonicação em gelo, utilizando-se 10 pulsos de 10 s a 75 W de potência, para o rompimento das células.

O sobrenadante contendo as proteínas solúveis foi coletado após centrifugação do lisado de bactérias a $16.000 \times g$ durante 30 min a 4°C .

A primeira etapa de purificação da proteína recombinante foi realizada em coluna de vidro empacotada manualmente com a resina Ni-NTA agarose (Quiagen). O sobrenadante (~ 14 ml) obtido após a lise das bactérias de 1 l do meio cultura foi incubado à temperatura ambiente com 2 ml da resina durante 20 min para possibilitar a interação da cauda de His₆ com os centros de Ni²⁺ da resina. Após essa etapa, a coluna foi lavada com 40 ml (4×10 ml) do tampão A [Tris 20 mM; NaCl 500 mM; imidazol 20 mM; BMER 5 mM; pH 7,4] e a eluição da proteína recombinante foi realizada em três etapas com 7,5 ml ($4 + 2 + 1,5$ ml) de tampão B [Tris 20 mM; NaCl 500 mM; imidazol 500 mM; BMER 5 mM; pH 7,4].

O material eluído da coluna Ni-NTA agarose foi submetido a gel filtração em coluna HiTrap Desalting (GE Healthcare) de 5 ml, previamente equilibrada com 40 ml de tampão fosfato [NaPi 50 mM; NaCl 100 mM; BMER 5 mM; EDTA 1 mM; pH 6,75], para diminuir a concentração de NaCl e remover o imidazol da solução com o objetivo de aumentar a eficiência na clivagem enzimática da cauda de His₆. Nesta etapa, realizada manualmente, a cada 1,5 ml de solução aplicados na coluna, foram coletados 2 ml.

Após a troca do tampão por gel filtração foram adicionados 100 μl da resina trombina-agarose (Sigma Aldrich) e a solução foi mantida por 3 h a temperatura ambiente sob agitação branda. A clivagem da cauda de histidina (1,9 kDa) foi confirmada por gel de SDS-PAGE antes da proteína ser submetida a uma segunda etapa de purificação.

A segunda etapa de purificação da proteína foi realizada por cromatografia líquida num equipamento FPLC Akta Purifier (GE Healthcare) equipado com uma coluna preparativa de gel filtração Hiloal 26/600 Superdex 75 (GE Healthcare), previamente equilibrada com 600 ml de tampão fosfato [NaPi 50 mM; NaCl 100 mM; BMER 1 mM; EDTA 1 mM; pH 6,75] a um fluxo de 0,8 ml/min. As amostras de proteína obtidas após a clivagem da cauda de His₆ foram concentradas em filtros Amicon Ultra-15 3 kDa (Millipore) até um volume final de 4,5 ml antes de serem aplicadas na coluna. A cromatografia foi então conduzida sob um fluxo de 0,8 ml/min com um volume total de ~ 320 ml de tampão fosfato. Frações de 4 ml foram coletadas automaticamente e avaliadas por SDS-PAGE

```

*****
C1 366 GSHMEFQSDHRELIGDYSKAFLLQTVDGKHQDLKYISPETMVALLTGKFSNIVDKFVIVDCRYPYEYEGG 435
C2 366 GSHMEFQSDHRELIGDYSKAFLLQTVDGKHQDLKYISPETMVALLTGKFSNIVDKFVIVDCRYPYEYEGG 435

*****
C1 436 HIKTAVNLPLERDAESFLLKSPIAPCSLDKRVILIFHCEFSSESGPRMCRFIRERDRAVNDYPSLYPEM 505
C2 436 HIKTAVNLPLERDAESFLLKSPIAPCSLDKRVILIFHCEFSSESGPRMCRFIRERDRAVNDYPSLYPEM 505

*****
C1 506 YILKGGYKEFFPQHNPFCPEQDYRPMNHEAFKDELKTFRKTRSWAGERSRRELCSRLQDQ 566
C2 506 YILKGGYKEFFPQHNPFCPEQDYRPMNHE 534

```

Figura 10 – Sequência primária das construções *C1* e *C2*.

quanto à sua homogeneidade.

A sequência de aminoácidos das construções *C1* e *C2* da Cdc25B obtidas após a clivagem da cauda de His₆ estão representadas na figura 10. O N-terminal das construções possui um segmento extra (366–GSHMEFQ–372) referente aos resíduos: (i) remanescentes após a clivagem com trombina (Figura 9); (ii) codificados nas sequências adaptadoras presentes no *primer* direto usado na amplificação do cDNA (Seção 2.1.1; Tabela 1).

2.1.3 RMN

Os experimentos de RMN foram realizados com o objetivo de assinalar os átomos da cadeia principal das construções *C1* e *C2* da Cdc25B, além de medir acoplamentos dipolares residuais (RDC), taxas de relaxação (T_1 e T_2) de ^{15}N e NOE heteronucleares $\{^1\text{H}\} - ^{15}\text{N}$ da cadeia principal da construção *C1*.

Todos os espectros foram adquiridos a 298 K em um magneto Bruker AscendTM 18.8 T (800 MHz na frequência de próton) equipado com uma sonda criogênica TCI e um console Avance III.

2.1.3.1 Assinalamento da cadeia principal

Os espectros de RMN utilizados no assinalamento da cadeia principal foram adquiridos de amostras de 550 μl acondicionadas em tubos de 5 mm com $\sim 200 \mu\text{M}$ das construções *C1* ou *C2* em tampão fosfato de sódio [NaPi 20 mM; NaCl 50 mM; DTT 5 mM; BMER 5 mM; EDTA 1 mM; 5% $^2\text{H}_2\text{O}$; pH 6,75], marcadas isotopicamente com ^{13}C e ^{15}N . A construção *C1* também foi deuturada parcial e aleatoriamente (Seção 2.1.2).

Foram adquiridos espectros heteronucleares tridimensionais (HNCA, HN(CO)CA,

HNCO, HN(CA)CO, HN(CA)CB, HN(CO)CACB) e bidimensionais (HSQC) $^1\text{H}-^{15}\text{N}$ para o assinalamento da cadeia principal utilizando-se sequências de pulsos convencionais⁵⁴. Os espectros foram processados com o programa NMRPipe⁵⁵ e o assinalamento da cadeia principal da proteína feito com o auxílio do programa CCPNMR Analysis 2.4⁵⁶.

Os espectros de $^1\text{H}-^{15}\text{N}$ HSQC foram adquiridos com 1024×128 pontos complexos e centro ajustado para 4,7 ppm e 118 ppm nas dimensões direta (^1H) e indireta (^{15}N), respectivamente. Foram utilizadas janelas espectrais de 18 ppm na dimensão de ^1H e 44 ppm na dimensão de ^{15}N .

Todos os experimentos de tripla ressonância foram adquiridos com $768 \times 40 \times 40$ pontos complexos nas dimensões de $^1\text{H} \times ^{15}\text{N} \times ^{13}\text{C}$, exceto o espectro de tripla correlação HNCO, adquirido com $1024 \times 40 \times 40$ pontos complexos. Foram utilizadas janelas espectrais de 18 ppm na dimensão de ^1H e 44 ppm na dimensão de ^{15}N com o centro do espectro em 4,7 ppm e 18 ppm, respectivamente. As janelas espectrais, centros dos espectros na dimensão de ^{13}C e número de *scans* variaram de acordo com o tipo de experimento:

Experimento	Centro	Janela espectral	Scans
HNCA	54 ppm	40 ppm	16
HN(CO)CA	54 ppm	40 ppm	32
HNCACB/HN(CO)CACB	39 ppm	75 ppm	32
HNCO	176 ppm	30 ppm	16
HN(CA)CO	176 ppm	30 ppm	32

Ao final do processo de assinalamento da Cdc25B, os deslocamentos químicos (CS) dos núcleos assinalados nas construções *C1* e *C2* foram referenciados segundo procedimentos adotados por Zhang *et al.*⁵⁷, utilizando o programa Shiftcor 1.0⁵⁸, e usados na previsão de: (i) ângulos diedrais da cadeia principal, utilizando o programa TALOS-N⁵⁹; (ii) desordem conformacional através do cálculo do índice de desordem conformacional (RCI)⁶⁰; (iii) estruturas secundárias através do cálculo dos índices do deslocamento químico (CSI)⁶¹.

Com o intuito de confirmar o assinalamento do sítio ativo (*P-loop*), espectros de $^1\text{H}-^{15}\text{N}$ HSQC foram adquiridos durante titulação com 0, 1, 2, 4, 8 e 16 mM de fosfato

inorgânico (Pi) de uma amostra de 50 μ M da construção *C1* inicialmente em tampão MES [MES 20 mM; NaCl 50 mM; DTT/EDTA 1 mM; BMER 5 mM; 5% $^2\text{H}_2\text{O}$; pH 6,75]. A titulação foi feita a partir de uma solução estoque de tampão NaPi 500 mM, NaCl 50 mM, pH 6,75.

Perturbações do deslocamento químico (CSP) de grupos amida da cadeia principal foram calculadas por:

$$CSP = \sqrt{(\Delta\delta_{1H})^2 + 0.15(\Delta\delta_{15N})^2}, \quad (2.1)$$

onde $\Delta\delta_X$ é a diferença entre os deslocamentos químicos de $X = ^1\text{H}$ ou ^{15}N .

2.1.3.2 Previsão de ângulos diedrais φ/ψ da cadeia principal

Os ângulos diedrais φ/ψ da cadeia principal da construção *C1* da Cdc25B em solução foram previstos a partir dos deslocamentos químicos assinalados com o uso do programa TALOS-N⁵⁹. Este programa baseia-se no uso de rede neural artificial (ANN) de duplo nível treinada sobre um banco de dados composto por proteínas que possuem estrutura cristalográfica de alta resolução e deslocamentos químicos da cadeia principal medidos experimentalmente ou previstos pelo programa SPARTA+⁶².

No segundo nível da ANN, uma matriz de densidade com 324 (18×18) estados torcionais dos ângulos φ e ψ do resíduo central de pentapeptídeos da sequência primária da proteína de interesse ($D_{1 \times 324}^{\varphi/\psi}$) é obtida utilizando-se uma matriz com 1620 nós ($X_{1 \times 1620}$) com os deslocamentos químicos secundários e índices de similaridade com os pentapeptídeos do banco de dados:

$$D_{1 \times 324}^{\varphi/\psi} = f_2(f_1(X_{1 \times 1620} \times W_{1620 \times 360}^1 + b_{1 \times 360}^1) \times W_{360 \times 324}^2 + b_{1 \times 324}^2),$$

onde $f_1(x) = (1 - e^{-2x})/(1 + e^{-2x})$, $f_2(x) = x$, $\mathbf{W}^1$ e $\mathbf{b}^1$ são matrizes peso e tendência, respectivamente, para a conexão entre as camadas de entrada e intermediária e $\mathbf{W}^2$ e $\mathbf{b}^2$ são matrizes semelhantes para a conexão entre as camadas intermediária e de saída.

A matriz de densidade $D_{1 \times 324}^{\varphi/\psi}$ é obtida após otimização das matrizes peso e tendência com um conjunto de treino. Então, os 25 melhores acordos entre os pentapeptídeos da proteína de interesse e do banco de dados são utilizados na previsão dos ângulos torcionais. Caso todos estejam agrupados em uma única região do gráfico de Ramachandran, a

previsão é única e obtidas pela média dos valores ponderada por um fator de similaridade entre o segmento da proteína de interesse e os 25 segmentos do banco de dados com melhor acordo. Caso as melhores concordâncias estejam agrupadas em duas regiões distintas do gráfico de Ramachandran, a previsão é ambígua. A previsão não é feita para resíduos que apresentam altos valores de RCI⁶⁰.

2.1.3.3 Previsão de desordem conformacional

O índice de desordem conformacional (RCI) foi calculado através do método desenvolvido por Berjanskii e Wishart⁶⁰:

$$RCI = (\kappa|\Delta\delta_{C\alpha}| + \lambda|\Delta\delta_{C'}| + |\Delta\delta_{C\beta}| + \mu|\Delta\delta_N| + \nu|\Delta\delta_{H\alpha}|)^{-1}$$

onde os coeficientes $\kappa = 4,8$, $\lambda = 4,8$, $\mu = 3,96$ e $\nu = 5,96$ foram parametrizados para ajustar os valores experimentais de deslocamento químico secundário ($\Delta\delta$) com flutuações observada em simulações de dinâmica molecular de 1585 resíduos distribuídos 14 proteínas-modelo.

2.1.3.4 Acoplamento dipolar residual

O acoplamento dipolar entre dois núcleos é um efeito causado pela influência do campo magnético de um núcleo sobre o momento magnético de outro núcleo, e pode ser expresso por ^{63, §16.3}:

$$D_{ij} = D_{max} \frac{1}{2} (3 \cos^2 \zeta - 1), \quad (2.2)$$

onde $D_{max} = h\gamma_i\gamma_j/2\pi^2r_{ij}^3$, γ é a constante giromagnética, r_{ij} é a distância entre os núcleos i e j e ζ é o ângulo formado entre o vetor internuclear $\vec{ij}$ e o campo magnético estático $\vec{B}_0$.

Devido ao rápido tombamento das moléculas em solução, o acoplamento dipolar resultante é nulo pois a probabilidade do vetor $\vec{ij}$ assumir um valor ζ é a mesma para todos os ângulos, fazendo com que $D_{ij} = 0$:

$$\int_0^\pi \frac{1}{2} (3 \cos^2 \zeta - 1) \sin \zeta d\zeta = \left. \frac{-\cos^3 \zeta}{2} + \frac{\cos \zeta}{2} \right|_0^\pi = 0. \quad (2.3)$$

Contudo, em condições anisotrópicas do meio quando uma pequena fração ($\sim 1/1000$) das moléculas está alinhada em relação ao campo $\vec{B}_0$, observa-se um acoplamento dipolar residual (RDC) que pode ser expresso em Hz a partir da equação (2.2):

$$D_{ij}(\theta, \phi) = D_{max} \left[A_a \frac{1}{2} (3 \cos^2 \theta - 1) + A_r \frac{3}{4} \sin^2 \theta \cos 2\phi \right], \quad (2.4)$$

onde θ e ϕ são ângulos polares do vetor internuclear $\vec{ij}$ no sistema de referência do tensor de alinhamento e $A_a = \frac{3}{2}A_{zz}$ e $A_r = A_{xx} - A_{yy}$. A_{zz} , A_{yy} e A_{xx} são componentes diagonais da matriz do tensor de alinhamento com $|A_{zz}| > |A_{yy}| > |A_{xx}|$.

A equação equação (2.4) pode ser simplificada para:

$$D_{ij}(\theta, \phi) = D_a \left[(3 \cos^2 \theta - 1) + \frac{3}{2}R \sin^2 \theta \cos 2\phi \right], \quad (2.5)$$

onde $D_a = \frac{3}{4}D_{max}A_{zz}$ é a magnitude do tensor de alinhamento e a rombicidade $R = 2(A_{xx} - A_{yy})/3A_{zz}$.

RDCs de grupos amida da cadeia principal da construção *C1* foram medidos em meio líquido-cristalino formado pela mistura polietilenoglicol (PEG) C12E5/ η -hexanol/ H_2O , conforme procedimentos descritos por Rückert e Otting⁶⁴. Resumidamente, pequenas alíquotas de η -hexanol (0,2–1 μ l), seguidas de agitação em *vórtex*, foram adicionadas a amostras de 500 μ l com 200 μ M da proteína em tampão fosfato [NaPi 20 mM; NaCl 50 mM; BMER 5 mM; 5% 2H_2O ; pH 6,75], contendo C12E5 5% (v:v) até a obtenção da fase líquido-cristalina. No total, foram adicionados à amostra um 7,4 μ l de η -hexanol.

Os valores de RDCs (D_{NH}) foram obtidos pela diferença entre a separação das duas linhas do sinal de ^{15}N medidos em meio isotrópico (na ausência do meio de alinhamento) e anisotrópico através de experimentos *in-phase/anti-phase* (IPAP)⁶⁵:

$$D_{NH} = ({}^1J_{NH} + D_{NH})^{aniso} - {}^1J_{NH}^{iso}. \quad (2.6)$$

Os D_{NH} observados e a estrutura cristalográfica da Cdc25B (PDB: 1QB0)⁴³ foram utilizados para a obtenção do tensor de alinhamento com o auxílio do programa DC contido no pacote de programas NMRPipe⁵⁵. Neste cálculo, os elementos da matriz de Saupe⁶⁶ foram obtidos através da decomposição em valores singulares⁶⁷, considerando-se somente os D_{NH}^{obs} cujos átomos de ^{15}N apresentaram fatores B cristalográficos abaixo de um

desvio padrão da média ($B_N < \overline{B_N} + \sigma$). As coordenadas dos átomos de hidrogênio foram construídas na estrutura cristalográfica da Cdc25B (PDB: 1QB0) com a ferramentas do pacote de programas GROMACS 4.5.3⁶⁸ e parâmetros do campo de força Amber99SB-ILDN⁶⁹. A energia do sistema foi minimizada em vácuo até uma força máxima inferior a 1000 kJ/mol · nm, a fim de otimizar a orientação dos vetores NH da cadeia principal da proteína. Para restringir a posição dos átomos pesados foi aplicada uma constante harmônica de 1000 kJ/mol · nm².

2.1.3.5 Taxas de relaxação R_1 e R_2 e NOE heteronuclear $\{^1\text{H}\}-^{15}\text{N}$

Medidas de relaxação longitudinal R_1 e no referencial girante $R_{1\rho}$ dos grupos amida da cadeia principal, foram calculadas a partir de experimentos pseudo-3D de $^1\text{H}-^{15}\text{N}$ HSQC, cuja terceira dimensão foi utilizada somente para a aquisição de 12 espectros com diferentes intervalos de relaxação. Para medir R_1 foram utilizados intervalos de relaxação de 0,1, 0,2, 0,3, 0,5, 0,6, 0,8, 0,9, 1,2, 1,4, 1,6, 1,8 e 2,0 segundos e para medir $R_{1\rho}$ foram utilizados intervalos de 0,010, 0,024, 0,040, 0,054 ($2\times$), 0,080, 0,096, 0,112, 0,124, 0,140, 0,160 e 0,180 segundos. Todos espectros foram adquiridos com 16 *scans* coletados em intervalos de 3 s.

O comportamento da intensidade dos picos (S) em função dos diferentes intervalos de relaxação τ utilizados na aquisição dos espectros foi ajustado a uma função de decaimento exponencial para a obtenção das taxas de relaxação:

$$S(\tau) = S(0) [\exp(-R_z\tau)], \quad (2.7)$$

onde $z = 1, 1\rho$ e $S(0)$ é a intensidade do sinal no tempo $\tau = 0$.

A incerteza de S foi obtida pelo desvio padrão da linha base do espectro no cálculo de R_1 e pela diferença entre as medidas obtidas na aquisição de dois espectros com o mesmo tempo de relaxação (0,054 s) no cálculo de $R_{1\rho}$. A incerteza das medidas de relaxação R_1 e $R_{1\rho}$ foi obtida através de 100 interações de Monte Carlo.

Valores de R_2 foram obtidos a partir das medidas de $R_{1\rho}$, obtidas com sequência de pulsos otimizada para o emprego de *spinlock* fraco (até 1 kHz)⁷⁰, de acordo com a seguinte

expressão:

$$R_{1\rho} = R_1 \cos^2 \theta + R_2 \sin^2 \theta, \quad (2.8)$$

onde $\theta = \arctan \frac{\Omega_1}{\Omega}$ é o ângulo de inclinação entre o campo magnético efetivo e o eixo z , Ω_1 a frequência de precessão ao redor do campo de *spinlock* (B_1) e Ω o *offset* do núcleo.

Valores mais confiáveis (menores erros) de R_2 foram obtidos com a medida de $R_{1\rho}$ em três diferentes valores na frequência carreadora de ^{15}N : 113, 119,5 e 128 ppm. Os valores coincidentes entre dois conjuntos de dados foram considerados para as análises, sendo que aquele que apresentou menor χ^2 foi selecionado.

Medidas de NOE heteronuclear dos *spins* de ^{15}N ($\{^1\text{H}\} - ^{15}\text{N}$) foram obtidas pela relação entre a intensidade dos sinais de espectros bidimensionais $^1\text{H} - ^{15}\text{N}$ adquiridos na presença ou na ausência de saturação dos *spins* de prótons:

$$\text{NOE} = \frac{S_z}{S_z^o}, \quad (2.9)$$

onde S_z e S_z^o são as intensidades dos sinais na presença ou na ausência de saturação em ^1H , respectivamente.

2.1.3.6 Cálculo dos parâmetros dinâmicos pelo formalismo de Lipari-Szabo

Os principais mecanismos de relaxação de *spins* de ^{15}N em uma proteína são a interação dipolar com o próton diretamente ligado e a anisotropia do deslocamento químico ($\Delta\sigma$). Esse fenômeno é provocado por flutuações estocásticas (movimento Browniano) da interação dipolar e do deslocamento químico causadas por movimentações internas e pelo tombamento da molécula em solução. Toda a informação sobre a dinâmica de reorientação molecular está contida na função de densidade espectral ($J(\omega)$). Dessa forma, R_1 , R_2 e NOE podem ser descritos em função de $J(\omega)$ amostrada em frequências específicas associadas à transição entre os diferentes estados do *spin* nuclear^{63, §19.3} e da perda de

coerência da magnetização transversal no caso de R_2 :

$$R_1 = \frac{d^2}{10} [J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N)] + \frac{2}{15} \omega_N^2 \Delta\sigma^2 J(\omega_N) \quad (2.10)$$

$$R_2 = \frac{d^2}{20} [4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)] \\ + \frac{1}{45} \omega_N^2 \Delta\sigma^2 [4J(0) + 3J(\omega_N)] + R_{ex} \quad (2.11)$$

$$\text{NOE} = 1 + \frac{\gamma_H}{\gamma_N} \frac{d^2}{R_1} \frac{1}{10} [6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)], \quad (2.12)$$

onde $d = (h\gamma_H\gamma_N/\pi r_{\text{NH}}^3)$; h é a constante de Planck; γ_H e γ_N são as razões giromagnéticas de ^1H e ^{15}N , $r_{\text{NH}} = 1.02\text{\AA}$ é a distância da ligação $^1\text{H}-^{15}\text{N}$, ω_H e ω_N são as frequências de Larmor de ^1H e ^{15}N , $\Delta\sigma = -160$ ppm, e R_{ex} descreve a contribuição de trocas conformacionais que ocorrem na escala de tempo de μs – ms para a relaxação da magnetização transversal, R_2 .

Para uma macromolécula, pode-se assumir que os movimentos internos e os movimentos de reorientação global (difusão rotacional) são independentes. Neste caso, é possível decompor $J(\omega)$ em τ_m e τ_c , conforme proposto no formalismo de Lipari e Szabo^{71, 72}. Nesta abordagem, a amplitude das movimentações intramoleculares na escala de ps–ns é descrita pelo parâmetro de ordem S^2 . Para um tensor de difusão rotacional isotrópico, $J(\omega)$ pode ser expressa por:

$$J(\omega) = \frac{2}{5} \left[S^2 \frac{\tau_c}{1 + \omega^2 \tau_c^2} + (1 - S^2) \frac{\tau_m}{1 + \omega^2 \tau_m^2} \right], \quad (2.13)$$

onde $\tau_m = \tau_i \tau_c / (\tau_i + \tau_c)$ e τ_c e τ_i são os tempos de correlação global e de movimentações internas, respectivamente.

O formalismo de Lipari-Szabo foi posteriormente estendido por Clore *et al.*⁷³ para a inclusão de termos que descrevem movimentações internas em duas escalas de tempo:

$$J(\omega) = \frac{2}{5} \left[\frac{S^2 \tau_c}{1 + (\omega \tau_c)^2} + \frac{(1 - S_f^2) \tau_f'}{1 + (\omega \tau_f')^2} + \frac{(S_f^2 - S^2) \tau_s'}{1 + (\omega \tau_s')^2} \right], \quad (2.14)$$

onde $\tau_f' = \tau_f \tau_c / (\tau_f + \tau_c)$ e $\tau_s' = \tau_s \tau_c / (\tau_s + \tau_c)$, τ_f e τ_s ($\tau_f < \tau_s \ll \tau_c$) são tempos de correlação para movimentações internas rápidas e lentas, respectivamente, $S^2 = S_f^2 S_s^2$ é o parâmetro de ordem generalizado e S_f^2 e S_s^2 são os parâmetros de ordem para movimentações rápidas e lentas, respectivamente.

Na presença de um tensor de difusão rotacional axialmente simétrico, a equação (2.14) pode ser reescrita com os componentes do tensor:

$$J(\omega) = \frac{2}{5} \sum_{j=1}^3 A_j \left[\frac{S^2 \tau_j}{1 + (\omega \tau_j)^2} + \frac{(1 - S_f^2) \tau'_f}{1 + (\omega \tau'_f)^2} + \frac{(S_f^2 - S^2) \tau'_s}{1 + (\omega \tau'_s)^2} \right], \quad (2.15)$$

onde $\tau_1^{-1} = 6D_{\perp}$, $\tau_2^{-1} = D_{\parallel} + 5D_{\perp}$, $\tau_3^{-1} = 4D_{\parallel} + 2D_{\perp}$, $A_1 = (3 \cos^2 \beta - 1)^2/4$, $A_2 = 3 \sin^2 \beta \cos^2 \beta$, $A_3 = (3/4) \sin^4 \beta$. β é o ângulo formado entre o vetor NH e o eixo principal do tensor.

O tensor de difusão rotacional da proteína foi calculado segundo procedimentos descritos por Tjandra *et al.*⁷⁴, utilizando o programa R2R1_diffusion⁷⁵. Para esse cálculo são inicialmente excluídos os resíduos que apresentam NOE < 0.65 para então excluir aqueles resíduos sujeitos a trocas conformacionais, identificados por:

$$(\overline{T_2} - T_{2,n})/\overline{T_2} - (\overline{T_1} - T_{1,n})/\overline{T_1} > 1,5 \times \sigma, \quad (2.16)$$

onde $T_2 = 1/R_2$, $T_1 = 1/R_1$, $T_{2,n}$ e $T_{1,n}$ são os valores de T_2 e T_1 para o resíduo n , $\overline{T_2}$ e $\overline{T_1}$ representam os valores médios e σ é o desvio padrão de $(\overline{T_2} - T_{2,n})/\overline{T_2} - (\overline{T_1} - T_{1,n})/\overline{T_1}$. Utilizou-se neste procedimento a estrutura central do estado conformacional mais populado na simulação de dinâmica molecular da construção *C1* da Cdc25B (Capítulo 3), cujas coordenadas foram rotadas com o programa PDBinertia⁷⁶ para diagonalização da matriz com os momentos de inércia.

O programa FAST-Modelfree⁷⁷ foi usado para o cálculo dos parâmetros dinâmicos dos *spins* de ^{15}N e otimização do tensor de difusão rotacional. Este programa interage com o programa Modelfree⁷⁸ e automatiza a análise das medidas de relaxação utilizando critérios descritos por Mandel *et al.*⁷⁹ para a seleção de modelos (Tabela 3) para $J(\omega)$, usando a equação (2.14) no caso de um tensor de difusão isotrópico ou a equação (2.15) para um tensor de difusão axialmente simétrico.

Para um modelo de difusão rotacional axialmente simétrico foram utilizados os seguintes valores iniciais no ajuste dos parâmetros dinâmicos: tempo de correlação global $\tau_c = 12,6 \text{ ns}$, tensor de difusão rotacional $D_{\parallel/\perp} = 1,26$ e ângulos de rotação $\theta = 0^\circ$ e $\phi = 0^\circ$, pois as coordenadas dos átomos da proteínas foram rodadas para o sistema de referências do tensor de inércia. Durante a análise *model-free*, os parâmetros do tensor

Tabela 3 – Modelos para o ajuste dos parâmetros dinâmicos da equação (2.15) do formalismo de Lipari-Szabo.

Modelo	Parâmetros ajustados	Pressupostos
1	S^2	$\tau_f = R_{ex} = 0, S_s^2 = 1$
2	S^2, τ_f	$R_{ex} = 0, S_s^2 = 1$
3	S^2, R_{ex}	$\tau_f = 0, S_s^2 = 1$
4	S^2, τ_f, R_{ex}	$S_s^2 = 1$
5	S_f^2, S_s^2, τ_s	$\tau_f = 0$

de difusão rotacional e τ_c , que afetam globalmente os *spins*, e os parâmetros dinâmicos locais foram ajustados com 600 simulações de Monte Carlo e os valores convergiram após 5 iterações para o modelo de difusão axialmente simétrico e após 2 iterações para o modelo de difusão isotrópica. Somente os *spins* com três medidas de relaxação (R_1 , R_2 e NOE) foram utilizados no ajuste dos parâmetros dinâmicos.

As propriedades difusionais também foram calculadas por modelo de bolas⁸⁰ a partir de estruturas tridimensionais da Cdc25B, utilizando o programa HYDRONMR 7c⁸⁰, a fim de compará-las com as propriedades obtidas por medidas de relaxação. Neste cálculo, o raio dos átomos pesados foi extrapolado a 3.2 Å para considerar a camada de hidratação da proteína e utilizada uma viscosidade do solvente de 0,01 poise.

2.2 Resultados

2.2.1 Clonagem, expressão e purificação

Foram clonadas construções do domínio catalítico da Cdc25B completo (*C1*) e com 32 resíduos do C-terminal truncados (*C2*), da qual retiraram-se a α -hélice e os 16 resíduos cuja densidade eletrônica não foi resolvida no experimento de difração de raios-X (Figura 10). Essas construções foram utilizadas para avaliar os contatos do C-terminal com o núcleo proteico.

O sequenciamento das construções *C1* e *C2* evidenciou que os insertos apresentaram total similaridade com o cDNA da Cdc25B (NCBI: XM_005260900.1; Figura 11).

A ligação destes insertos no plasmídeo de expressão pET-28a foi confirmada por

Cdc25B	*****	120
C1	AGTGACCCAGAGAGCTGATTGGAGATTACTCTAAGGCCTTCCTCTACAGACAGTAGACGGAAGCAACCAAGACCTCAAGTACATCTCACCAGAAACGATGGTGGCCCTATTGACGGGC	120
C2	AGTGACCCAGAGAGCTGATTGGAGATTACTCTAAGGCCTTCCTCTACAGACAGTAGACGGAAGCAACCAAGACCTCAAGTACATCTCACCAGAAACGATGGTGGCCCTATTGACGGGC	120
	1.....10.....20.....30.....40.....50.....60.....70.....80.....90.....100.....110.....120	
Cdc25B	*****	240
C1	AAGTTCAGCAACATCCTGGATAAGTTTGTGATTGACTCCAGATACCCCTATGAAATATGAAGCGGGCACATCAAGACTGCGGTGAATTCGCCCTGGAACCGGACCCGAGAGCTTC	240
C2	AAGTTCAGCAACATCCTGGATAAGTTTGTGATTGACTCCAGATACCCCTATGAAATATGAAGCGGGCACATCAAGACTGCGGTGAATTCGCCCTGGAACCGGACCCGAGAGCTTC	240
	130.....140.....150.....160.....170.....180.....190.....200.....210.....220.....230.....240	
Cdc25B	*****	360
C1	CTACTGAAGAGCCCATCGCGCCCTGTAGCCTGGACAAGAGAGTCATCCTCATTTTCCACTGTGAATTCATCTGAGCGTGGGCCCGCATGTGCCGTTTCATCAGGGAACGAGACCGT	360
C2	CTACTGAAGAGCCCATCGCGCCCTGTAGCCTGGACAAGAGAGTCATCCTCATTTTCCACTGTGAATTCATCTGAGCGTGGGCCCGCATGTGCCGTTTCATCAGGGAACGAGACCGT	360
	250.....260.....270.....280.....290.....300.....310.....320.....330.....340.....350.....360	
Cdc25B	*****	480
C1	GCTGTCAAGCACTACCCAGCCCTCTACTACCTGAGATGTATATCCTGAAAGGGCGGCTACAAGAGTTCCTCCCTCAGCACCCGAACTTCTGTGAACCCAGGACTACCGGCCATGAAC	480
C2	GCTGTCAAGCACTACCCAGCCCTCTACTACCTGAGATGTATATCCTGAAAGGGCGGCTACAAGAGTTCCTCCCTCAGCACCCGAACTTCTGTGAACCCAGGACTACCGGCCATGAAC	480
	370.....380.....390.....400.....410.....420.....430.....440.....450.....460.....470.....480	
Cdc25B	*****	582
C1	CACGAGGCCCTTCAAGGATGAGCTAAAGACCTTCGGCCTCAAGACTCGCAGCTGGGCTGGGGAGCGGAGCCGGCGGGAGCTCTGTAGCCGGCTGCAGGACCAG	582
C2	CACGAGGCCCTTCAAGGATGAGCTAAAGACCTTCGGCCTCAAGACTCGCAGCTGGGCTGGGGAGCGGAGCCGGCGGGAGCTCTGTAGCCGGCTGCAGGACCAG	486
	490.....500.....510.....520.....530.....540.....550.....560.....570.....580..	

Figura 11 – Alinhamento entre o sequenciamento das construções *C1*, *C2* e o cDNA da Cdc25B.

eletroforese em gel de agarose após dupla digestão com as enzimas de restrição *NdeI* e *BamHI*. Este resultado está apresentado na figura 12, onde pode-se observar bandas referentes ao plasmídeo linearizado (5 kbp) e aos insertos das construções *C1* (~ 600 bp) e *C2* (~ 500 bp) liberados após a digestão.

Figura 12 – Dupla digestão dos plasmídeos pET-28a_ *C1* e pET-28a_ *C2* com as enzimas de restrição *NdeI* e *BamHI*. PB: padrão de pares de base.

A expressão das construções *C1* e *C2* da Cdc25B contendo a cauda de His₆ no N-terminal (*C1*: 25,6 kDa e *C2*: 21,7 kDa) foi confirmada pelo aparecimento de bandas com a massa molecular esperada das proteínas após 2 horas de indução do meio de cultivo bacteriano (Figura 13). Amostras coletadas após 4 e 17 horas de indução mostraram o aumento na quantidade de proteína expressa.

Apesar da construção *C2* ter sido expressa na forma insolúvel (P; Figura 14),

Figura 13 – Perfis de expressão de proteínas das bactérias transformadas com os vetores pET-28a_ *C1* e pET-28a_ *C2*. Meio de cultivo não induzido (2) e após 0,5 h (3), 2 h (4), 4 h (5) e 17 h (6) de indução com IPTG.

também foi possível obter uma quantidade considerável na forma solúvel (S; Figura 14). O rendimento da expressão foi de ~ 10 mg de proteína solúvel em 1 l de meio de cultivo, quantificada por espectrofotometria (Abs_{280nm}) após a etapa de purificação em coluna Ni-NTA agarose. Resultados semelhantes foram obtidos para a construção *C1*.

Figura 14 – Perfil proteico das etapas do primeiro passo de purificação da construção *C2* em coluna de Ni^{2+} . MM: padrão de massa molecular; P: *pellet* de bactérias lisadas; S: fração solúvel obtida após lise bacteriana; FT: fração que não se ligou à resina; W: lavagem com imidazol 20 mM; E: fração eluída com imidazol 500 mM.

Após a primeira etapa de purificação em coluna Ni-NTA agarose, a construção *C1* ainda não se mostrou homogênea devido à presença de contaminantes com ~ 100 kDa (banda E; Figura 15A). Antes da segunda etapa de purificação, a proteína foi submetida a clivagem enzimática da cauda de His_6 com o uso da resina trombina-agarose. Foi possível confirmar a clivagem pela diferença no padrão de migração das bandas referentes à proteína antes e depois da incubação com a enzima (Figura 15A). Resultados semelhantes foram obtidos para a construção *C2*.

Devido à presença de contaminantes com massa molecular muito diferente da proteína de interesse, amostras das construções *C1* e *C2* obtidas após purificação em coluna Ni-NTA agarose foram submetidas a uma nova etapa de purificação por gel filtração.

O cromatograma dessa etapa de purificação para a construção *C1* revelou a presença de um pico majoritário eluído com ~ 180 ml e outros eluídos a menores volumes (Figura 15B). No cromatograma também estão destacadas as frações utilizadas para a composição da amostra final usada nos experimentos de RMN (Figura 15B inserto). Resultados semelhantes foram obtidos para a construção *C2*.

Figura 15 – Composição das amostras obtidas após as etapas de purificação da construção *C1*. (A) Gel de SDS-PAGE com o perfil proteico da amostra purificada em coluna de Ni^{2+} (E) e submetida a clivagem com trombina da cauda de His_6 (T). MW: padrão de massa molecular; FT: fração que não se ligou à resina; W: lavagem com imidazol 20 mM. (B) Cromatograma da segunda etapa de purificação por gel filtração. No inserto está apresentado um gel de SDS-PAGE com as frações utilizadas para a composição da amostra usada nos experimentos de RMN (traço vermelho).

2.2.2 Assinalamento da cadeia principal

Optou-se inicialmente pela aquisição dos espectros RMN para o assinalamento da construção *C2* devido a sua maior estabilidade em comparação à construção *C1* da Cdc25B. O assinalamento das ressonâncias dos núcleos da construção *C2* auxiliou o assinalamento da construção *C1*, já que houve sobreposição de grande parte dos sinais nos espectros de $^1\text{H}-^{15}\text{N}$ HSQC (Figura 28). A seguir, será apresentado em mais detalhes o assinalamento da construção *C1*.

O processo de assinalamento inicia-se com o ordenamento sequencial dos sistemas de *spins*, utilizando os deslocamentos químicos dos sinais nos espectros tridimensionais, para posteriormente atribuí-los à sequência primária da proteína. Este procedimento é possível pois alguns resíduos possuem deslocamentos químicos de $^{13}\text{C}^\alpha$ e $^{13}\text{C}^\beta$ característicos. Na

figura 16 estão ilustradas regiões do espectro tridimensional HNCA da construção *C1* que formam sistemas de *spins* sequenciais que foram atribuídos ao segmento 536F–546K da proteína.

Figura 16 – Regiões do espectro tridimensional HNCA utilizadas no assinalamento da cadeia principal da α -hélice C-terminal (536F–546K) da construção *C1*.

O espectro de ^1H – ^{15}N HSQC da construção *C1* (Figura 17) apresentou boa dispersão e resolução dos sinais, exceto na região de $\delta\ ^1\text{H} \sim 8$ ppm onde ocorre sobreposição. O número de picos identificados referentes ao grupo NH da cadeia principal foi próximo ao esperado. Também foi possível observar na região de $\delta\ ^{15}\text{N} \sim 112$ ppm grande parte dos sinais esperados para os grupos NH_2 das cadeias laterais de Gln e Asn (Tabela 4) e na região de $\delta\ ^1\text{H} \sim 10$ ppm um pico referente ao grupo NH da cadeia lateral do resíduo 550W. Picos referentes a grupos $\text{NH}\epsilon$ da cadeia lateral de Arg foram identificados pelo deslocamento dos seus sinais após aquisição de outro espectro com uma janela espectral diferente, já que esses sinais normalmente aparecem rebatidos.

Tabela 4 – Sinais esperados, observados e assinalados no espectro de ^1H – ^{15}N HSQC da construção *C1* (Figura 17).

Grupos NH	Esperados	Observados	Assinalados
Cadeia principal	187	173	143
Cadeia lateral (Gln+Asn)	24	20	–

Ao final do processo, foram assinalados 73% dos *spins* da cadeia principal (^1H , ^{15}N , $^{13}\text{C}^\alpha$, $^{13}\text{C}^\beta$, $^{13}\text{C}'$) da construção *C1*. Dos 201 resíduos, 161 tiveram ao menos um átomo

Figura 17 – Espectro de $^1\text{H}-^{15}\text{N}$ HSQC da construção *C1* adquirido a 298 K de amostra contendo $\sim 200\ \mu\text{M}$ da proteína em tampão fosfato [NaPi 20 mM; NaCl 50 mM; DTT 5 mM; BMER 5 mM; EDTA 1 mM; 5% $^2\text{H}_2\text{O}$; pH 6,75]. Os resíduos assinalados estão coloridos em azul na estrutura cristalográfica inserida na figura.

assinalado, incluindo os 16 resíduos C-terminais ausentes na estrutura cristalográfica. Também foram assinalados os resíduos 472H e a cisteína catalítica 473C do *P-loop* (472H–479R). Nota-se na estrutura cristalográfica apresentada na figura 17 que a região não assinalada da proteína é rica em *loops*.

O espectro de $^1\text{H}-^{15}\text{N}$ HSQC assinalado da construção *C1* está apresentado na figura 18. Apesar da grande sobreposição na região de $\delta\ ^1\text{H} \sim 8\ \text{ppm}$ foi possível assinalar grande parte dos picos (Figura 18B). Nota-se também que diversos resíduos do segmento C-terminal ausentes na estrutura cristalográfica estão presentes nessa região do espectro (pontos vermelhos; Figura 18).

Como o intuito de confirmar o assinalamento da cisteína catalítica, presente no *P-loop*, foi realizada uma titulação com NaPi (concentrações de 0–16 mM) de uma amostra da construção *C1* em tampão MES marcada isotopicamente com ^{15}N . Espectros de $^1\text{H}-^{15}\text{N}$ HSQC foram adquiridos ao longo da titulação (Figura 19) e as frequências dos sinais com 0 e 16 mM de NaPi foram utilizadas no cálculo do CSP (Equação (2.1); Figura 20A)

Figura 18 – Assinalamento do espectro de ^1H – ^{15}N HSQC construção *C1* adquirido a 298 K de amostra contendo $\sim 200\ \mu\text{M}$ de proteína em tampão fosfato de sódio [NaPi 20 mM; NaCl 50 mM; DTT 5 mM; BMER 5 mM; EDTA 1 mM; 5% $^2\text{H}_2\text{O}$; pH 6,75]. **(A)** Região do espectro referente aos grupos amida da cadeia principal. **(B)** Detalhe da região de $\delta\ ^1\text{H} \sim 8\ \text{ppm}$. Os resíduos do C-terminal ausentes na estrutura cristalográfica estão com uma marca vermelha.

O deslocamento dos sinais no espectro durante a titulação da proteína com um ligante indica a mudança no ambiente químico do resíduo correspondente e sua possível participação na formação do complexo proteína-ligante. Assim, os resíduos da Cdc25B que apresentaram elevados valores de CSP após a titulação com NaPi (Figura 20A) puderam ser mapeados na estrutura cristalográfica (Figura 20B). Nota-se que os resíduos com CSP mais elevados fazem parte dos sítios cristalográficos *P1* e *P2* de ligação a fosfato. O *P-loop*

Figura 19 – Espectros sobrepostos de ^1H – ^{15}N HSQC adquiridos durante a titulação da construção *C1* com NaPi. Os espectros foram coletados em tampão MES com 0, 1, 2, 4, 8 ou 16 mM de NaPi.

(472H–479R), que possui a cisteína catalítica 473, está presente no sítio P1.

O estado de oxidação das cisteínas pode ser avaliado pelo CS de $^{13}\text{C}^\beta$. Cisteínas envolvidas em pontes dissulfeto apresentam CS $^{13}\text{C}^\beta \sim 42$ ppm enquanto cisteínas reduzidas este valor é ~ 28 ppm^{81, 82}. Todas as cisteínas assinaladas da construção *C1* possuem valores de CS de $^{13}\text{C}^\beta$ condizentes com o estado reduzido. Para exemplificar, na figura 21 estão apresentados CS das cisteínas 426, 473 e 484.

O programa TALOS-N⁵⁹ foi utilizado na previsão dos ângulos diedrais φ e ψ e de estruturas secundárias na Cdc25B. No geral, os ângulos previstos estão de acordo com os ângulos extraídos da estrutura cristalográfica (Figura 22), apresentando coeficientes de correlação de Pearson (ρ) de 0,86 e 0,88 para φ e ψ , respectivamente. Essa correlação aponta para o correto assinalamento da cadeia principal e sugere a semelhança estrutural entre o domínio catalítico da Cdc25B em solução e cristalográfico. Em acordo com

Figura 20 – Perturbações do deslocamento químico da construção *C1* obtidas por titulação com NaPi. (A) CSP calculado com a equação (2.1) com valores de CS em 0 e 16 mM de NaPi. (B) Resíduos que apresentaram valores de CSP significativos. Os íons SO_4^{2-} cristalográficos estão representados como bastão.

Figura 21 – Regiões do espectro tridimensional HN(CA)CB onde são apresentados os picos de correlação $^1\text{H} \times ^{13}\text{C}^\beta$ das cisteínas 426, 473 e 484. Estes resíduos estão representados como bastão na estrutura à esquerda.

estes resultados, os índices do deslocamento químico (CSI)⁸³ também apontam para a similaridade entre as estruturas cristalográficas e em solução (Figura 23).

As previsões efetuadas pelo programa TALOS-N podem ser classificadas como:

- ambíguas: previsões agrupadas em duas regiões distintas do gráfico de Ramachandran;
- únicas: previsões agrupadas em uma única região do gráfico de Ramachandran;
- flexíveis: resíduos com altos valores de RCI ⁶⁰.

Figura 22 – Comparação entre ângulos diedrais φ e ψ extraídos da estrutura cristalográfica e previstos pelo programa TALOS-N para a construção *C1*.

Figura 23 – Comparação entre a estruturas secundárias cristalográfica e previstas por TALOS-N⁵⁹ ou CSI⁸³. Os segmentos assinalados da proteína estão sublinhados.

Aproximadamente 10% dos resíduos assinalados apresentaram previsão ambígua dos diedrais. Contudo, exceto para o resíduo 435G, em todas previsões ambíguas há uma conformação φ/ψ em acordo com a estrutura cristalográfica. Os ângulos diedrais dos resíduos 437, 438, 439, 473 e 547 apresentaram previsões únicas mas distintas dos valores extraídos da estrutura cristalográfica (Figura 22).

Para facilitar as análises posteriores, três segmentos da porção C-terminal da Cdc25B foram nomeados de acordo com suas características estruturais (Figura 23):

- 366G–386F: NTD (N-terminal desordenado);
- 387L–546K: núcleo proteico;
- 525P–533H: LCT (*linker* C-terminal);
- 534E–546K: HCT (hélice C-terminal);
- 547T–566Q: CTD (C-terminal desordenado).

2.2.3 Flexibilidade conformacional

A flexibilidade conformacional da cadeia principal da Cdc25B foi investigada por algoritmos de previsão que utilizam os CS dos *spins* (^1H , ^{15}N , $^{13}\text{C}'$, $^{13}\text{C}^\alpha$ e $^{13}\text{C}^\beta$), pelo ajuste dos parâmetros de ordem S^2 a partir de medidas de relaxação (R_1 , R_2 e NOE $\{^1\text{H}\}-^{15}\text{N}$) e pelos RDCs dos grupos amida da cadeia principal da construção *C1*.

Inicialmente, pode-se verificar que a maioria dos sinais do grupo amida dos resíduos C-terminais (549–566) localizam-se na região de $\delta\ ^1\text{H} \sim 8$ ppm e $\delta\ ^{15}\text{N} \sim 120$ ppm no espectro de HSQC (Figura 18B) que é característica da maioria dos resíduos em segmentos desordenados⁸⁴ (Tabela 5). As exceções são resíduos de Ser, Thr e Gly em regiões desordenadas, pois apresentam CS de NH fora da região de $^1\text{H} \sim 8$ ppm $^{15}\text{N} \sim 120$ ppm. Contudo, esses resíduos do C-terminal da Cdc25B (549S, 551A, 552G, 555S e 561S) também apresentam CS de NH compatíveis com segmentos desordenados⁸⁵.

Em acordo com essa observação, as populações em estrutura secundária previstas pelo método $\delta 2\text{D}$ ⁸⁶ e o índice de desordem conformacional (RCI)⁶⁰, que utilizam os CS no cálculo, também apontam para a ausência de estruturas secundárias e a flexibilidade conformacional da região C-terminal da Cdc25B, respectivamente (Figura 24). Além disso, esses resultados apontam para a flexibilidade da região N-terminal e uma relativa rigidez conformacional no núcleo da proteína.

Além de prevista através dos CS, a flexibilidade conformacional da cadeia polipeptídica pode ser avaliada pelo ajuste do parâmetro de ordem obtido através do formalismo

Tabela 5 – Deslocamentos químicos de ^1H e ^{15}N da cadeia principal do segmento C-terminal 549–566 observados e calculados para regiões desordenadas.

Resíduo	$\delta \text{ } ^1\text{H}$ (ppm)		$\delta \text{ } ^{15}\text{N}$ (ppm)	
	Obs.	Calc. ^a	Obs.	Calc. ^a
549S	8.28	8.50	116.07	117.57
550W	8.18	7.90	123.31	123.03
551A	8.17	7.60	125.37	124.70
552G	7.68	8.38	107.03	108.04
553E	8.22	8.27	121.01	120.70
554R	8.39	8.40	121.47	122.46
555S	8.23	8.43	116.05	117.78
556R	8.26	8.42	122.62	123.52
557R	8.25	8.47	121.32	123.10
558E	8.38	8.47	121.11	122.28
559L	8.20	8.64	122.33	123.89
560C	8.25	8.51	119.17	119.85
561S	8.32	8.47	118.22	118.97
562R	8.30	8.38	122.70	123.19
563L	8.24	8.32	122.86	123.70
564Q	8.30	8.36	119.92	120.69
565D	8.31	8.43	121.45	121.82

^a valores estimados de acordo com Tamiola *et al.*⁸⁵

de Lipari-Szabo, utilizando medidas de relaxação da magnetização longitudinal (R_1) e transversal (R_2) e de NOE heteronucleares (Seção 2.1.3.6). Para tal finalidade, inicialmente é necessário obter-se uma estimativa do tensor de difusão rotacional e do tempo de correlação global (τ_c) da proteína.

O tensor de difusão rotacional e o τ_c podem ser estimados através da relação R_2/R_1 dos *spins* nucleares, que teoricamente independe de movimentações internas rápidas⁸⁷, e uma estrutura tridimensional da proteína para obter-se a orientação dos vetores internucleares. Assim, a relação R_2/R_1 dos *spins* de ^{15}N que não estão sujeitos a troca conformacional e/ou movimentações mais lentas que algumas centenas de ps fornecem uma estimativa do tensor de difusão rotacional. Valores de R_2/R_1 foram obtidos para um total de 85 resíduos da Cdc25B. Dentre eles foram excluídos da análise, 16 resíduos (371, 372, 375, 377, 418, 419, 547, 548, 551, 553, 554, 558, 559, 563, 564, 565) que apresentaram $\text{NOE} < 0.65$, 2 resíduos (413 e 534) sujeitos a troca conformacional (Equação (2.16)) e o resíduo 452 cujo valor χ^2 mostrou-se muito elevado numa estimativa inicial. Assim, os

Figura 24 – Populações de estruturas secundárias e flexibilidade conformacional da construção C1 estimadas pelo método $\delta 2D^{86}$. (A) e por RCI⁶⁰ (B), respectivamente.

valores de relaxação de 66 resíduos foram utilizados na estimativa do tensor de difusão rotacional e do τ_c (Tabela 6).

As medidas de relaxação da Cdc25B foram melhor ajustadas a um tensor de difusão rotacional axialmente simétrico, constatado pelo menor valor de χ^2 (Tabela 6). Contudo, devido ao maior número de parâmetros ajustáveis no tensor axialmente simétrico (4) em relação ao isotrópico (1), torna-se fundamental a realização de um teste F ^{63, §19.8} para avaliar se essa melhora é estatisticamente significativa. O teste F para o modelo de difusão axialmente simétrico (Tabela 6) apresentou valor superior ao mínimo calculado para 66 medidas com 3 parâmetros adicionais e confiabilidade de 95% ($F_{(3;62;0,05)} = 2,76$), ou seja, a probabilidade de que a melhora no ajuste dos dados tenha ocorrido por acaso é menor do que 5%. Nota-se também que um tensor de difusão assimétrico está em melhor acordo com as propriedades difusionais calculadas para modelos hidrodinâmicos da estrutura cristalográfica da Cdc25B (Tabela 6).

No geral, as medidas de relaxação dos resíduos do núcleo da Cdc25B (387–546) apontam para uma relativa rigidez conformacional em solução comparada aos segmentos terminais (Figura 25). Todos os resíduos, exceto 377, 418, 419 e 456, apresentaram NOE heteronucleares > 0.65 . Foram observadas taxas de relaxação transversal R_2 mais elevadas

Tabela 6 – Parâmetros do tensor de difusão rotacional experimentais (RMN) e calculados (hidrodinâmicos).

	Modelo	τ_c (ns)	$\frac{2D_z}{D_x+D_y}$	D_x/D_y	θ	ϕ	ψ	χ_{red}^2	$F_{(3;62)}^a$
RMN ^b	D_{iso}	12,17	1	1				29,5	
	D_{axi}	12,09	1,26	1	0°	0°		25,7	4,19
hidrodinâmico ^c	I ^d	13,48	1,33	1,04	6°	5°	7°		
	II ^e	12,63	1,34	1,04	5°	7°	8°		
	III ^f	16,34	1,41	1,03	2°	5°	5°		

^a teste F do modelo axial em relação ao modelo isotrópico do tensor de difusão rotacional ($F_{(3;62;0,05)} = 2,76$)

^b obtidos a partir de medidas de relaxação com o uso do programa R2R1_diffusion⁷⁵

^c calculado por modelo de bolas com o uso do programa HYDRONMR 7c⁸⁰

^d baseado nas coordenadas dos resíduos 374–550 da estrutura cristalográfica (PDB: 1QB0)

^e baseado nas coordenadas cristalográficas contendo os resíduos que valores de NOE > 0,65 (res. 378–546)

^f baseado nas coordenadas da estrutura central do *cluster* mais populado da simulação de dinâmica molecular da construção C1 da Cdc25B (res. 366–566)

em relação aos terminais, exceto para os resíduos 392, 393, 452 e 456, onde $R_2 < 15 \text{ s}^{-1}$. De maneira similar, as taxas de relaxação R_1 também apresentaram-se diferentes nos terminais, onde mostraram-se mais elevadas. Assim, esses resultados apontam para a flexibilidade conformacional nas regiões terminais da Cdc25B e uma relativa rigidez no núcleo, exceto pelos resíduos 392, 452 e 456 que apresentam baixos valores de R_2/R_1 . Também é possível notar que os resíduos 413 e 534 apresentaram elevados valores de R_2/R_1 , o que indica a presença de troca conformacional na escala de μs – ms .

Os parâmetros dinâmicos do formalismo de Lipari-Szabo da Cdc25B foram calculados a partir da medidas de relaxação, utilizando-se um tensor de difusão rotacional axialmente simétrico. Foi utilizado esse modelo de difusão devido ao melhor ajuste das medidas de relaxação e ao acordo com modelos hidrodinâmicos (Tabela 6). Os parâmetros de ordem S^2 obtidos no formalismo de Lipari-Szabo são inversamente proporcionais à amplitude de movimentações na escala de ps–ns, refletindo assim a flexibilidade da cadeia polipeptídica numa escala de tempo restrita. A figura 26 mostra que as regiões terminais da Cdc25B apresentam menores valores de S^2 , evidenciando assim a maior flexibilidade conformacional dos segmentos terminais em relação ao núcleo proteico. De maneira semelhante, nota-se a maior flexibilidade dos segmentos em *loop* quando comparados às regiões da proteína em estrutura secundária.

Figura 25 – Medidas de relaxação dos *spins* de ^{15}N da cadeia principal da construção C1. São apresentadas as taxas de relaxação longitudinal (R_1) e transversal (R_2), os valores de NOE heteronucleares $\{^1\text{H}\}-^{15}\text{N}$ e a razão R_2/R_1 .

Em comparação com os segmentos do núcleo da proteína (res. 387–546) em estrutura secundária (α -hélice, fita- β ou volta- β), que apresentam $\overline{S^2} = 0,96 \pm 0,04$, as regiões em *loop* possuem $\overline{S^2} = 0,89 \pm 0,09$. Já os resíduos do C-terminal (549–566) e N-terminal (366–376) apresentam $\overline{S^2} = 0,32 \pm 0,1$, mostrando-se assim muito inferiores aos resíduos não terminais cujo $\overline{S^2} = 0,93 \pm 0,07$. Assim, esses resultados apontam para a desordem conformacional das regiões terminais da Cdc25B e para a maior flexibilidade de regiões em *loop* comparadas àquelas em estrutura secundária.

Durante o ajuste dos S^2 pelo programa FAST-Modelfree (Seção 2.1.3.6), os *spins* de ^{15}N dos resíduos 377 e 452 não puderam ser atribuídos com confiabilidade estatística a nenhum dos modelos dinâmicos descritos na tabela 3 (Seção 2.1.3.6). O modelo 1 foi o mais utilizado no ajuste dos parâmetros de ordem S^2 , totalizando 38 resíduos, dentre os quais nenhum está presente nas regiões terminais da proteína. Os modelos 2 ou 3, onde são ajustados dois parâmetros dinâmicos, foram adotados para 19 resíduos. E os modelos 4 ou 5, onde são ajustados três parâmetros dinâmicos, foram utilizados em 23 resíduos. Em 20 resíduos foram utilizados os modelos 3 ou 4, onde ajusta-se o termo da

Figura 26 – Parâmetros do formalismo de Lipari-Szabo ajustados a partir das medidas de relaxação (R_1 , R_2 e NOE) da construção *C1*. São apresentados os parâmetros de ordem S^2 (A) e as contribuições de R_{ex} (B).

equação (2.11) que descreve contribuições de troca conformacional (R_{ex}) para a relaxação da magnetização transversal.

Os resíduos que possivelmente apresentaram contribuições de troca conformacional encontram-se distribuídos pela estrutura tridimensional da Cdc25B, localizando-se em regiões de *loop* e estrutura secundária e, principalmente, nos segmentos LCT e HCT (página 64; inserto na figura 26B). Os resíduos 413 e 534 apresentaram as contribuições mais significativas de R_{ex} (Figura 26B), evidenciadas também por suas elevadas razões R_2/R_1 (Figura 25), após o ajuste dos parâmetros dinâmicos pelo formalismo de Lipari-Szabo.

Outra evidência experimental que aponta para a flexibilidade conformacional da região C-terminal (res. 549–566) da Cdc25B são os valores de RDC com $D_{NH}^{obs} \sim 0$ (Figura 27A). Esta observação pode ter duas interpretações opostas: (i) os vetores NH dessa região poderiam ser pouco flexíveis e possuir $\zeta \sim 54,7^\circ$ (Equação (2.2)) ou (ii) poderiam assumir um tombamento mais isotrópico em relação ao restante da proteína, quando os vetores NH não assumem uma orientação preferencial em relação ao campo magnético (Equação (2.3)). Como os valores de S^2 evidenciaram que a região C-terminal

é desordenada, essa última característica atenuaria seu alinhamento em relação ao núcleo proteico estruturado. Assim, os vetores NH dessa região não assumiriam uma orientação preferencial em meio anisotrópico, acarretando em $D_{\text{NH}}^{\text{obs}} \sim 0$.

Além disso, o acordo global entre os RDCs observados e calculados (Figura 27B), evidenciado pelo fator Q^{88} de 0,26, indica a semelhança entre as estruturas cristalográfica e em solução da Cdc25B, conforme sugerido pela previsão dos ângulos diedrais φ/ψ (Figura 22). Contudo, os resíduos que apresentam as diferenças mais significativas entre os RDCs observados e calculados (Figura 27B) apontam para inconsistências locais entre a estrutura cristalográfica e em solução e/ou a presença de flexibilidade conformacional.

Figura 27 – Acoplamentos dipolar residual dos grupos amida da cadeia principal (D_{NH}) da construção C1. Os valores observados ($D_{\text{NH}}^{\text{obs}}$) são apresentados em função dos resíduos (A) e dos valores calculados ($D_{\text{NH}}^{\text{calc}}$) a partir do tensor de alinhamento obtido da estrutura cristalográfica (B). Fator $Q = \sqrt{\sum_{i=1}^N 5(D_{\text{NH}}^{\text{calc}} - D_{\text{NH}}^{\text{obs}})^2 / (2ND_a^2 + 4R^2)^{88}}$. A magnitude da diferença entre $D_{\text{NH}}^{\text{obs}}$ e $D_{\text{NH}}^{\text{calc}}$ também é apresentada (C).

2.2.4 Contatos entre a região C-terminal e o núcleo proteico

Na seção 2.2.3 demonstrou-se através de previsões computacionais e medidas de RMN que a região C-terminal da Cdc25B é flexível em solução. Com o intuito de avaliar possíveis contatos entre a região C-terminal flexível e o núcleo proteico, foram medidas

perturbações do deslocamento químico (Equação (2.1)) dos grupos NH da cadeia principal após remoção do segmento C-terminal da construção *C1*. Para tal finalidade foram utilizados os CS de ^1H e ^{15}N da construção *C2*, da qual foram truncados 32 resíduos do C-terminal (535–566) referentes aos segmentos HCT e CTD (Página 64).

A sobreposição de grande parte dos sinais nos espectros de ^1H – ^{15}N HSQC das construções *C1* e *C2* (Figura 28) evidencia que a remoção da região C-terminal não afetou de maneira significativa o enovelamento do núcleo proteico. Assim, o deslocamento de alguns sinais deve estar relacionado a mudanças no ambiente químico provocadas principalmente pela remoção de resíduos do C-terminal que fazem contatos transientes com o núcleo da Cdc25B.

Figura 28 – Sobreposição dos espectros de ^1H – ^{15}N HSQC das construções *C1* (laranja) e *C2* (azul). Setas indicam os picos que sofreram perturbações mais evidentes.

Nota-se nos espectros da figura 28 que os picos que sofreram deslocamentos mais evidentes são do segmento 428–432 e dos resíduos 533H e 534E do LCT. Além do LCT, quatro segmentos da proteína coloridos em vermelho (426–436), laranja (442–449), amarelo

(482–485) e azul (510–514) na figura 29 e resíduos do P-loop sofreram perturbações significativas do deslocamento químico (Figura 29). Assim, esses resultados sugerem que resíduos do núcleo proteico podem realizar contatos com o C-terminal da construção *C1* da Cdc25B. Contudo, não é possível avaliar quais são os resíduos envolvidos em contatos núcleo–C-terminal.

Figura 29 – Contatos entre a região C-terminal e o núcleo da construção *C1*. (A) CSP calculados entre os CS das construções *C1* e *C2* pela equação (2.1). Os segmentos com CSP > 0.05 ppm estão identificados com as mesmas cores na estrutura cristalográfica (B). A linha tracejada representa o valor médio de CSP do segmento 369–524.

2.3 Discussão

A presença de bandas com peso molecular distinto das construções *C1* e *C2* da Cdc25B na fração eluída com imidazol 20 mM durante a primeira etapa de purificação em coluna Ni-NTA agarose (W; Figura 15A) demonstra a eficácia desse passo de lavagem para a eluição de proteínas que se ligaram inespecificamente à resina. Contudo, a presença de proteínas contaminantes na amostra obtida após a eluição com imidazol 500 mM (E; Figura 15A) demonstrou a necessidade de se utilizar uma segunda etapa para a purificação.

Com o intuito de avaliar a clivagem da cauda de His₆, foram aplicadas numa mesma raia do gel de SDS-PAGE amostras da proteína submetida (T) ou não (E) à reação com trombina (E+T; Figura 15A). A separação entre as bandas conferida após a remoção da

cauda de His₆ (~ 2 kDa) ficou evidente no gel, demonstrando assim a especificidade da reação enzimática.

Como a massa molecular das construções *C1* ou *C2* (~ 20 kDa) é significativamente menor que a dos contaminantes presentes após a purificação em coluna de Ni²⁺ (E; Figura 15A), optou-se por implementar uma etapa de gel filtração para a obtenção de uma amostra mais homogênea para a aquisição dos espectros de RMN. O cromatograma dessa etapa de purificação (Figura 15B) revelou um pico majoritário e bem definido eluído em ~ 180 ml, referente à proteína de interesse e cujas frações foram utilizadas na composição da amostra para RMN. Outros dois picos secundários foram eluídos em menores volumes e, portanto, referem-se aos contaminantes com peso molecular de ~ 100 kDa. Além da remoção das proteínas contaminantes, esta etapa de gel filtração também foi eficiente na separação da cauda de His₆, já que sua massa molecular (~ 2 kDa) está abaixo do limite inferior de separação da coluna (3 kDa) especificado pelo fabricante.

Apesar da quantidade considerável de proteína expressa na forma insolúvel (P; Figura 14), após as duas etapas de purificação obteve-se um rendimento, quantificado por Abs_{280nm}, de ~ 10 mg/l. Esta quantidade foi suficiente para a obtenção de 4 amostras de 500 μ l a 200 μ M de proteína para a aquisição dos espectros de RMN.

A expressão de proteínas em meio de cultivo preparado com ²H₂O causa a substituição aleatória de ¹H por ²H na cadeia polipeptídica. Essa estratégia contribui para diminuir a contribuição de interações dipolares para a taxa de relaxação da magnetização transversal (Equação (2.11)), pois a constante giromagnética (γ) de ²H é aproximadamente seis vezes menor que a constante giromagnética de ¹H. Assim, a diminuição de R_2 prolonga o decaimento livre da indução (FID), contribuindo para o aumento da resolução e sensibilidade dos sinais no espectrômetro. Contudo, obteve-se um menor rendimento da construção *C1* expressa em meio de cultivo preparado com ²H₂O comparada à expressão em meio preparado com H₂O. Essa diminuição no rendimento já foi relatado para outras proteínas⁸⁹ e deve-se, possivelmente, à diminuição do metabolismo bacteriano em meio de cultivo contendo ²H₂O.

O processo de assinalamento foi inicialmente conduzido com a construção *C2* devido à sua maior estabilidade em solução, constatada visualmente pela menor precipitação

após os longos períodos necessários para a aquisição dos espectros 3D de RMN. Esse procedimento facilitou o assinalamento da construção *C1* pois foi possível transferir o assinalamento entre os espectros de $^1\text{H}-^{15}\text{N}$ HSQC devido à sobreposição de grande parte dos sinais (Figura 28).

No espectro de $^1\text{H}-^{15}\text{N}$ HSQC da construção *C1* (Figura 17) observou-se um número de sinais de grupos amida da cadeia principal inferior ao esperado (Tabela 4). Dentre os sinais observados, 20 não puderam ser assinalados devido à ausência de correlações nos espectros tridimensionais ou à presença de sistemas de *spins* que não puderam ser atribuídos a nenhum segmento da sequência primária. Dessa maneira, não foi possível assinalar completamente a cadeia principal da Cdc25B.

O assinalamento da construção *C1* está de acordo com o assinalamento publicado em 2014 por Lund e Cierpicki⁴⁹. Contudo, a construção utilizada por esses autores (res. 372–551) possui uma mutação na cisteína catalítica (C473S) que aumenta a estabilidade da proteína e não contem os últimos 15 resíduos do CTD que foram assinalados na construção *C1* (pontos vermelhos; Figura 18). Nesse trabalho foi assinalado um maior número (125) de grupos amida da cadeia principal quando comparado ao número de grupos assinalados na construção *C1* (119). Este fato, conforme discutido pelo autores, pode ser explicado pela maior estabilidade do mutante C473S em relação à Cdc25B nativa, o que certamente confere uma melhor qualidade aos espectros. Os resíduos 434G, 509K, 511G e 529R apresentaram diferenças significativas nos CS de ^{15}N , pois a janela espectral utilizada pelos autores durante a aquisição do espectro de HSQC foi menor do que a utilizada aqui. Assim, os sinais desses resíduos apareceram rebatidos no espectro e, portanto, apresentaram valores incorretos na dimensão de ^{15}N ⁴⁹.

Em particular, o assinalamento da cisteína catalítica 473 presente no *P-loop* foi confirmado após a titulação com NaPi (Figura 20) de onde foram identificados os sítios cristalográficos P1 e P2 de ligação a Pi. Os demais resíduos não observados do *P-loop* possivelmente apresentam sinais alargados. Este alargamento de sinais do sítio ativo da Cdc25B (*P-loop*; Figura 17) pode ser causado por um equilíbrio ($\text{A} \xrightleftharpoons[k_2]{k_1} \text{B}$) em regime de troca intermediária^{90, Cap. IV}, quando a diferença entre a frequência de ressonância dos estados ($\Delta\nu = \nu_A - \nu_B$) é semelhante à taxa de conversão (k). Este fato já foi relatado

no assinalamento do P-loop de uma PTP (proteína tirosina fosfatase) de baixo peso molecular⁹¹, onde notou-se o desaparecimento de alguns sinais no espectro de HSQC após titulação com NaPi. Sendo assim, os autores concluíram que o equilíbrio de associação de Pi ao P-loop está em regime de troca intermediária. Contudo, no caso da construção *C1* da Cdc25B não houve o surgimento e nem o desaparecimento de sinais durante a titulação com Pi (Figura 19). Portanto, é provável que a ausência de sinais do P-loop está relacionada a um equilíbrio conformacional independente da ligação a Pi. Vale ainda ressaltar que a região não assinalada da Cdc25B é rica em *loops* (Insero na figura 17) que também podem estar sujeitos a trocas conformacionais, provocando o desaparecimento de alguns sinais esperados, conforme contabilizado pelos picos no espectro de $^1\text{H}-^{15}\text{N}$ HSQC da construção *C1* (Tabela 4).

Lund e Cierpicki⁴⁹ também reportaram a ausência de flexibilidade conformacional em diferentes escalas de tempo no segmento HCT, utilizando medidas de relaxação de uma construção do domínio catalítico da Cdc25B da qual foram truncados 15 resíduos do CTD (*C1.d15*). Em contraste com esse resultado, as análises de *model-free* da construção *C1* revelaram importantes contribuições de R_{ex} nos resíduos 538 e 545 do HCT (Figura 26B). Assim, é provável que os 15 resíduos no C-terminal da construção *C1* confirmam ao HCT uma flexibilidade conformacional distinta da observada por esses autores na construção *C1.d15*.

Vale ainda ressaltar que o elevado valor de R_2 do resíduo 534E da construção *C1* (Figura 25), que sugere a presença de movimentações lentas no LCT, também foi observado por Lund e Cierpicki⁴⁹ na construção *C1.d15* selvagem. Contudo, em contraste com a selvagem, os resíduos do LCT no mutante C473S apresentaram valores de R_2 compatíveis com o restante da proteína⁴⁹ (Figura 7), indicando a ausência de movimentações lentas. Portanto, é provável que a mutação da cisteína catalítica da Cdc25B tenha alterado a dinâmica conformacional do segmento LCT.

Buhrman *et al.*⁴⁷ propuseram que a formação de uma ponte dissulfeto entre as cisteínas 473 (catalítica) e 426 é um mecanismo de inibição da atividade enzimática da Cdc25B. Portanto, a cisteína catalítica encontra-se reduzida e, possivelmente, ionizada⁴⁵ na enzima ativa. Através do CS de $^{13}\text{C}^\beta$ foi possível verificar que as todas as cisteínas

assinaladas não estão envolvidas na formação de pontes dissulfeto (Figura 21), sugerindo que a enzima está ativa nas condições experimentais utilizadas. Contudo, seus estados de ionização não puderam ser determinados, sendo possível avaliá-los através da dependência dos CS de $^{13}\text{C}^\beta$ em função do pH, conforme realizado por Lim *et al.*⁹².

No geral, a consistência entre os ângulos diedrais previstos por TALOS-N e os extraídos da estrutura cristalográfica (Figura 22) evidenciam o correto assinalamento da proteína e confirmam a eficiência das previsões deste programa. A proporção de ambiguidades e inconsistências (frente à estrutura cristalográfica) na previsão dos ângulos obtida para a construção *C1* ($\sim 10\%$) também está de acordo com aquelas obtidas na validação do programa⁵⁹. A inconsistência entre os diedrais previstos e observados do segmento 437–439 (Figura 22) sugere que a conformação cristalográfica desta região é pouco populada em solução. Esse segmento faz contatos com outra molécula no cristal, caracterizados por ligações de hidrogênio entre o grupo ϵNH^{3+} da 438K e a carbonila dos resíduos 382Y e 383S e a carbonila da 439T e o grupo amino terminal (Figura 30). Além disso, o parâmetro de ordem do resíduo 439 ($S^2 = 0,81$), abaixo do valor médio obtido para os resíduos do núcleo ($\overline{S^2} = 0,94$), e a contribuição de R_{ex} para a relaxação da magnetização transversal (Anexo B), evidenciam a presença de flexibilidade conformacional em diferentes escalas de tempos. Assim, a inconsistência entre os diedrais previstos e observados na estrutura cristalográfica do segmento 437–439 sugere que a cristalização da proteína estabilizou uma conformação local pouco populada em solução.

Figura 30 – Contatos entre diferentes moléculas na unidade assimétrica observados na estrutura cristalográfica da Cdc25B. O segmento 437–439 está representado em laranja, oxigênio em vermelho e nitrogênio em azul. A distância doador–aceptor das ligações de hidrogênio estão ilustradas por linhas tracejadas.

Para a previsão dos ângulos diedrais φ/ψ , o algoritmo de TALOS-N utiliza os CS

do próprio resíduo, de dois posteriores e dois anteriores. Dessa maneira, são utilizados os CS de pentapeptídeos da sequência de aminoácidos da proteína na previsão dos ângulos diedrais do resíduo central. Como o resíduo 473C localiza-se no término do segmento assinalado 468–473 é possível que a falta de CS de dois resíduos posteriores tenham contribuído para a previsão de um ângulo ψ que está em desacordo com a estrutura cristalográfica (Figura 31).

Figura 31 – Gráfico de Ramachandran gerado pelo programa TALOS-N com a previsão dos ângulos diedrais de 473C. O valor dos ângulos φ/ψ da estrutura cristalográfica está representado em azul e as 25 melhores concordâncias do banco de dados estão representadas em vermelho. Em cinza está delimitada a região do gráfico de Ramachandran de resíduos de cisteína presentes no banco de dados de TALOS-N.

A semelhança entre as estruturas secundárias cristalográficas e previstas por TALOS-N ou CSI (Figura 23) e os RDCs observados e calculados (Figura 27B) evidenciam a similaridade estrutural entre o domínio catalítico da Cdc25B no cristal e em solução. Esses resultados estão de acordo com aqueles obtidos por Lund e Cierpicki⁴⁹ e também sugerem que o desenovelamento da hélice C-terminal observado por simulações de dinâmica molecular⁴⁸ e de Monte Carlo baseada em biblioteca de rotâmeros⁴⁶ não é observado experimentalmente. Essas discordâncias podem estar relacionadas com imprecisões nos campos de força e/ou com o reduzido número de contatos nativos presentes entre a hélice C-terminal e o núcleo da proteína, utilizados para estabilizar as estruturas geradas na simulação por Monte Carlo⁴⁶.

Em acordo com o experimento de difração de raios-X⁴³, onde não foi observada densidade eletrônica para os 16 resíduos C-terminais (551–566) da Cdc25B, os parâmetros de ordem S^2 do CTD mostram a grande amplitude de movimentos internos na escala de ps–ns que este segmento apresenta em solução (Figura 26), sugerindo sua desordem

conformacional. Esta característica atenuaria o alinhamento dos vetores NH no meio anisotrópico, refletindo em valores de $D_{\text{NH}} \sim 0$ (Equação (2.3)), conforme observado na figura 27. Dessa maneira, os D_{NH} também apontam para a desordem no CTD com flutuações conformacionais numa escala de tempo mais ampla (ps–ms) que as obtidas pelo formalismo de Lipari-Szabo. Como o CTD participa no reconhecimento do substrato natural da Cdc25B⁵⁰, o complexo Cdk2/CiclinaA, este fato parece ser mais um exemplo da participação de regiões intrinsecamente desordenadas na sinalização celular¹⁵.

Em geral, os resultados de CSP entre as construções *C1* e *C2* são coerentes com a estrutura cristalográfica, uma vez que as regiões da proteína próximas ao C-terminal apresentaram os maiores valores (Figura 29). Os resíduos 532, 533 e 534 do LCT apresentaram valores de CSP muito elevados em comparação aos demais, indicando perturbação desta região após remoção dos 32 resíduos do C-terminal da construção *C1*. De fato, altos valores de RCI demonstram que o segmento LCT fica desestruturado na construção *C2* (dado não apresentado). Assim, a desestruturação do LCT pode explicar a perturbação do deslocamento químico observada nos segmentos vermelho e azul da figura 29. Além disso, esses resultados sugerem que os segmentos do núcleo coloridos em laranja e amarelo na figura 29B fazem contatos com o HCT e/ou CTD. Devido à menor distância do CTD, é provável que este segmento desordenado da região C-terminal da Cdc25B popule estados conformacionais que realizam contatos com os segmentos amarelo e/ou laranja (Figura 29).

Devido à proximidade com o sítio catalítico (*P-loop*; Figura 29), a região C-terminal da Cdc25B pode afetar a afinidade de inibidores competitivos. Nesse sentido, a obtenção de um *ensemble* conformacional da Cdc25B, sobretudo do segmento CTD, pode contribuir para uma estimativa mais precisa da afinidade de inibidores competitivos.

3 Modelagem molecular do domínio catalítico

3.1 Metodologia

3.1.1 Simulações de dinâmica molecular com potenciais clássicos

Simulações de dinâmica molecular podem ser baseadas no uso de potenciais clássicos para a descrição da energia potencial ($\mathcal{E}_t$) de um sistema, constituídos por um termo referente à energia das ligações covalentes ($\mathcal{E}_{cov}$) e outro referente às energias de interações não-covalentes ($\mathcal{E}_{ncov}$):

$$\mathcal{E}_t = \mathcal{E}_{cov} + \mathcal{E}_{ncov}$$

O termo covalente é composto pela somatória das energias de ligação, ângulo e diedrais próprios e impróprios de todos os átomos covalentemente ligados. As energias da distância de ligação ($\mathcal{E}_{lig}$) entre os átomos i e j , do ângulo de ligação ($\mathcal{E}_{ang}$) entre os átomos i , j e k e do diedral impróprio ($\mathcal{E}_{imp}$) entre os átomos i (central), j , k e l podem ser descritas por potenciais harmônicos:

$$\mathcal{E}_{lig}(r_{ij}) = \sum_{lig} \frac{k_r}{2} (r_{ij} - r_{ij}^0)^2 \quad (3.1)$$

$$\mathcal{E}_{ang}(\theta_{ijk}) = \sum_{ang} \frac{k_\theta}{2} (\theta_{ijk} - \theta_{ijk}^0)^2 \quad (3.2)$$

$$\mathcal{E}_{imp}(\zeta_{ijkl}) = \sum_{imp} \frac{k_\zeta}{2} (\zeta_{ijkl} - \zeta_{ijkl}^0)^2, \quad (3.3)$$

onde r , θ e ζ são a distância $\vec{ij}$, o ângulo formado por i , j e k e a distância entre i e o plano $jl k$, respectivamente, com valores de equilíbrio r^0 , θ^0 e ζ^0 e k são constantes de força específicas para cada conjunto de átomos.

O termo que descreve a energia dos ângulos diedrais ($\mathcal{E}_{died}$) de átomos i , j , k e l sequencialmente ligados, pode ser descrito por uma soma de funções periódicas com diferentes períodos e constantes de força:

$$\mathcal{E}_{died}(\phi_{ijkl}) = \sum_{died} \sum_n \frac{k_\phi^n}{2} (1 + \cos(n\phi - \phi_f)), \quad (3.4)$$

onde ϕ é o ângulo formado pelos planos ijk e jkl , n é a periodicidade do ângulo e ϕ_f é a fase.

O termo não-covalente ($\mathcal{E}_{ncov}$) descreve a energia de interação entre átomos da mesma ou de diferentes moléculas que não estão ligados covalentemente. Neste termo, estão incluídas contribuições eletrostáticas ($\mathcal{E}_{elet}$), descritas pela lei de Coulomb e da energia de Lennard-Jones ($\mathcal{E}_{LJ}$), utilizada como uma aproximação para o cálculo das forças de van der Waals:

$$\mathcal{E}_{elet}(r_{ij}) = \sum_{i < j} \frac{q_i q_j}{r_{ij}} \quad (3.5)$$

$$\mathcal{E}_{LJ}(r_{ij}) = \sum_{i < j} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right], \quad (3.6)$$

onde i e j representam pares de átomos com cargas parciais q localizados a uma distância r_{ij} , ϵ_{ij} e σ_{ij} são o valor mínimo e sua posição no potencial de Lennard-Jones, respectivamente.

Para calcular-se $\mathcal{E}_t$ de um sistema molecular, utilizam-se parâmetros fornecidos por campos de força, como Amber⁹³ ou CHARMM⁹⁴ que foram desenvolvidos para biomoléculas.

No procedimento de dinâmica molecular computacional a trajetória dos átomos de uma molécula é descrita pela propagação das coordenadas iniciais ao longo do tempo. Para realizar este procedimento, as funções empíricas de energia descritas nas equações (3.1)–(3.6) são utilizadas para o cálculo da força exercida em cada átomo ($\vec{F}_i$) com a equação (3.7). Tais forças são usadas na integração da equação de Newton (equação (3.8)) para que sejam geradas novas posições e velocidades dos átomos em curtos passos de tempo:

$$\vec{F}_i = -\frac{\partial \mathcal{E}_t}{\partial \mathbf{r}_i} \quad (3.7)$$

$$\vec{F}_i = m_i \frac{\partial^2 \mathbf{r}_i}{\partial t^2} = m_i \mathbf{a}_i, \quad (3.8)$$

onde m é a massa, $\mathbf{r}$ é a matriz de coordenada espacial e $\mathbf{a}$ é o vetor da aceleração do átomo i e t é o tempo.

O algoritmo de Verlet⁹⁵ é um dos métodos mais simples e eficientes para integrar a equação (3.8) a fim de gerar novas coordenadas atômicas em um intervalo de tempo Δt :

$$\mathbf{r}(t + \Delta t) = 2\mathbf{r}(t) - \mathbf{r}(t - \Delta t) + \mathbf{a}(t)\Delta t^2. \quad (3.9)$$

Os intervalos de tempo Δt (Equação (3.9)) utilizados na integração da equação (3.8) são normalmente da ordem de femtosegundos. Assim, simulações de dinâmica molecular, comumente realizadas por um período de nanosegundos, podem gerar milhões de geometrias de uma molécula.

A sequência de aminoácidos do modelo estrutural usado nas simulações de dinâmica molecular corresponde à construção *C1* do domínio catalítico da Cdc25B (Figura 10) usada na aquisição das medidas experimentais de RMN. Essa construção contém 7 resíduos N-terminais não observados na proteína nativa em virtude da presença de sequências adaptadoras no *primer* de clonagem e do sítio de trombina no vetor de expressão (GSHMEFQ+373S–566Q). No experimento de difração de raio-X do domínio catalítico da Cdc25B⁴³, a densidade eletrônica dos últimos 16 resíduos do C-terminal não foi resolvida. Também não se conhece a estrutura do segmento GSHMEFQ adicionado no N-terminal. Assim, ambas regiões foram manualmente adicionadas à estrutura cristalográfica (PDB: 1QB0) numa conformação estendida. Moléculas cristalográficas de água, íons e β -mercaptoetanol foram removidos. Os átomos de hidrogênio foram construídos considerando as cadeias laterais de Asp, Glu, Arg e Lys carregadas e os resíduos e Cys reduzidos e neutros, exceto 473C cujo estado ionizado é aceito como cataliticamente ativo⁴⁵. Somente o íon PO_4^{3-} presente no *P-loop* da estrutura cristalográfica foi mantido na estrutura inicial.

Essa estrutura da Cdc25B foi usada numa simulação de dinâmica molecular (DM) de 20 ns com solvente implícito onde o raio atômico de Born foi calculado pelo modelo OBC⁹⁶. A posição dos átomos pesados presentes na estrutura cristalográfica foi restringida, aplicando-se uma constante harmônica de 1000 kJ/mol $\cdot$ nm² em todas as direções. Assim, somente os segmentos terminais construídos manualmente amostraram o espaço conformacional. Dessa simulação foram retiradas duas geometrias da Cdc25B (**A** e **B**) que posteriormente foram utilizadas como coordenadas iniciais em simulações longas de 2 μ s (Tabela 7) com solvente explícito.

A construção do modelo, a análise dos resultados e as dinâmicas moleculares foram conduzidas com o pacote de programas GROMACS 4.6.1⁶⁸. Foram utilizados os campos de força Amber99SB-ILDN^{97, 98, 69} ou CHARMM36^{99, 100} para os átomos da proteína e

íons e o modelo TIP3P¹⁰¹ para as moléculas de água. No total, 3 trajetórias de dinâmica molecular (Tabela 7) foram realizadas por 2 μ s a 300 K e 1 atm com passos de integração de 2 fs em um *ensemble* NPT, onde o número de moléculas, a pressão e a temperatura do sistema são mantidas constantes. Foram utilizados o algoritmo de Berendsen *et al.*¹⁰² para ajustar a pressão do sistema a cada 0,5 ps e o método de reescalonamento de velocidades de Bussi *et al.*¹⁰³ para controlar a temperatura. O algoritmo LINCS¹⁰⁴ foi usado para restringir a distância das ligações covalentes. Condições periódicas de contorno foram tratadas em caixas dodecaédricas replicadas no espaço. A estrutura inicial da proteína foi solvatada e centralizada na caixa com uma distância de pelos menos 10 Å das bordas onde também foram adicionados cinco íons Cl^- e dois Na^+ para neutralizar as cargas do sistema. O método PME¹⁰⁵ foi usado no cálculo das interações eletrostáticas de longo alcance com espaçamento máximo de 0,15 nm na grade de Fourier e interpolação de quarta ordem. As forças de Van der Waals foram truncadas a distâncias maiores de 0,9 nm.

Tabela 7 – Trajetórias de dinâmica molecular de 2 μ s do domínio catalítico da Cdc25B.

Trajeto�ria	Estrutura inicial ^a	Campo de for�a
<i>DM1</i>	A	Amber99SB-ILDN
<i>DM2</i>	B	Amber99SB-ILDN
<i>DM3</i>	B	CHARMM36

^a retirada da dinmica molecular em solvente implcito

As simulaes de dinmica molecular de 2 μ s (Tabela 7) foram concatenadas em uma nica trajetria de 6 μ s (*DM1+DM2+DM3*), que no pode ser interpretada como uma trajetria de DM contnua. Desta trajetria concatenada foram obtidas 60.000 estruturas da Cdc25B retiradas da simulao a cada intervalo de 100 ps para as anlises descritas nas sees seguintes.

A flexibilidade conformacional de um tomo i durante uma simulao de DM pode ser avaliada pela flutuao mdia (RMSF) de suas coordenadas $\mathbf{r}_i$ ao longo do tempo t :

$$\text{RMSF}_i = \sqrt{\frac{1}{N} \sum_{t=1}^N \|\bar{\mathbf{r}}_i - \mathbf{r}_i^t\|^2}, \quad (3.10)$$

onde N o nmero de estrutura e $\bar{\mathbf{r}}$ a coordenada mdia.

A similaridade entre duas estruturas proteicas foi expressa pelo desvio médio quadrático (RMSD) entre suas coordenadas atômicas:

$$\text{RMSD}(\mathbf{s}, \mathbf{t}) = \sqrt{\frac{1}{N} \sum_{j=1}^N \|\mathbf{s}_j - \mathbf{t}_j\|^2}, \quad (3.11)$$

onde a somatória é feita sobre N átomos dos conjuntos de coordenadas $\mathbf{s}$ e $\mathbf{t}$.

3.1.2 Previsão dos deslocamentos químicos

As estruturas da Cdc25B retiradas da simulação de dinâmica molecular de 6 μs foram usadas para a previsão dos CS de ^1H , ^{15}N , $^{13}\text{C}^\alpha$, $^{13}\text{C}^\beta$ e $^{13}\text{C}'$ com o programa SPARTA+⁶². Dessa maneira, foi possível obter-se as flutuações dos CS de diferentes núcleo da proteína ao longo do tempo da simulação. Todos resíduos nos arquivos PDB gerados pelo programa `trjconv` do pacote GROMACS foram renomeados de acordo com o código convencional de 3 letras antes da previsão dos CS.

A previsão dos CS pelo programa SPARTA+ é baseada na dependência geométrica, dinâmica e da similaridade de um resíduo e seus vizinhos. Essa dependência é estabelecida através rede neurais artificiais (ANN) treinadas sobre um banco de dados composto por proteínas cuja estrutura e CS são conhecidos. Para tal finalidade, os parâmetros estruturais de tripeptídeos da estrutura primária (ângulos diedrais φ , ψ , χ_1 e χ_2 , ligações de H, parâmetros de ordem, similaridade de sequência) são representados na camada de entrada da ANN por uma matriz com 113 nós ($X_{1 \times 113}$). Os deslocamentos químicos secundários previstos ($\Delta\delta^{prev}$) na camada de saída são obtidos por:

$$\Delta\delta_{1 \times 1}^{prev} = f_2(f_1(X_{1 \times 113} \times W_{113 \times 30}^1 + b_{1 \times 30}^1) \times W_{113 \times 30}^2 + b_{1 \times 30}^2),$$

onde $f_1(x) = (1 - e^{-2x})/(1 + e^{-2x})$, $f_2(x) = x$, $\mathbf{W}^1$ e $\mathbf{b}^1$ são matrizes peso e tendência, respectivamente, para a conexão entre as camadas de entrada e intermediária e $\mathbf{W}^2$ e $\mathbf{b}^2$ são matrizes semelhantes para a conexão entre as camadas intermediária e de saída.

Para a previsão dos CS (δ^{prev}) são então adicionados aos valores de $\Delta\delta^{prev}$ as contribuições do efeito de anel aromático (δ^{AA}), campo elétrico (δ^{CE}) e deslocamento químico para estruturas desordenadas (δ^{RC}):

$$\delta^{prev} = \Delta\delta^{prev} + \delta^{AA} + \delta^{CE} + \delta^{RC}. \quad (3.12)$$

Valores médios dos CS (δ_{DM}) de $^{13}\text{C}^\alpha$, $^{13}\text{C}^\beta$, $^{13}\text{C}'$, ^1H e ^{15}N foram calculados a partir dos valores previstos de todas estruturas retiradas das trajetórias de DM e comparados aos valores experimentais de CS (δ_{Exp}). Também foram calculadas médias cumulativas dos deslocamentos químicos ($\delta_{\text{DM}}^{\text{cml}}$), utilizando um número crescente de estruturas retiradas longo das trajetórias, com o intuito avaliar o tempo de convergência.

A previsão dos CS também foi feita para a mesma trajetória onde os 32 resíduos do C-terminal foram apagados dos arquivos PDB com o objetivo de simular os deslocamentos químicos da construção *C2*. Optou-se por essa abordagem, ao invés de se calcular os deslocamentos químicos a partir de uma simulação de DM da construção *C2*, para se evitar diferenças de amostragem conformacional. Assim, os contatos entre os resíduos do C-terminal e o núcleo proteico seriam as únicas contribuições para eventuais diferenças entre os CS previstos nas duas construções.

Utilizando-se essa abordagem, foi possível calcular as perturbações do deslocamento químico (Equação (2.1)) a partir dos CS de ^1H e ^{15}N previstos das construções *C1* e *C2* a fim de compará-los com os valores experimentais (Figura 29).

Os deslocamentos químicos obtidos da simulação de dinâmica molecular da construção *C1* (δ_{DM}) e previstos da estrutura cristalográfica da Cdc25B (PDB:1QB0) (δ_{PDB}) foram utilizados no cálculo de um fator Γ que reflete o acordo entre os CS experimentais e previstos:

$$\mathcal{P}^x = |\delta_{\text{PDB}}^x - \delta_{\text{Exp}}^x| - |\delta_{\text{DM}}^x - \delta_{\text{Exp}}^x| \quad (3.13)$$

$$\Gamma_i = \frac{\sum_x \mathcal{P}_{\text{norm}}^x}{W_i}, \quad (3.14)$$

onde $\mathcal{P}_{\text{norm}}^x = \mathcal{P}^x / \mathcal{P}_{\text{max}}^x$, $x = (\text{H}, \text{N}, \text{C}^\alpha, \text{C}^\beta, \text{C}')$ e W_i é o número de átomos assinalados para o resíduo i . Assim, Γ_i pode assumir valores $-1 \leq \Gamma_i \leq 1$. Valores de $\Gamma_i < 0$ indicam que a descrição cristalográfica do resíduo i está em maior acordo com δ_{Exp} . Quando $\Gamma_i > 0$, δ_{DM} está em maior acordo com δ_{Exp} .

3.2 Resultados e discussão

3.2.1 Simulações da construção *C1*

No capítulo 2 foi demonstrado através de diferentes medidas de RMN e de previsões computacionais que o domínio catalítico da Cdc25B em solução possui regiões terminais flexíveis e um núcleo enovelado de acordo com a estrutura cristalográfica⁴³. Além disso, verificou-se que a região C-terminal flexível pode fazer contatos com diferentes regiões do núcleo da proteína. Com o intuito de investigar com mais detalhes a dinâmica conformacional da Cdc25B, simulações longas de dinâmica molecular foram realizadas (Seção 3.1.1) e avaliadas, principalmente, através da comparação entre propriedades calculadas e experimentais (Seção 3.1.2).

As geometrias iniciais da proteína usadas nas simulações de DM com solvente explícito (Tabela 7) foram retiradas de trajetórias de dinâmica molecular com solvente implícito (Seção 3.1.1). Essas estruturas iniciais diferem principalmente no segmento C-terminal. Enquanto na estrutura inicial da trajetória *DM1*, o C-terminal encontra-se mais exposto ao solvente, alguns resíduos na estrutura inicial das simulações *DM2* e *DM3* fazem contatos com a hélice C-terminal (Figura 32).

Figura 32 – Estruturas iniciais usadas nas simulações *DM1* e *DM2/DM3* da Cdc25B. Os segmentos terminais cujas estruturas foram amostradas por simulação em solvente implícito estão representados em vermelho. O íon PO_4^{3-} (Pi) presente no P-loop da estrutura cristalográfica (PDB: 1QB0) está representado na forma de bastão.

Os valores de RMSD das estruturas da Cdc25B obtidos ao longo das trajetórias de DM (Tabela 7) foram calculados, utilizando as coordenadas iniciais como referência

(Figura 33). No início das simulações ocorre um rápido aumento nos valores RMSD devido à relaxação das coordenadas atômicas no campo de força. Esse aumento inicial do RMSD foi mais expressivo para o CTD nas três trajetórias, onde apresentou valores $\sim 0,5 - 1,5$ nm (verde; Figura 33). Como suas coordenadas iniciais (Figura 32) foram obtidas de dinâmica molecular com solvente implícito, a descrição explícita do solvente nas trajetórias *DM1-3* além da flexibilidade intrínseca do CTD em solução podem ter contribuído para uma alteração mais expressiva na sua geometria. Por outro lado, o RMSD do núcleo nas simulações *DM1-3* flutuou em $\sim 0,1$ nm (Figura 33), sugerindo a estabilidade da estrutura cristalográfica em solução. Contudo, valores mais elevados foram observados após $\sim 0,8$ μ s na *DM3* em decorrência de uma mudança conformacional no LCT que pode ser notada pelo aumento do RMSD para $\sim 0,5$ nm nesse tempo da simulação (vermelho; Figura 33).

A desestruturação do HCT, observada em simulação de dinâmica molecular de 60 ns⁴⁸ com o campo de força OPLS-AA¹⁰⁶, não foi confirmada nas trajetórias *DM1-3*, onde os valores de RMSD $\sim 0,2$ nm (ciano; Figura 33) apontam para sua estabilidade durante as simulações. Essa observação sugere que o campo OPLS-AA desestabiliza o HCT. Foi demonstrado que modificações na composição da região terminal influenciam a helicidade de peptídeos poli-Ala¹⁰⁷ em simulações de DM. Assim, a inclusão dos resíduos C-terminais (551–566), ausentes na dinâmica molecular de 60 ns⁴⁸, também pode ter contribuído para aumentar a estabilidade da hélice.

Os valores RMSF (Figura 34) das simulações da Cdc25B apontam para a flexibilidade das regiões terminais e uma relativa rigidez conformacional do núcleo, conforme sugerido pelas medidas de RMN da construção *C1* (Figuras 25–27). Além desses acordos qualitativos, os valores de RMSF das simulações podem ser diretamente comparados com os RMSF estimados dos RCI^{60, 108} (Figura 34). Já que os RCI são calculados a partir dos CS, a comparação com os RMSF das trajetórias é uma maneira de comparar as simulações de dinâmica molecular com medidas experimentais.

A flexibilidade do CTD mostrou-se diferente nas três trajetórias, apresentando maior valor médio de RMSF na *DM1* ($\overline{\text{RMSF}} = 0,56$ nm) comparada com a *DM2* ($\overline{\text{RMSF}} = 0,39$ nm) e *DM3* ($\overline{\text{RMSF}} = 0,22$ nm). No entanto, o melhor acordo visual entre

Figura 33 – Desvio médio quadrático (RMSD) das coordenadas de N da cadeia principal das trajetórias de dinâmica molecular *DM1*, *DM2* e *DM3* da Cdc25B (Tabela 7). Os valores de RMSD foram calculados para diferentes segmentos da proteína, utilizando as coordenadas iniciais de cada simulação como referência.

Figura 34 – Flutuação média (RMSF) de N da cadeia principal dos resíduos da Cdc25B. Os valores foram calculados das simulações de dinâmica molecular (*DM1–3*) e estimados dos valores de RCI⁶⁰.

os valores de RMSF_{RCI} e os RMSF das trajetórias *DM1* e *DM2* (Figura 34) indica que a flexibilidade do CTD nessas trajetórias está razoavelmente bem representada frente às evidências experimentais.

No geral, os perfis de RMSF das trajetórias *DM1–3* são similares entre si e estão em acordo com o RMSF_{RCI} (Figura 34). As exceções localizam-se na região C-terminal da simulação *DM3*, onde resíduos no LCT/HCT (527–534) apresentaram maior flexibilidade

conformacional (Figura 34). A maior flexibilidade desses resíduos na *DM3* é decorrente da formação de ligações de hidrogênio entre os resíduos 534 O...H₅₃₈ e 535 O...H₅₃₉ que ocorrem em $\sim 0,8 \mu\text{s}$ de simulação (Figura 33), tornando a hélice HCT contínua e promovendo a abertura do LCT (Figura 35). A formação de uma hélice contínua a partir de ligações de hidrogênio entre os segmentos 534–536 e 537–546 durante a simulação de DM sugere a existência de um equilíbrio conformacional entre o HCT contínuo e descontínuo. Essa observação está em acordo com as populações em α -hélice dos segmentos 534–536 e 537–546 de $\sim 50\%$ e $> 70\%$ (Figura 24), respectivamente, calculadas pelo método $\delta 2D$ ⁸⁶.

Durante os $6 \mu\text{s}$ acumulados de todas as simulações, essa transição de estados ($\text{HCT}_{\text{descontínuo}} \longrightarrow \text{HCT}_{\text{contínuo}}$) ocorre duas vezes (Figura 35). Assim, é provável que cinética da transição ocorra em uma escala de tempo da ordem de μs . Em acordo com essa observação, as contribuições de troca conformacional (R_{ex}) para a relaxação dos *spins* de ^{15}N dos resíduos 527, 528 e 533 do LCT e 534 e 538 do HCT (Figura 26B) indicam a presença de movimentações na escala de μs – ms . Assim, contribuições de R_{ex} presentes nos resíduos do segmento LCT/HCT podem ser um reflexo dessa mudança conformacional observada na trajetória *DM3*. Contudo, essa interpretação deve ser feita com cautela, já que a transição $\text{HCT}_{\text{descontínuo}} \longrightarrow \text{HCT}_{\text{contínuo}}$ ocorreu somente na simulação de DM com parâmetros do campo de força CHARMM36, apesar da sua parametrização ter sido motivada principalmente pela superestabilização de segmentos em hélice observada em versões anteriores do campo de força CHARMM^{107, 109}. Além disso, a significância estatística dessa transição não é expressiva no momento, já que foram realizadas somente três simulações. Assim, para se avaliar com maior confiabilidade a transição $\text{HCT}_{\text{descontínuo}} \longrightarrow \text{HCT}_{\text{contínuo}}$, um número maior de simulações com diferentes condições iniciais devem ser realizadas.

Nos períodos iniciais das três simulações de DM foi observada a dissociação do íon PO_4^{3-} do *P-loop* (472–479) (Figura 36A) e a formação de outros complexos estabilizados pela ligação entre o fosfato e cadeias laterais de resíduos de Arg ou Lys da região C-terminal, como ilustrado na figura 36B. Apesar desse desacordo entre a dinâmica molecular e a presença do íon SO_4^{2-} complexado ao *P-loop* na estrutura cristalográfica (PDB: 1QB0), torna-se importante observar que os cristais da Cdc25B foram obtidos em concentrações

Figura 35 – Mudança conformacional da cadeia principal dos segmentos HCT e LCT observada na simulação *DM3* (Figura 33). A formação da ligação de H 534O...H538 é representada pela distância entre O de 534E e H de 538D.

de SO_4^{2-} cerca de 1000 vezes maiores que as usadas nas simulações. Assim, nas condições cristalográficas, o equilíbrio é deslocado para a formação do complexo SO_4^{2-} -P-loop. Além disso, a dissociação do complexo PO_4^{3-} -P-loop logo no início das simulações de DM pode ter sido favorecida pelo estado cataliticamente ativo (Cys-S⁻)⁴⁵ da cisteína 473. Os sítios alternativos de ligação a PO_4^{3-} evidenciados pelas simulações de DM da Cdc25B podem explicar a maior estabilidade da proteína em tampão NaPi em comparação a tampão Tris.

3.2.2 Comparação dos modelos estruturais a partir de medidas experimentais

A estrutura cristalográfica da Cdc25B é constituída pelo segmento 374–550 da proteína. Assim, os modelos estruturais desse segmento representados pela geometria cristalográfica ou pelo conjunto de estruturas das simulações de DM foram comparados entre si nas seções 3.2.2.1 e 3.2.2.2 através do acordo entre os CS previstos e experimentais.

Os deslocamentos químicos experimentais reportam uma média ponderada pela energia livre dos estados visitados por uma proteína em solução. Logo, os CS médios calculados a partir de trajetórias de DM (δ_{DM}) podem apresentar maior acordo com os valores experimentais (δ_{Exp}) do que os CS previstos a partir de poucas ou uma única estrutura estática (cristalográfica)^{110, 111} (δ_{PDB}).

Figura 36 – Descomplexação do íon PO_4^{3-} do P-loop durante as trajetórias de DM. (A) Distâncias entre os centros de massa do P-loop (472–479) e do íon PO_4^{3-} . (B) Estrutura representativa de um complexo formado entre o íon PO_4^{3-} e cadeias laterais de Arg durante a trajetória DM1.

Contudo, em geral os δ_{DM} da Cdc25B não apresentaram maior acordo com os δ_{Exp} quando comparados aos δ_{PDB} . Na tabela 8 observa-se que os RMSD dos δ_{PDB} em relação aos valores experimentais são similares aos RMSD dos δ_{DM} . Esses resultados estão em desacordo com os trabalhos de Markwick *et al.*; Robustelli *et al.*^{110, 111}, onde foram observados maiores acordos entre δ_{DM} e δ_{Exp} quando comparados aos δ_{PDB} previstos de proteínas com estruturas cristalográficas com resolução de 2,7 Å e 2,8 Å, respectivamente. A estrutura cristalográfica da Cdc25B, por sua vez, apresentou resolução significativamente maior (1,9 Å). Assim, é provável que as simulações de DM conduzidas nesses trabalhos tenham amostrado melhor os mínimos locais da geometria proteica, aumentando significativamente o acordo com entre δ_{DM} e δ_{Exp} .

Essa dependência da resolução da estrutura cristalográfica nos desvios entre CS previstos e experimentais já foi relatada no trabalho de Robustelli *et al.*¹¹¹. Neste trabalho, foram observados valores semelhantes de RMSD de δ_{DM} e δ_{PDB} previstos de uma proteína com estrutura cristalográfica de alta resolução (1,5 Å), demonstrando que neste caso a geometria cristalográfica é uma representação média razoável dos estados conformacionais

visitados pela proteína em solução.

Somente os valores médios de CS calculados a partir das simulações de DM concatenadas (*DM1+2+3*) serão avaliados a seguir.

Tabela 8 – Desvio médio quadrático (RMSD) entre os deslocamentos químicos experimentais e calculados da estrutura cristalográfica ou trajetórias de DM da Cdc25B. Os valores de RMSD foram calculados apenas para os resíduos presentes na estrutura cristalográfica (373–550).

	RMSD ($\delta_{\text{PDB/DM}}, \delta_{\text{Exp}}$)					
	^1H	^{15}N	$^{13}\text{C}^\alpha$	$^{13}\text{C}^\beta$	$^{13}\text{C}'$	Soma
PDB	0,52	2,53	0,98	1,07	1,18	6,28
<i>DM1</i>	0,51	2,60	1,10	1,13	1,23	6,57
<i>DM2</i>	0,50	2,61	1,08	1,14	1,20	6,53
<i>DM3</i>	0,51	2,76	1,13	1,22	1,21	6,83
<i>DM1+2+3</i>^a	0,50	2,58	1,06	1,13	1,19	6,46

^a trajetórias *DM1*, *DM2* e *DM3* concatenadas

Nota-se que para a grande maioria dos resíduos, os desvios de δ_{DM} ou δ_{PDB} em relação aos valores experimentais são menores que as incertezas intrínsecas do programa SPARTA+⁶² (Anexo C) na previsão dos deslocamentos químicos para cada tipo atômico (incertezas para $^{13}\text{C}^\alpha$: 0,94 ppm; $^{13}\text{C}^\beta$: 1,14 ppm; $^{13}\text{C}'$: 1,09 ppm; ^1H : 0,49 ppm; ^{15}N : 2,45 ppm). Assim, não é possível afirmar para tais resíduos se os CS experimentais estão em melhor acordo com os valores previstos da estrutura cristalográfica ou calculados das trajetórias de DM. Contudo, desvios significativos também foram observados, indicando que em determinadas regiões da proteína a geometria cristalográfica local ou os estados conformacionais populados nas simulações estão em melhor acordo com os dados espectroscópicos (Anexo C). Os átomos que apresentaram melhoria significativa no acordo com os CS experimentais estão listados na tabela 9.

As regiões da proteína que apresentaram melhor acordo com a estrutura cristalográfica apontam para a presença de deficiências no campo de força ou amostragem insuficiente das simulações. Já as regiões que mostraram-se em melhor acordo com as trajetórias de DM sugerem a presença de artefatos de cristalização ou a flexibilidade conformacional, uma vez que essas regiões poderiam ser melhor representadas por um conjunto de estruturas ao invés de uma estrutura estática. Essas possibilidades serão exploradas a seguir na análise

das trajetórias do CS das regiões da proteína que apresentaram melhor acordo com a estrutura cristalográfica ou com as simulações de DM.

Tabela 9 – Deslocamentos químicos da Cdc25B em significativo maior acordo com os valores previstos da estrutura cristalográfica (PDB) ou com os valores previstos das trajetórias de dinâmica molecular (DM). Nesta tabela estão listados os átomos cuja magnitude do desvio $|\delta_{\text{PDB}} - \delta_{\text{Exp}}| - |\delta_{\text{DM}} - \delta_{\text{Exp}}|$ (Anexo C) é maior que a precisão do programa SPARTA+ na previsão do deslocamento químico.

PDB		DM	
Resíduo	Núcleo	Resíduo	Núcleo
415	$^{13}\text{C}^\alpha$	395	$^{13}\text{C}^\beta$
424	$^{13}\text{C}^\alpha$	395	^{15}N
428	$^{13}\text{C}^\alpha$	396	^1H
448	$^{13}\text{C}^\alpha$	408	^1H
452	$^{13}\text{C}^\alpha$	418	$^{13}\text{C}^\alpha$
455	$^{13}\text{C}^\alpha$	423	$^{13}\text{C}^\beta$
491	$^{13}\text{C}'$	441	$^{13}\text{C}^\beta$
494	$^{13}\text{C}^\alpha$	452	^{15}N
494	^{15}N	531	^1H
495	$^{13}\text{C}^\beta$	536	$^{13}\text{C}^\alpha$
496	$^{13}\text{C}^\beta$	536	$^{13}\text{C}^\beta$
497	^{15}N	537	^1H
522	$^{13}\text{C}^\alpha$	549	$^{13}\text{C}'$
529	^{15}N	549	$^{13}\text{C}^\alpha$
531	$^{13}\text{C}^\alpha$	549	^1H
533	$^{13}\text{C}^\alpha$	549	^{15}N
534	^{15}N		
539	$^{13}\text{C}^\alpha$		

Em virtude do elevado número de átomos da Cdc25B assinalados (753), optou-se por agrupar os deslocamentos químicos calculados em cada resíduo no cálculo de um fator Γ (Equação (3.14)) para facilitar a análise dos dados. Valores negativos de Γ apontam para regiões da proteína onde o maior acordo entre CS experimentais e previstos é um reflexo da geometria cristalográfica. Já as regiões onde Γ é positivo sugerem que o maior acordo com os CS experimentais é um reflexo dos CS previstos de um conjunto de geometrias amostradas durante as simulações de DM (Figura 37A). Assim, foi possível conduzir uma análise sistemática e direcionada para explicar porque regiões da proteína são melhor representas pela geometria cristalográfica ou por um conjunto das conformações amostras nas simulações de DM. Vale ressaltar que as interpretações dos fatores Γ devem considerar a significância dos CS calculados (Tabela 9), já que a imprecisão de SPARTA+ na previsão

não é considerada para o cálculo de Γ .

Figura 37 – Comparação entre os deslocamentos químicos experimentais e previstos da estrutura cristalográfica ou calculados a partir da simulação *DM1+2+3*. (A) Fator Γ (Equação (3.14)) calculado para os resíduos assinalados da Cdc25B. (B) Flutuação média de N da cadeia principal calculada da trajetória *DM1+2+3* (linha preta) ou a partir dos fatores B cristalográficos¹¹² (linha verde): $\text{RMSF}_i = \sqrt{3B_i/8\pi^2}$.

Para orientar as análises das simulações de dinâmica molecular, a sequência primária da Cdc25B foi dividida em segmentos cujos resíduos consistentemente apresentam valores de Γ com mesmo sinal (Figura 37A):

- $\Gamma < 0$: 424–429, 443–456, 491–497 e 531–534;
- $\Gamma > 0$: 375–376, 394–398, 535–537 e 549–550.

Os segmentos com $\Gamma < 0$ localizam-se próximos a sítios cristalográficos de ligação a $\text{SO}_4^{2-}/\text{PO}_4^{3-}$ (Figura 38), sugerindo sua importância na estabilização da geometria cristalográfica. Assim, a ausência do íon PO_4^{3-} do sítio P2 nas estruturas iniciais das simulações de DM e a descomplexação do íon PO_4^{3-} do sítio P1 (Figura 36) podem ter favorecido a desestruturação dessas regiões durante as simulações, conforme discutido a seguir na seção 3.2.2.1. Vale ainda ressaltar que o programa SPARTA+ não considera a perturbação do deslocamento químico causada por ligantes. Portanto, a perturbação do íon

PO_4^{3-} no CS dos resíduos que formam seus sítios de ligação não é incluída no programa, podendo ser uma fonte adicional de erros na previsão dos CS.

Já os segmentos que apresentam $\Gamma > 0$, estão localizados nos terminais da proteína (375–376 e 549–550), na alça 394–398 e na hélice C-terminal (535–537). Dessa forma, a representação desses segmentos por seus estados conformacionais amostrados nas simulações de DM contribuíram para sua melhor descrição estrutural. Esse fato pode ocorrer em virtude da flexibilidade conformacional intrínseca das regiões terminais ou da presença de imprecisões na estrutura cristalográfica corrigidas nas simulações, conforme discutido na seção 3.2.2.2.

Figura 38 – Segmentos do domínio catalítico da Cdc25B cujos deslocamentos químicos experimentais estão em melhor acordo com a geometria cristalográfica ($\Gamma < 0$; verde) ou as geometrias amostradas nas simulações de DM ($\Gamma > 0$; vermelho). Os íons cristalográficos SO_4^{2-} presentes nos sítios P1 e P2 estão representados em bastão.

Para cada um desses segmentos foram analisadas as trajetórias dos deslocamentos químicos das simulações de DM daqueles núcleos cujo melhor acordo dos δ_{Exp} com δ_{DM} ou δ_{PDB} são significativos (Tabela 9).

3.2.2.1 Concordância com a estrutura cristalográfica

Nesta seção serão apresentas e discutidas as características estruturais dos segmentos da Cdc25B cujos CS previstos a partir da estrutura cristalográfica apresentaram maior acordo com os valores experimentais.

Os segmentos 491–497 e 531–534 da Cdc25B apresentaram $\text{RMSF}_{\text{DM}} > \text{RMSF}_{\text{fatorB}}$ e $\Gamma < 0$ (Figura 37), sugerindo que a amostragem nas simulações de DM exagera a flexibilidade conformacional dessas regiões em solução. Nota-se que o desvio entre δ_{Exp}

e δ_{PDB} em diversos resíduos desses segmentos é significativamente menor em relação aos δ_{DM} (Tabela 9).

Nas trajetórias *DM2* e *DM3*, observa-se uma maior flutuação nos RMSD das coordenadas de N da cadeia principal no segmento 491–497 (Figura 39A). Já que $\Gamma < 0$ (Figura 37A), a flexibilidade conformacional do segmento 491–497 nas simulações de DM na ausência de $\text{SO}_4^{2-}/\text{PO}_4^{3-}$ complexado a este sítio não é uma boa representação das conformações importantes em solução com concentrações destes íons que promovem a formação do complexo (condições dos experimentos de RMN e difração de raios-X). A comparação entre δ_{Exp} e δ_{DM} pode ser observada na trajetória de CS previstos de ^{15}N do resíduo 497Y (Figura 39B), onde verifica-se que valores mais próximos aos experimentais são observados de estruturas com menores RMSD estrutural em relação à geometria cristalográfica (Figura 39A).

O segmento 491–497 está localizado no sítio P2 de ligação a PO_4^{3-} observado após titulação com NaPi da proteína em solução (Figura 19; página 61) e presente na estrutura cristalográfica (Figura 39C). Contudo, o íon PO_4^{3-} não foi adicionado ao sítio P2 nas simulações, o que possivelmente favoreceu sua desestabilização nas trajetórias *DM2* e *DM3* e acarretou nos elevados valores de RMSF observados no segmento 491–497 (Figura 37B).

A expressiva flutuação estrutural do segmento 531–534 (Figura 37) está relacionada à mudança conformacional observada em $\sim 0,8 \mu\text{s}$ da simulação *DM3* (Figura 34) que promove a abertura do LCT e a formação de ligações de hidrogênio que tornam o HCT contínuo (Figura 35). Os valores negativos de Γ desse segmento sugerem que a conformação do HCT contínua não é significativamente populada em solução. Contudo, as contribuições de R_{ex} (Figura 26; página 69) indicam a presença de movimentações na escala de μs – ms nesta região, conforme discutido na página 87. Dessa maneira, o desacordo entre δ_{DM} e δ_{Exp} em diversos resíduos do segmento 531–534 (Tabela 9) pode ser um reflexo de deficiências no balanço populacional das conformações aberta e fechada do LCT, principalmente na simulação *DM3* (CHARMM36). Portanto, embora a conformação do HCT contínuo deva ser pouco estável em solução, a simulação *DM3* sugere a ocorrência da transição $\text{HCT}_{\text{descontínuo}} \longrightarrow \text{HCT}_{\text{contínuo}}$ que poderia explicar a presença de R_{ex} em diversos resíduos do segmento LCT/HCT (Figura 26) após o ajuste das medidas de relaxação pelo formalismo

Figura 39 – Desestabilização do sítio P2 de ligação a PO_4^{3-} observado nas simulações de DM. (A) O desvio médio das coordenadas de N do segmento 491–497 do sítio P2 na trajetória *DM1+2+3* foi calculado em relação às coordenadas iniciais. (B) Trajetória do deslocamento químico de N do resíduo 497 previsto pelo programa SPARTA+ das estruturas obtidas da simulação *DM1+2+3*. (C) Geometria do sítio P2 cristalográfico onde estão representados os contatos entre o íon SO_3^{2-} e átomos da proteína. O segmento 491–497 está colorido em verde e as distâncias doador–aceptor em ligações de hidrogênio menores que 3,5 Å estão representados por linhas tracejadas.

de Lipari-Szabo.

Valores negativos de Γ também são observados em resíduos do segmento 443–456 (Figura 37) que na estrutura cristalográfica é parcialmente estruturado em α -hélice (446–454). Essa hélice mostrou-se instável na simulação *DM1* e permaneceu estruturada ao longo das simulações *DM2* e *DM3* (Figura 40A). Contudo, os dados espectroscópicos do segmento 446–454 utilizados na previsão de estruturas secundárias (Figura 23; página 63) também apontam para a presença de uma hélice em solução.

O desacordo entre δ_{Exp} e δ_{DM} provocado pela desestruturação da hélice 446–454 na simulação *DM1* pode ser observado nas trajetórias do deslocamento químico de $^{13}\text{C}^\alpha$ dos resíduos 448D e 452F (Figura 40B). Nota-se que desvios mais expressivos entre os deslocamentos químicos previstos das simulações e os valores experimentais são observados com a desestruturação da hélice 446–454 ao longo de quase toda a trajetória *DM1* e nos períodos finais da simulação *DM2*. Esses resultados sugerem uma pequena desestabilização helicoidal pelos parâmetros do campo de força Amber99SB-ILDN (*DM1* e *DM2*), conforme observado por Best e Hummer¹¹³.

Figura 40 – Desestruturação do segmento 446–454 nas simulações de DM. (A) Estruturas secundárias do segmento 446–458 visitadas ao longo das simulações de DM. À direita estão representadas as geometrias cristalográfica (verde) e retirada da trajetória *DM1* (vermelho). (B) Trajetórias dos deslocamentos químicos de $^{13}\text{C}\alpha$ dos resíduos 448D e 452F previstos com SPARTA+ ao longo da simulação *DM1+2+3*.

3.2.2.2 Concordância com as trajetórias simuladas

Nesta seção serão apresentas e discutidas as características estruturais dos segmentos da Cdc25B cujos CS previstos a partir das trajetórias de DM apresentaram maior acordo com os valores experimentais.

O melhor acordo entre δ_{DM} e δ_{Exp} de $^{13}\text{C}\beta$ do resíduo 395H e ^1H do resíduo 396Q está relacionado às conformações do diedral ψ_{395} amostradas na trajetória *DM1+2+3*. Enquanto na estrutura cristalográfica esse ângulo é $\sim 90^\circ$, nas simulações são amostrados valores que flutuam em torno de 127° (Figura 41).

A mudança no diedral ψ_{395} observada na simulação de DM possivelmente relaciona-se à presença de contatos cristalográficos que estabilizam uma geometria pouco populada em solução. Os resíduos 395H e 396Q estão localizados em uma alça que faz contatos com outra molécula na unidade assimétrica do cristal (Figura 42). Assim, o melhor acordo

Figura 41 – Trajetórias dos deslocamentos químicos de $^{13}\text{C}^\beta$ do resíduo 395H e ^1H do resíduo 396Q previstos da simulação *DM1+2+3*. Os deslocamentos químicos experimentais (Exp) e previstos da estrutura cristalográfica estão representados por linhas azul e verde, respectivamente. Estruturas representativas do estado do diedral ψ_{395} na estrutura cristalográfica e populado na trajetória de dinâmica molecular são apresentadas à direita da figura.

entre δ_{DM} e δ_{Exp} demonstram que a geometria dessa alça foi melhor descrita nas simulações de DM, apontando também para a existência de artefatos de cristalização nessa região da proteína.

Figura 42 – Contatos cristalográficos do segmento 395–397 observados no experimento de difração de raios-X. As duas unidades assimétricas da proteína representadas estão coloridas em verde e cinza. Os átomos C, O e N estão representados em verde/cinza, vermelho e azul, respectivamente. As linhas tracejadas representam as distâncias das ligações de hidrogênio.

A flexibilidade conformacional dos resíduos terminais da Cdc25B foi observada experimentalmente (Figura 26; página 69) e confirmada nas simulações de DM (Figura 34). Assim, espera-se que medidas experimentais desses resíduos sejam melhor estimadas a partir de um *ensemble* conformacional em comparação às medidas estimadas de uma única

estrutura estática. Consistente com essa expectativa, os resíduos 375H, 376R, 549S e 550W apresentaram valores negativos de Γ (Figura 37), demonstrando que o conjunto das geometrias visitadas durante as simulações de DM é mais representativo dos estados conformacionais populados em solução quando comparados com a estrutura cristalográfica.

Por exemplo, os δ_{DM} de $^{13}\text{C}^\alpha$, $^{13}\text{C}'$, ^{15}N e ^1H do resíduo 549S estão em maior acordo com os valores experimentais dos que os deslocamentos químicos previstos da estrutura cristalográfica (Tabela 9). Para esse resíduo, maior acordo é decorrente dos diferentes estados conformacionais visitados ao longo das simulações de DM (Figura 43). Esses estados são caracterizados por diferenças nos $\delta_{\text{DM}}^{\text{H}}$, cujas previsões são particularmente sensíveis a torções do diedral ψ do resíduo anterior (ψ_{548})⁶². Na figura 43 é possível verificar que o acordo entre δ_{DM} e δ_{Exp} sugere que além da geometria cristalográfica, o estado conformacional caracterizado por $\psi_{548} \sim 160^\circ$ também é significativamente populado em solução. Vale ainda ressaltar, que em acordo com as simulações, $D_{\text{NH}549} \sim 0$ Hz aponta para sua flexibilidade conformacional em solução e a diferença $|D_{\text{NH}}^{\text{obs}} - D_{\text{NH}}^{\text{calc}}| \sim 5$ Hz indica que a geometria cristalográfica é inconsistente com o RDC medido para esse resíduo (Figura 27; página 70).

De maneira similar, as trajetórias do deslocamento químico de ^1H e ^{15}N do resíduo 550W também sugerem a presença de um outro estado conformacional em solução, caracterizado por $\psi_{549} \sim -10^\circ$, além da geometria observada na estrutura cristalográfica da Cdc25B (Figura 44).

3.2.3 Deslocamentos químicos previstos da região C-terminal desordenada

Devido à ausência de densidade eletrônica do segmento 551–565 no experimento de difração de raios-X do domínio catalítico⁴³, a avaliação de sua dinâmica conformacional em solução foi baseada exclusivamente nas simulações de DM através da comparação entre δ_{DM} e δ_{Exp} . Os resíduos 551–565 são parte do segmento CTD (547–566), cujos parâmetros de ordem S^2 (Figura 26; página 69) evidenciaram sua desordem conformacional em solução. Assim, o segmento 551–565 foi nomeado CTD*. É importante salientar que o programa SPARTA+ não prevê os deslocamentos químicos de resíduos terminais. No caso da

Figura 43 – Trajetória dos deslocamentos químicos de ^1H do resíduo 549S previstos da simulação *DM1+2+3*. Os descolamentos químicos experimentais (Exp) e previstos da estrutura cristalográfica estão representados por linhas azul e verde, respectivamente. Estruturas representativas dos estados do diedral ψ_{548} na estrutura cristalográfica e populados durante a simulação são apresentadas à direita da figura.

construção *C1* do domínio catalítico Cdc25B, os resíduos 366 e 566.

De maneira geral, os deslocamentos químicos do segmento CTD* previstos das simulações *DM1* e *DM2* apresentaram melhor acordo com os valores experimentais em comparação com a simulação *DM3* (Tabela 10), sugerindo que o campo de força Amber99SB-ILDN descreve melhor a distribuição conformacional desta região do que o campo de força CHARMM36. Contudo, acordos ainda mais expressivos com os deslocamentos químicos experimentais foram obtidos das trajetórias *DM1* e *DM2* concatenadas (Tabela 10; *DM1+2*), sugerindo uma melhor amostragem dos estados conformacionais populados pela região C-terminal. Esta observação está em acordo com a semelhança entre os RMSF estruturais das simulações *DM1* e *DM2* e os RMSF_{RCI} dos resíduos do CTD* (Figura 34).

Nota-se que os menores valores de RMSD de CS calculados de médias cumulativas dos δ_{DM} (RMSD^{cm1}) do segmento CTD* nas simulações *DM1* e *DM2* são atingidos por volta de 1 e 1,5 μs , respectivamente (Figura 45). Por outro lado, os menores valores de

Figura 44 – Estados conformacionais populados pelo segmento CTD na trajetória *DM1+2+3*. São apresentadas as trajetória dos deslocamentos químicos de ^1H e ^{15}N do resíduo 550W previstos da simulação *DM1+2+3*. Os valores médios dos deslocamentos químicos previstos da simulação, os deslocamentos químicos experimentais (Exp) e previstos da estrutura cristalográfica estão representados em vermelho, azul e verde, respectivamente.

Tabela 10 – Desvio médio padrão (RMSD) dos deslocamentos químicos do segmento CTD* (551-565) calculados das trajetórias de DM da Cdc25B.

	RMSD ($\delta_{\text{DM}}, \delta_{\text{Exp}}$)					
	^1H	^{15}N	$^{13}\text{C}^\alpha$	$^{13}\text{C}^\beta$	$^{13}\text{C}'$	Soma
<i>DM1</i>	0,27	1,57	0,91	0,94	0,69	4,38
<i>DM2</i>	0,28	1,49	0,67	0,85	0,63	3,92
<i>DM3</i>	0,36	1,94	0,94	1,03	0,86	5,13
<i>DM1+2</i>^a	0,20	1,28	0,55	0,77	0,49	3,29
<i>DM1+2+3</i>^b	0,22	1,40	0,56	0,75	0,57	3,50

^a trajetórias *DM1* e *DM2* concatenadas

^b trajetórias *DM1*, *DM2* e *DM3* concatenadas

RMSD^{cm1} na simulação *DM3* são atingidos em tempos inferiores a 0,25 μs . Além disso, os valores mínimos de RMSD^{cm1} nas simulações *DM1* e *DM2* são inferiores aos da simulação *DM3* (Figura 45). Essas observações indicam que a flexibilidade conformacional do CTD* é melhor representada nas simulações *DM1* e *DM2*, assim como apontado pela semelhança entre a flutuação média (RMSF) dos resíduos calculada das trajetórias e derivadas dos

RCI (Figura 34).

Na figura 45 ainda é possível notar que os valores mínimos de RMSD^{cml} do CTD* ocorrem em tempos mais elevados das simulações quando comparados com os valores mínimos de RMSD^{cml} do núcleo proteico enovelado, obtidos em tempo inferiores a $0,2\ \mu\text{s}$. Esse fato decorre do maior tempo de simulação necessário para amostrar o espaço conformacional de regiões desordenadas da proteína. O tempo de simulação inferior a $0,2\ \mu\text{s}$ necessário para a observação dos valores mínimos de RMSD^{cml} do núcleo proteico da Cdc25B, sugere que simulações de DM mais curtas são capazes de descrever melhor os δ_{Exp} . Esta observação está de acordo com o trabalho de Robustelli *et al.*¹¹¹, onde não houve mudanças expressivas nos valores de RMSD calculados para os deslocamentos químicos previstos de simulações de DM de 100 e 1000 ns de uma proteína com enovelamento pouco flexível. O aumento nos RMSD^{cml} do núcleo após $\sim 0,2\ \mu\text{s}$ também podem ser interpretados como (i) um reflexo das imprecisões do campo de força que, acumuladas ao longo das trajetórias, afetam a descrição da dinâmica conformacional da proteína ou (ii) a amostragem insuficiente de movimentações mais lentas da cadeia principal na escala de μs –ms.

De maneira geral os desvios entre δ_{DM} e δ_{Exp} dos átomos do segmento CTD* são menores do que a precisão do programa SPARTA+ na previsão dos deslocamentos químicos (Figura 46). Contudo, os desvios significativos dos valores experimentais que foram observados para δ_{DM} de $^{13}\text{C}^{\alpha}$ e $^{13}\text{C}^{\beta}$ dos resíduos 556R e 563L, ^1H do resíduo 552G e ^{15}N do resíduo 560C apontam para uma amostragem conformacional insuficiente do CTD*. Desta forma, não é possível interpretar de maneira inequívoca as trajetórias de deslocamento químico dos resíduos do CTD* através da avaliação de seu acordo com os deslocamentos químicos experimentais.

Para exemplificar, nota-se nas trajetórias do deslocamento químico de $^{13}\text{C}^{\alpha}$ e $^{13}\text{C}^{\beta}$ do resíduo 556R a presença de dois estados conformacionais caracterizados pelos diedrais ψ com valores de -10° e 170° (Figura 47). Assim, o desacordo significativo entre δ_{DM} e δ_{Exp} (Figura 46) pode ser um reflexo da amostragem insuficiente da distribuição dos estados visitados ou de imprecisões no campo de força, já que os δ_{Exp} sugerem a predominância de um ângulo $\psi = -10^\circ$ em solução.

Figura 45 – Desvios médios quadráticos (RMSD^{cml}) entre médias cumulativas dos deslocamentos químicos previstas ao longo das trajetórias de DM ($\delta_{\text{DM}}^{\text{cml}}$) e os δ_{Exp} de $^{13}\text{C}^{\alpha}$ e $^{13}\text{C}^{\beta}$ dos resíduos do CTD* (551–566) e do núcleo (387–546).

3.2.4 Contatos entre a região C-terminal e o núcleo proteico

As diferenças entre os CS do grupo NH da cadeia principal medidos para as construções *C1* e *C2* forneceram evidências de contatos existentes entre a região C-terminal e o núcleo proteico (Figura 29; página 72).

Estes contatos também foram observados na simulação *DM1+2+3* (Figura 48). Nota-se na matriz simétrica da distância média entre C^{α} obtida da simulação (Figura 48) a existência de contatos entre HCT/CTD $\times$ segmento vermelho (426–436), HCT $\times$ segmento laranja (443–449) e HCT/CTD $\times$ segmento amarelo (482–487). Contudo, através da metodologia experimental, não foi possível determinar especificamente quais os resíduos que fazem esses contatos. Assim, as perturbações do deslocamento químico foram calculadas utilizando os CS previstos para as estruturas obtidas das simulações *DM1+2+3* mas com os 32 resíduos do C-terminal truncados (Seção 3.1.2).

Nota-se na figura 49 uma semelhança razoável entre os perfis de CSP calculados e experimentais. A magnitude da perturbação calculada no segmento 426–436 (vermelho;

Figura 46 – Desvios entre δ_{DM} da simulação *DM1+2+3* e δ_{Exp} dos resíduos do CTD* res. 551–565 da Cdc25B. A precisão do programa SPARTA+ na previsão dos CS está representada por linhas tracejadas.

Figura 48) é menor do que os valores experimentais (CSP $\sim 0,5$ ppm) pois, possivelmente há um efeito adicional de perturbação conformacional no segmento LCT causadas pela remoção do HCT e CTD que não foram considerados no cálculo. O altos valores de RCI observados nos resíduos do LCT da construção *C2* (dados não apresentados) sugerem que essa perturbação é decorrente da sua desordem conformacional conferida após a remoção dos 32 resíduos do C-terminal da construção *C1*.

Os contatos entre o C-terminal da Cdc25B e o segmento amarelo (482–487; Figura 48) foram avaliados com as trajetórias de δ_{DM} do resíduo 482R, pois seu grupo NH da cadeia principal apresentou o mais elevado valor de CSP calculado do segmento (Figura 49). A abordagem utilizada para o cálculo de CSP considera essencialmente efeitos de corrente na previsão dos CS pela proximidade de anéis aromáticos⁶². Todos os ângulos torcionais são idênticos pois a mesma trajetória de DM foi usada nas previsões (Seção 3.1.2). Dessa forma, a proximidade da cadeia lateral do resíduo 550W do CTD (único resíduo aromático do segmento) e o resíduo 482R na previsão dos deslocamentos químicos de $^{13}C^{\alpha}$, $^{13}C^{\beta}$ e $^{13}C'$ foi avaliada (Figura 50). Nota-se que os CS previstos do resíduo 482R são influenciados principalmente pela torção do diedral χ_1 da cadeia lateral. Por sua vez, o

Figura 47 – Trajetória dos deslocamentos químicos de $^{13}\text{C}^\alpha$ e $^{13}\text{C}^\beta$ do resíduo 556R previstos dinâmica molecular *DM1+2+3*. O valor médio dos deslocamentos químicos previstos e os deslocamentos químicos experimentais (Exp) estão representados em vermelho e azul, respectivamente.

estado conformacional do diedral χ_1 é influenciado pela formação de um contato, ausente na estrutura cristalográfica, tipo cátion- π entre a cadeia lateral de 482R e o anel aromático do 550W (Figura 51). A formação desse contato fica evidente após ~ 5000 ns da simulação, quando a distância 550W–482R é mínima e o diedral χ_1 é estabilizado em $\sim -70^\circ$. Nesses períodos também observa-se um melhor acordo entre os δ_{DM} e δ_{Exp} de $^{13}\text{C}^\alpha$, $^{13}\text{C}^\beta$ e $^{13}\text{C}'$, sugerindo que conformações do CTD onde a interação entre os resíduos 550W e 482R está presente são significativamente populados em solução.

Outra maneira de avaliar a presença desse contato em solução é através da somatória dos desvios entre δ_{DM} e δ_{Exp} de $^{13}\text{C}^\alpha$, $^{13}\text{C}^\beta$ e $^{13}\text{C}'$ de 482R em função da proximidade e orientação do anel aromático de 550W. Observa-se na figura 51 que desvios menores são obtidos quando ocorre o contato cátion- π numa conformação específica, ou seja, quando o ângulo Φ entre as retas normais aos planos do anel indólico e do grupo guanidina é $\sim 0^\circ$ e a distância é mínima. Desvios mais elevados ocorrem quando as normais dos planos estão perpendiculares ($\cos \Phi \sim 0$) ou fazem um ângulo de 180° ($\cos \Phi \sim -1$), apesar da relativa proximidade entre as cadeias laterais (distâncias < 1 nm). Essa análise sugere que

Figura 48 – Matriz simétrica da distância média de C^α obtida da simulação *DM1+2+3*. Foram observados contatos entre a região C-terminal e o núcleo proteico. À esquerda estão coloridos em vermelho, laranja, amarelo e azul, os segmentos onde foram medidos experimentalmente perturbações significativas do deslocamento químico (Figura 29; página 72).

Figura 49 – Perturbações entre os deslocamentos químicos dos grupos amida das construções *C1* e *C2* medidas experimentalmente (vermelho) e calculadas a partir de estruturas obtidas da simulação *DM1+2+3* (preto).

o contato cátion- π entre 482R e 550W ocorre através de uma orientação específica do anel indólico representada no inserto da figura 51.

Figura 50 – Trajetória dos deslocamentos químicos de $^{13}\text{C}^\alpha$, $^{13}\text{C}^\beta$, $^{13}\text{C}'$ do resíduo 482R previstos pelo programa SPARTA+ da simulação *DM1+2+3*. Os deslocamentos químicos experimentais (Exp) e previstos da estrutura cristalográfica estão representados em azul e verde, respectivamente. Abaixo estão representadas as trajetórias do ângulo diedral χ_1 de 482R e a distância entre o centro de massa dos átomos pesados das cadeias laterais de 550W e 482R.

Figura 51 – Contato tipo cátion- π formado entre as cadeias laterais de 482R e 550W durante as simulações de DM. Φ é o ângulo formado entre os planos do anel indólico de 550W e do grupo guanidina de 482R.

4 Conclusões

No capítulo 2 foi apresentado um estudo sobre flexibilidade conformacional do domínio catalítico da Cdc25B em solução através de diferentes medidas de espectroscopia de RMN.

Assim como verificado por Lund e Cierpicki⁴⁹, observou-se uma semelhança global entre as estruturas do núcleo proteico da Cdc25B em solução e no cristal. A estabilidade da hélice C-terminal também foi confirmada, em desacordo com trabalhos anteriores^{46, 48}. Além disso, os sítios P1 e P2 de ligação a PO_4^{3-} presentes na estrutura cristalográfica também foram observados em solução.

Através de diferentes abordagens verificou-se pela primeira vez que os últimos 15 resíduos da região C-terminal são desordenados em solução, consistente com a ausência de densidade eletrônica desta região no experimento de difração de raios-X⁴³. Esse segmento da Cdc25B está envolvido no reconhecimento do seu substrato proteico⁵⁰, tornando-se assim, mais um exemplo da importância de regiões terminais desordenadas de proteínas no reconhecimento molecular⁵¹.

Modelos estruturais da Cdc25B representados pela estrutura cristalográfica (PDB: 1QB0) ou pelo conjunto de conformações amostradas por simulações de DM da construção C1 foram comparados no capítulo 3 frente às medidas espectroscópicas. De forma geral, desvios dos CS previstos em relação aos valores experimentais foram semelhantes para os dois modelos estruturais (DM e PDB), em desacordo com os trabalhos de Markwick *et al.*; Robustelli *et al.*^{110, 111}. Conforme discutido na página 89, esse fato pode ser um reflexo da alta resolução da estrutura cristalográfica (1,9 Å) da Cdc25B. A concordância entre as estruturas cristalográfica e em solução foi também substantiada pelo acordo dos RDCs observados e calculados.

Determinadas regiões da proteína foram significativamente melhor representadas pela estrutura cristalográfica ou pelo conjunto de geometrias populadas nas simulações. As regiões em maior acordo com a estrutura cristalográfica podem ser reflexo de uma amostragem incompleta do espaço conformacional nas simulações, de possíveis deficiências

no campo de força, como no caso da desestruturação da hélice 446–454, ou de erros na composição do sistema, como no caso da desestruturação do sítio P2. Por outro lado, o maior acordo com as trajetórias de DM sugere a presença de artefatos de cristalização ou de estados conformacionais significativamente populados em solução que não são observados na estrutura cristalográfica.

O ajuste das medidas de relaxação pelo formalismo de Lipari-Szabo revelou a presença de movimentações lentas na escala de μs – ms nas regiões proteicas denominadas LCT/HCT. Em acordo com essa observação, durante a simulação de DM ocorreu uma transição que provocou mudanças significativas na cadeia principal deste segmento, conforme discutido na página 87. Assim, a simulação possivelmente amostrou estados conformacionais envolvidos na transição $\text{HCT}_{\text{descontínuo}} \longrightarrow \text{HCT}_{\text{contínuo}}$ e auxiliou a interpretação atomística da troca conformacional (R_{ex}) observada nessa região da proteína. Contudo, essa transição apresenta baixa significância estatística pois foi observada em somente uma das simulações. Assim, a interpretação deve ser feita com cautela.

As medidas de CSP entre as construções *C1* e *C2*, sugeriram a presença de contatos intramoleculares transientes entre a região C-terminal e o núcleo proteico da Cdc25B. Estes contatos foram razoavelmente descritos pelas simulações, onde foi possível avaliá-los com mais detalhes. Por exemplo, verificou-se que geometrias do CTD visitadas nas trajetórias onde ocorre o contato cátion- π entre os resíduos 550W e 482R apresentaram maior acordo com os CS experimentais, sugerindo que este contato é importante em solução.

Dessa maneira, a partir das simulações computacionais foi possível interpretar com detalhe atômico as medidas experimentais de RMN, elucidar a flexibilidade conformacional e os contatos intramoleculares do domínio catalítico da fosfatase Cdc25B em solução.

Referências

- 1 DONALD, V.; G., V. J. *Biochemistry*. 4. ed. Hoboken, NJ, EUA: Wiley, 2011. 1248 p. ISBN 978-0470-57095-1.
- 2 FISCHER, E. Einfluss der Configuration auf die Wirkung der Enzyme. *Berichte der Dtsch. Chem. Gesellschaft*, v. 27, n. 3, p. 2985–2993, 1894. Disponível em: <<http://dx.doi.org/10.1002/cber.18940270364>>.
- 3 LICHTENTHALER, F. W. 100 Years“Schlüssel-Schloss-Prinzip”: What Made Emil Fischer Use this Analogy? *Angew. Chemie Int. Ed. English*, v. 33, n. 2324, p. 2364–2374, 1995. Disponível em: <<http://dx.doi.org/10.1002/anie.199423641>>.
- 4 KOSHLAND, D. E. Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proc. Natl. Acad. Sci.*, v. 44, n. 2, p. 98–104, 1958. Disponível em: <<http://www.pnas.org/content/44/2/98.full.pdf>>.
- 5 BOEHR, D. D.; NUSSINOV, R.; WRIGHT, P. E. The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.*, v. 5, n. 11, p. 789–796, 2009. Disponível em: <<http://dx.doi.org/10.1038/nchembio1209-954d>>.
- 6 HENZLER-WILDMAN, K. a. *et al.* A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature*, v. 450, n. 7171, p. 913–916, 2007. Disponível em: <<http://dx.doi.org/10.1038/nature06407>>.
- 7 BEDEM, H. van den; FRASER, J. S. Integrative, dynamic structural biology at atomic resolution—it’s about time. *Nat. Methods*, v. 12, n. 4, p. 307–318, 2015. Disponível em: <<http://dx.doi.org/10.1038/nmeth.3324>>.
- 8 FRAUENFELDER, H.; SLIGAR, S.; WOLYNES, P. The energy landscapes and motions of proteins. *Science*, v. 254, n. 5038, p. 1598–1603, 1991. Disponível em: <<http://dx.doi.org/10.1126/science.1749933>>.
- 9 KARPLUS, M.; KURIYAN, J. Molecular dynamics and protein function. *Proc. Natl. Acad. Sci.*, v. 102, n. 19, p. 6679–85, 2005. Disponível em: <<http://dx.doi.org/10.1073/pnas.0408930102>>.
- 10 TSAI, C. J. *et al.* Folding funnels, binding funnels, and protein function. *Protein Sci.*, v. 8, n. September 2000, p. 1181–1190, 1999. Disponível em: <<http://dx.doi.org/10.1110/ps.8.6.1181>>.
- 11 DILL, K. A.; CHAN, H. S. From Levinthal to pathways to funnels. *Nat. Struct. Biol.*, v. 4, n. 1, p. 10–19, 1997. Disponível em: <<http://dx.doi.org/10.1038/nsb0197-10>>.
- 12 LEE, R. van der *et al.* Classification of intrinsically disordered regions and proteins. *Chem. Rev.*, v. 114, n. 13, p. 6589–631, 2014. Disponível em: <<http://dx.doi.org/10.1021/cr400525m>>.
- 13 UVERSKY, V. N. Introduction to intrinsically disordered proteins (IDPs). *Chem. Rev.*, v. 114, n. 13, p. 6557–60, 2014. Disponível em: <<http://dx.doi.org/10.1021/cr500288y>>.

- 14 OATES, M. E. *et al.* D²P²: database of disordered protein predictions. *Nucleic Acids Res.*, v. 41, n. Database issue, p. D508–16, 2013. Disponível em: <<http://dx.doi.org/10.1093/nar/gks1226>>.
- 15 DUNKER, A. K. *et al.* Flexible nets: The roles of intrinsic disorder in protein interaction networks. *FEBS J.*, v. 272, n. 20, p. 5129–5148, 2005. Disponível em: <<http://dx.doi.org/10.1111/j.1742-4658.2005.04948.x>>.
- 16 UVERSKY, V. N.; DUNKER, A. K. Understanding protein non-folding. *Biochim. Biophys. Acta*, v. 1804, n. 6, p. 1231–64, 2010. Disponível em: <<http://dx.doi.org/10.1016/j.bbapap.2010.01.017>>.
- 17 WRIGHT, P. E.; DYSON, H. J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.*, v. 16, n. 1, p. 18–29, 2014. Disponível em: <<http://dx.doi.org/10.1038/nrm3920>>.
- 18 GALEA, C. A. *et al.* Regulation of cell division by intrinsically unstructured proteins: intrinsic flexibility, modularity, and signaling conduits. *Biochemistry*, v. 47, n. 29, p. 7598–609, 2008. Disponível em: <<http://dx.doi.org/10.1021/bi8006803>>.
- 19 HARPER, J. W. *et al.* Inhibition of cyclin-dependent kinases by p21. *Mol. Biol. Cell*, v. 6, n. 4, p. 387–400, 1995. Disponível em: <<http://dx.doi.org/10.1091/mbc.6.4.387>>.
- 20 LACY, E. R. *et al.* p27 binds cyclin-CDK complexes through a sequential mechanism involving binding-induced protein folding. *Nat. Struct. Mol. Biol.*, v. 11, n. 4, p. 358–64, 2004. Disponível em: <<http://dx.doi.org/10.1038/nsmb746>>.
- 21 RUSSO, A. A. *et al.* Crystal structure of the p27Kip1 cyclin-dependent-kinase inhibitor bound to the cyclin A-Cdk2 complex. *Nature*, v. 382, n. 6589, p. 325–31, 1996. Disponível em: <<http://dx.doi.org/10.1038/382325a0>>.
- 22 MICHELSEN, K. *et al.* Ordering of the N-terminus of human MDM2 by small molecule inhibitors. *J. Am. Chem. Soc.*, v. 134, n. 41, p. 17059–67, 2012. Disponível em: <<http://dx.doi.org/10.1021/ja305839b>>.
- 23 KRISHNAN, N. *et al.* Targeting the disordered C terminus of PTP1B with an allosteric inhibitor. *Nat. Chem. Biol.*, v. 10, n. 7, p. 558–66, 2014. Disponível em: <<http://dx.doi.org/10.1038/nchembio.1528>>.
- 24 De Simone, A. *et al.* Toward an accurate determination of free energy landscapes in solution states of proteins. *J. Am. Chem. Soc.*, v. 131, n. 11, p. 3810–1, 2009. Disponível em: <<http://dx.doi.org/10.1021/ja8087295>>.
- 25 SALMON, L. *et al.* Nuclear magnetic resonance provides a quantitative description of protein conformational flexibility on physiologically important time scales. *Biochemistry*, v. 50, n. 14, p. 2735–2747, 2011. Disponível em: <<http://dx.doi.org/10.1021/bi200177v>>.
- 26 JENSEN, M. R. *et al.* Exploring free-energy landscapes of intrinsically disordered proteins at atomic resolution using NMR spectroscopy. *Chem. Rev.*, v. 114, n. 13, p. 6632–60, 2014. Disponível em: <<http://dx.doi.org/10.1021/cr400688u>>.
- 27 ROBUSTELLI, P. *et al.* Conformational dynamics of the partially disordered yeast transcription factor GCN4. *J. Chem. Theory Comput.*, v. 9, n. 11, p. 5190–5200, 2013. Disponível em: <<http://dx.doi.org/10.1021/ct400654r>>.

- 28 ALBERTS, B. *et al.* *Molecular Biology of the Cell*. 5. ed. New York, NY, EUA: Garland Science, 2008. 1392 p. ISBN 978-0-8153-4105-5.
- 29 PAVLETICH, N. P. Mechanisms of cyclin-dependent kinase regulation: structures of Cdk, their cyclin activators, and Cip and INK4 inhibitors. *J. Mol. Biol.*, v. 287, n. 5, p. 821–8, 1999. Disponível em: <<http://dx.doi.org/10.1006/jmbi.1999.2640>>.
- 30 OBAYA, A.; SEDIVY, J. Regulation of cyclin-Cdk activity in mammalian cells. *Cell. Mol. Life Sci.*, v. 59, n. 1, p. 126–142, 2002. Disponível em: <<http://dx.doi.org/10.1007/s00018-002-8410-1>>.
- 31 MALUMBRES, M.; BARBACID, M. Mammalian cyclin-dependent kinases. *Trends Biochem. Sci.*, v. 30, n. 11, p. 630–41, 2005. Disponível em: <<http://dx.doi.org/10.1016/j.tibs.2005.09.005>>.
- 32 BOUTROS, R.; DOZIER, C.; DUCOMMUN, B. The when and wheres of CDC25 phosphatases. *Curr. Opin. Cell Biol.*, v. 18, n. 2, p. 185–191, 2006. Disponível em: <<http://dx.doi.org/10.1016/j.ceb.2006.02.003>>.
- 33 KARLSSON-ROSENTHAL, C.; MILLAR, J. B. A. Cdc25: mechanisms of checkpoint inhibition and recovery. *Trends Cell Biol.*, v. 16, n. 6, p. 285–92, 2006. Disponível em: <<http://dx.doi.org/10.1016/j.tcb.2006.04.002>>.
- 34 ARESSY, B.; DUCOMMUN, B. Cell cycle control by the CDC25 phosphatases. *Anticancer. Agents Med. Chem.*, v. 8, n. 8, p. 818–824, 2008. Disponível em: <<http://dx.doi.org/10.2174/187152008786847756>>.
- 35 BOUTROS, R.; LOBJOIS, V.; DUCOMMUN, B. CDC25 phosphatases in cancer cells: key players? Good targets? *Nat. Rev. Cancer*, v. 7, n. 7, p. 495–507, 2007. Disponível em: <<http://dx.doi.org/10.1038/nrc2169>>.
- 36 LAVECCHIA, A.; Di Giovanni, C.; NOVELLINO, E. CDC25 Phosphatase Inhibitors: An Update. *Mini Rev. Med. Chem.*, v. 12, n. 1, p. 62–73, 2012. Disponível em: <<http://dx.doi.org/10.2174/138955712798868940>>.
- 37 JOHNSTON, P. A. *et al.* Cdc25B dual-specificity phosphatase inhibitors identified in a high-throughput screen of the NIH compound library. *Assay Drug Dev. Technol.*, v. 7, n. 3, p. 250–265, 2009. Disponível em: <<http://dx.doi.org/10.1089/adt.2008.186>>.
- 38 PARK, H. *et al.* Structure-based virtual screening approach to identify novel classes of Cdc25B phosphatase inhibitors. *Bioorg. Med. Chem. Lett.*, v. 19, n. 15, p. 4372–4375, 2009. Disponível em: <<http://dx.doi.org/10.1016/j.bmcl.2009.05.078>>.
- 39 LAVECCHIA, A. *et al.* Discovery of new inhibitors of Cdc25B dual specificity phosphatases by structure-based virtual screening. *J. Med. Chem.*, v. 55, n. 9, p. 4142–4158, 2012. Disponível em: <<http://dx.doi.org/10.1021/jm201624h>>.
- 40 LUND, G. *et al.* Inhibition of CDC25B phosphatase through disruption of protein-protein interaction. *ACS Chem. Biol.*, v. 10, n. 2, p. 390–394, 2015. Disponível em: <<http://dx.doi.org/10.1021/cb500883h>>.
- 41 RUDOLPH, J. Cdc25 phosphatases: Structure, specificity, and mechanism. *Biochemistry*, v. 46, n. 12, p. 3595–3604, 2007. Disponível em: <<http://dx.doi.org/10.1021/bi700026j>>.

- 42 FAUMAN, E. B. *et al.* Crystal structure of the catalytic domain of the human cell cycle control phosphatase, Cdc25A. *Cell*, v. 93, n. 4, p. 617–625, 1998. Disponível em: <[http://dx.doi.org/10.1016/S0092-8674\(00\)81190-3](http://dx.doi.org/10.1016/S0092-8674(00)81190-3)>.
- 43 REYNOLDS, R. A. *et al.* Crystal structure of the catalytic subunit of Cdc25B required for G2/M phase transition of the cell cycle. *J. Mol. Biol.*, v. 293, n. 3, p. 559–568, 1999. Disponível em: <<http://dx.doi.org/10.1006/jmbi.1999.3168>>.
- 44 ARANTES, G. M. *Biochem. J.*, v. 399, n. 2, p. 343–350, 2006. Disponível em: <<http://dx.doi.org/10.1042/BJ20060637>>.
- 45 ARANTES, G. M. The catalytic acid in the dephosphorylation of the Cdk2-pTpY/CycA protein complex by Cdc25B phosphatase. *J. Phys. Chem. B*, v. 112, n. 47, p. 15244–15247, 2008. Disponível em: <<http://dx.doi.org/10.1021/jp8070019>>.
- 46 MAMONOV, A. B. *et al.* General library-based monte carlo technique enables equilibrium sampling of semi-atomistic protein models. *J. Phys. Chem. B*, v. 113, n. 31, p. 10891–10904, 2009. Disponível em: <<http://dx.doi.org/10.1021/jp901322v>>.
- 47 BUHRMAN, G. *et al.* Structural mechanism of oxidative regulation of the phosphatase Cdc25B via an intramolecular disulfide bond. *Biochemistry*, v. 44, n. 14, p. 5307–5316, 2005. Disponível em: <<http://dx.doi.org/10.1021/bi047449f>>.
- 48 ARANTES, G. M. Flexibility and inhibitor binding in Cdc25 phosphatases. *Proteins*, v. 78, n. 14, p. 3017–3032, 2010. Disponível em: <<http://dx.doi.org/10.1002/prot.22826>>.
- 49 LUND, G.; CIERPICKI, T. Solution NMR studies reveal no global flexibility in the catalytic domain of CDC25B. *Proteins*, v. 82, n. 11, p. 2889–2895, 2014. Disponível em: <<http://dx.doi.org/10.1002/prot.24581>>.
- 50 WILBORN, M. *et al.* The C-terminal tail of the dual-specificity Cdc25B phosphatase mediates modular substrate recognition. *Biochemistry*, v. 40, n. 47, p. 14200–14206, 2001. Disponível em: <<http://dx.doi.org/10.1021/bi015638h>>.
- 51 UVERSKY, V. N. The most important thing is the tail: Multitudinous functionalities of intrinsically disordered protein termini. *FEBS Lett.*, v. 587, n. 13, p. 1891–1901, 2013. Disponível em: <<http://dx.doi.org/10.1016/j.febslet.2013.04.042>>.
- 52 NUNES-ALVES, A.; ARANTES, G. M. Ligand-Receptor Affinities Computed by an Adapted Linear Interaction Model for Continuum Electrostatics and by Protein Conformational Averaging. *J. Chem. Inf. Model.*, 2014. Disponível em: <<http://dx.doi.org/10.1021/ci500301s>>.
- 53 SANGER, F.; NICKLEN, S.; COULSON, A. R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.*, v. 74, n. 12, p. 5463–7, 1977. Disponível em: <<http://dx.doi.org/10.1073/pnas.74.12.5463>>.
- 54 SATTLER, M. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectrosc.*, v. 34, n. 2, p. 93–158, 1999. Disponível em: <[http://dx.doi.org/10.1016/S0079-6565\(98\)00025-9](http://dx.doi.org/10.1016/S0079-6565(98)00025-9)>.

- 55 DELAGLIO, F. *et al.* NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, v. 6, n. 3, p. 277–293, 1995. Disponível em: <<http://dx.doi.org/10.1007/BF00197809>>.
- 56 VRANKEN, W. F. *et al.* The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins*, v. 59, n. 4, p. 687–696, 2005. Disponível em: <<http://dx.doi.org/10.1002/prot.20449>>.
- 57 ZHANG, H.; NEAL, S.; WISHART, D. S. RefDB: A database of uniformly referenced protein chemical shifts. *J. Biomol. NMR*, v. 25, n. 3, p. 173–195, 2003. Disponível em: <<http://dx.doi.org/10.1023/A:1022836027055>>.
- 58 WISHART, D. S. *Shiftcor Version 1.0*. 2014. <<http://shiftcor.wishartlab.com/>>.
- 59 SHEN, Y.; BAX, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR*, v. 56, n. 3, p. 227–41, 2013. Disponível em: <<http://dx.doi.org/10.1007/s10858-013-9741-y>>.
- 60 BERJANSKII, M. V.; WISHART, D. S. A simple method to predict protein flexibility using secondary chemical shifts. *J. Am. Chem. Soc.*, v. 127, n. 43, p. 14970–14971, 2005. Disponível em: <<http://dx.doi.org/10.1021/ja054842f>>.
- 61 WISHART, D. S.; SYKES, B. D.; RICHARDS, F. M. The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry*, v. 31, n. 6, p. 1647–1651, 1992. Disponível em: <<http://dx.doi.org/10.1021/bi00121a010>>.
- 62 SHEN, Y.; BAX, A. SPARTA+: a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network. *J. Biomol. NMR*, v. 48, n. 1, p. 13–22, 2010. Disponível em: <<http://dx.doi.org/10.1007/s10858-010-9433-9>>.
- 63 RULE, G. S.; HITCHES, T. K. *Fundamentals of Protein NMR Spectroscopy*. Dordrecht, Holanda: Springer Netherlands, 2006. 530 p. ISBN 1-4020-3499-7.
- 64 RÜCKERT, M.; OTTING, G. Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments. *J. Am. Chem. Soc.*, v. 122, n. 32, p. 7793–7797, 2000. Disponível em: <<http://dx.doi.org/10.1021/ja001068h>>.
- 65 OTTIGER, M.; DELAGLIO, F.; BAX, A. Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J. Magn. Reson.*, v. 131, n. 2, p. 373–378, 1998. Disponível em: <<http://dx.doi.org/10.1006/jmre.1998.1361>>.
- 66 SAUPE, A. Recent Results in the Field of Liquid Crystals. *Angew. Chemie Int. Ed. English*, v. 7, n. 2, p. 97–112, 1968. Disponível em: <<http://dx.doi.org/10.1002/anie.196800971>>.
- 67 LOSONCZI, J. A. *et al.* Order matrix analysis of residual dipolar couplings using singular value decomposition. *J. Magn. Reson.*, v. 138, n. 2, p. 334–42, 1999. Disponível em: <<http://dx.doi.org/10.1006/jmre.1999.1754>>.
- 68 HESS, B. *et al.* GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.*, v. 4, n. 3, p. 435–447, 2008. Disponível em: <<http://dx.doi.org/10.1021/ct700301q>>.

- 69 LINDORFF-LARSEN, K. *et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins*, v. 78, n. 8, p. 1950–1958, 2010. Disponível em: <<http://dx.doi.org/10.1002/prot.22711>>.
- 70 MASSI, F. *et al.* NMR $R_{1\rho}$ rotating-frame relaxation with weak radio frequency fields. *J. Am. Chem. Soc.*, v. 126, n. 7, p. 2247–2256, 2004. Disponível em: <<http://dx.doi.org/10.1021/ja038721w>>.
- 71 LIPARI, G.; SZABO, A. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.*, v. 104, n. 17, p. 4546–4559, 1982. Disponível em: <<http://dx.doi.org/10.1021/ja00381a009>>.
- 72 LIPARI, G.; SZABO, A. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *J. Am. Chem. Soc.*, v. 104, n. 1, p. 4559–4570, 1982. Disponível em: <<http://dx.doi.org/10.1021/ja00381a010>>.
- 73 CLORE, G. M. *et al.* Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. *J. Am. Chem. Soc.*, v. 112, n. 12, p. 4989–4991, 1990. Disponível em: <<http://dx.doi.org/10.1021/ja00168a070>>.
- 74 TJANDRA, N. *et al.* Rotational diffusion anisotropy of human ubiquitin from N-15 NMR relaxation. *J. Am. Chem. Soc.*, v. 117, n. 50, p. 12562–12566, 1995. Disponível em: <<http://dx.doi.org/10.1021/ja00155a020>>.
- 75 PALMER, A. G. *R2R1_diffusion 1.11*. 2014. <http://www.palmer.hs.columbia.edu/software/r2r1_diffusion.html>.
- 76 PALMER, A. G. *PDBinertia 1.11*. 2014. <<http://www.palmer.hs.columbia.edu/software/pdbinertia.html>>.
- 77 COLE, R.; LORIA, J. P. FAST-Modelfree: A program for rapid automated analysis of solution NMR spin-relaxation data. *J. Biomol. NMR*, v. 26, n. 3, p. 203–213, 2003. Disponível em: <<http://dx.doi.org/10.1023/A:1023808801134>>.
- 78 PALMER, A. G. *ModelFree 4.2*. 2014. <<http://www.palmer.hs.columbia.edu/software/modelfree.html>>.
- 79 MANDEL, A. M.; AKKE, M.; PALMER, A. G. Backbone dynamics of Escherichia coli ribonuclease HI: correlations with structure and function in an active enzyme. *J. Mol. Biol.*, v. 246, n. 1, p. 144–163, 1995. Disponível em: <<http://dx.doi.org/10.1006/jmbi.1994.0073>>.
- 80 GARCÍA DE LA TORRE, J.; HUERTAS, M. L.; CARRASCO, B. HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J. Magn. Reson.*, v. 147, n. 1, p. 138–46, 2000. Disponível em: <<http://dx.doi.org/10.1006/jmre.2000.2170>>.
- 81 SHARMA, D.; RAJARATHNAM, K. ^{13}C NMR chemical shifts can predict disulfide bond formation. *J. Biomol. NMR*, v. 18, n. 2, p. 165–171, 2000. Disponível em: <<http://dx.doi.org/10.1023/A:1008398416292>>.

- 82 MARTIN, O. A. *et al.* Analysis of $^{13}\text{C}\alpha$ and $^{13}\text{C}\beta$ chemical shifts of cysteine and cystine residues in proteins: a quantum chemical approach. *J. Biomol. NMR*, v. 46, n. 3, p. 217–225, 2010. Disponível em: <<http://dx.doi.org/10.1007/s10858-010-9396-x>>.
- 83 WISHART, D. S.; SYKES, B. D. The ^{13}C chemical-shift index: a simple method for the identification of protein secondary structure using ^{13}C chemical-shift data. *J. Biomol. NMR*, v. 4, n. 2, p. 171–180, 1994. Disponível em: <<http://dx.doi.org/10.1007/BF00175245>>.
- 84 WISHART, D. S. *et al.* ^1H , ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J. Biomol. NMR*, v. 5, n. 1, p. 67–81, 1995. Disponível em: <<http://dx.doi.org/10.1007/BF00227471>>.
- 85 TAMIOLA, K.; ACAR, B.; MULDER, F. A. A. Sequence-specific random coil chemical shifts of intrinsically disordered proteins. *J. Am. Chem. Soc.*, v. 132, n. 51, p. 18000–18003, 2010. Disponível em: <<http://dx.doi.org/10.1021/ja105656t>>.
- 86 CAMILLONI, C. *et al.* Determination of secondary structure populations in disordered states of proteins using nuclear magnetic resonance chemical shifts. *Biochemistry*, v. 51, n. 11, p. 2224–31, 2012. Disponível em: <<http://dx.doi.org/10.1021/bi3001825>>.
- 87 KAY, L. E.; TORCHIA, D. A.; BAX, A. Backbone dynamics of proteins as studied by nitrogen-15 inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry*, v. 28, n. 23, p. 8972–8979, 1989. Disponível em: <<http://dx.doi.org/10.1021/bi00449a003>>.
- 88 CLORE, G. M.; GARRETT, D. S. R-factor, free R, and complete cross-validation for dipolar coupling refinement of nmr structures. *J. Am. Chem. Soc.*, v. 121, n. 39, p. 9008–9012, 1999. Disponível em: <<http://dx.doi.org/10.1021/ja991789k>>.
- 89 SIVASHANMUGAM, A. *et al.* Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli*. *Protein Sci.*, v. 18, n. 5, p. 936–948, 2009. Disponível em: <<http://dx.doi.org/10.1002/pro.102>>.
- 90 HORE, P. J. *Nuclear Magnetic Resonance*. New York, NY, EUA: Oxford University Press, 1995. 87 p. ISBN 0198556829.
- 91 STEHLE, T. *et al.* The apo-structure of the low molecular weight protein-tyrosine phosphatase A (MtpA) from *Mycobacterium tuberculosis* allows for better target-specific drug development. *J. Biol. Chem.*, v. 287, n. 41, p. 34569–34582, 2012. Disponível em: <<http://dx.doi.org/10.1074/jbc.M112.399261>>.
- 92 LIM, J. C. *et al.* A low pKa cysteine at the active site of mouse methionine sulfoxide reductase A. *J. Biol. Chem.*, v. 287, n. 30, p. 25596–25601, 2012. Disponível em: <<http://dx.doi.org/10.1074/jbc.M112.369116>>.
- 93 WEINER, S. J. *et al.* A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.*, v. 106, n. 3, p. 765–784, 1984. Disponível em: <<http://dx.doi.org/10.1021/ja00315a051>>.
- 94 BROOKS, B. R. *et al.* CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.*, v. 4, n. 2, p. 187–217, 1983. Disponível em: <<http://dx.doi.org/10.1002/jcc.540040211>>.

- 95 VERLET, L. Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.*, v. 159, n. 1, p. 98–103, 1967. Disponível em: <<http://dx.doi.org/10.1103/PhysRev.159.98>>.
- 96 ONUFRIEV, A.; BASHFORD, D.; CASE, D. A. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins*, v. 55, n. 2, p. 383–94, 2004. Disponível em: <<http://dx.doi.org/10.1002/prot.20033>>.
- 97 WANG, J.; CIEPLAK, P.; KOLLMAN, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.*, v. 21, n. 12, p. 1049–1074, 2000. Disponível em: <[http://dx.doi.org/10.1002/1096-987X\(200009\)21:12<1049::AID-JCC3>3.0.CO;2-F](http://dx.doi.org/10.1002/1096-987X(200009)21:12<1049::AID-JCC3>3.0.CO;2-F)>.
- 98 HORNAK, V. *et al.* Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins*, v. 65, n. 3, p. 712–725, 2006. Disponível em: <<http://dx.doi.org/10.1002/prot.21123>>.
- 99 MACKERELL, A. D. *et al.* All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*, v. 102, n. 18, p. 3586–616, 1998. Disponível em: <<http://dx.doi.org/10.1021/jp973084f>>.
- 100 BEST, R. B. *et al.* Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ , ψ and side-chain χ_1 and χ_2 dihedral angles. *J. Chem. Theory Comput.*, v. 8, n. 9, p. 3257–3273, 2012. Disponível em: <<http://dx.doi.org/10.1021/ct300400x>>.
- 101 JORGENSEN, W. L. *et al.* Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, v. 79, n. 2, p. 926–935, 1983. Disponível em: <<http://dx.doi.org/10.1063/1.445869>>.
- 102 BERENDSEN, H. J. C. *et al.* Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, v. 81, n. 8, p. 3684–3690, 1984. Disponível em: <<http://dx.doi.org/10.1063/1.448118>>.
- 103 BUSSI, G.; DONADIO, D.; PARRINELLO, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.*, v. 126, n. 1, p. 014101, 2007. Disponível em: <<http://dx.doi.org/10.1063/1.2408420>>.
- 104 HESS, B. *et al.* LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.*, v. 18, n. 12, p. 1463–1472, 1997. Disponível em: <[http://dx.doi.org/10.1002/\(SICI\)1096-987X\(199709\)18:12<1463::AID-JCC4>3.0.CO;2-H](http://dx.doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H)>.
- 105 DARDEN, T.; YORK, D.; PEDERSEN, L. Particle mesh Ewald: An $N \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.*, v. 98, n. 12, p. 10089–10092, 1993. Disponível em: <<http://dx.doi.org/10.1063/1.464397>>.
- 106 JORGENSEN, W. L. *et al.* Development and Testing of the OLPS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.*, v. 118, n. 15, p. 11225–11236, 1996. Disponível em: <<http://dx.doi.org/10.1021/ja9621760>>.
- 107 BEST, R. B.; BUCHETE, N.-V.; HUMMER, G. Are current molecular dynamics force fields too helical? *Biophys. J.*, v. 95, n. 1, p. L07–9, 2008. Disponível em: <<http://dx.doi.org/10.1529/biophysj.108.132696>>.

- 108 BERJANSKII, M. V.; WISHART, D. S. NMR: prediction of protein flexibility. *Nat. Protoc.*, v. 1, n. 2, p. 683–688, 2006. Disponível em: <<http://dx.doi.org/10.1038/nprot.2006.108>>.
- 109 FREDDOLINO, P. L. *et al.* Force field bias in protein folding simulations. *Biophys. J.*, v. 96, n. 9, p. 3772–80, 2009. Disponível em: <<http://dx.doi.org/10.1016/j.bpj.2009.02.033>>.
- 110 MARKWICK, P. R. L. *et al.* Enhanced conformational space sampling improves the prediction of chemical shifts in proteins. *J. Am. Chem. Soc.*, v. 132, n. 4, p. 1220–1, 2010. Disponível em: <<http://dx.doi.org/10.1021/ja9093692>>.
- 111 ROBUSTELLI, P.; STAFFORD, K. A.; PALMER, A. G. Interpreting protein structural dynamics from NMR chemical shifts. *J. Am. Chem. Soc.*, v. 134, n. 14, p. 6365–6374, 2012. Disponível em: <<http://dx.doi.org/10.1021/ja300265w>>.
- 112 KUZMANIC, A.; ZAGROVIC, B. Determination of Ensemble-Average Pairwise Root Mean-Square Deviation from Experimental B-Factors. *Biophys. J.*, v. 98, n. 5, p. 861–871, 2010. Disponível em: <<http://dx.doi.org/10.1016/j.bpj.2009.11.011>>.
- 113 BEST, R. B.; HUMMER, G. Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. *J. Phys. Chem. B*, v. 113, n. 26, p. 9004–15, 2009. Disponível em: <<http://dx.doi.org/10.1021/jp901540t>>.
- 114 BOLTON, W.; PERUTZ, M. F. Three Dimensional Fourier Synthesis of Horse Deoxyhaemoglobin at 2.8 Å Resolution. *Nature*, v. 228, n. 5271, p. 551–552, 1970. Disponível em: <<http://dx.doi.org/10.1038/228551a0>>.
- 115 MONOD, J.; WYMAN, J.; CHANGEUX, J.-P. On the nature of allosteric transitions: A plausible model. *J. Mol. Biol.*, v. 12, n. 1, p. 88–118, 1965. Disponível em: <[http://dx.doi.org/10.1016/S0022-2836\(65\)80285-6](http://dx.doi.org/10.1016/S0022-2836(65)80285-6)>.
- 116 KOSHLAND, D. E.; NÉMETHY, G.; FILMER, D. Comparison of Experimental Binding Data and Theoretical Models in Proteins Containing Subunits *. *Biochemistry*, v. 5, n. 1, p. 365–385, 1966. Disponível em: <<http://dx.doi.org/10.1021/bi00865a047>>.
- 117 YUAN, Y. *et al.* New look at hemoglobin allostery. *Chem. Rev.*, v. 115, n. 4, p. 1702–24, 2015. Disponível em: <<http://dx.doi.org/10.1021/cr500495x>>.

Anexos

ANEXO A – Mecanismos de regulação alostérica na Hemoglobina

A hemoglobina foi uma das primeiras proteínas a ter sua estrutura tridimensional em alta resolução determinada por difração de raios-X (PDB: 2DHB) ¹¹⁴. Esta proteína apresenta uma estrutura quaternária heterotetramérica composta por duas subunidades α e duas subunidades β (ou um dímero de $\alpha\beta$ protômeros). Cada subunidade possui um grupo heme com um átomo de ferro central que está envolvido na ligação de O_2 . Dessa forma, até quatro moléculas de O_2 podem se ligar na Hb. A curva de dissociação do O_2 apresenta um padrão cooperativo, ou seja, a ligação de cada molécula de O_2 aumenta a afinidade da hemoglobina pela ligação de O_2 adicionais. Diferentes mecanismos estruturais baseados no modelo de chave-fechadura ou de ajuste induzido foram propostos para explicar a cooperatividade da complexação entre O_2 e Hb. Estes modelos baseiam-se nas diferenças entre a estrutura quaternária da oxi-Hb e desoxi-Hb. A desoxi-Hb encontra-se numa conformação quaternária **T** (de “tenso”) de baixa afinidade por O_2 e a oxi-Hb num estado **R** (de “relaxado”) de alta afinidade por O_2 . Os estados **t** ou **r** são referentes à estrutura terciária das subunidades α e β ¹.

O modelo proposto por Monod, Wyman e Changeux ¹¹⁵ (modelo MWC) para explicar a cooperatividade alostérica na ligação de O_2 baseia-se na simetria estrutural da Hb. Inicialmente, assume-se um equilíbrio onde todas as subunidades da Hb encontram-se num estado **t** ou **r**, remetendo a um equilíbrio conformacional $R \rightleftharpoons T$ onde ambos estados (**R** e **T**) ligam O_2 com diferentes afinidades. Neste caso, o efeito cooperativo é decorrente da relação entre as afinidades por O_2 e as populações dos estados **R** e **T**. Aqui, o mecanismo de chave-fechadura fica implícito, já que mudanças conformacionais causadas pela complexação não são consideradas.

Já o modelo sequencial proposto por Koshland, Némethy e Filmer ¹¹⁶ (modelo KNF) assume a existência de um mecanismo de ajuste induzido para explicar o efeito cooperativo. Neste modelo, propõe-se que a ligação de uma molécula de O_2 a uma das subunidades induz

mudanças conformacionais nas outras subunidades da Hb. O efeito cooperativo estaria relacionado ao aumento da afinidade por O_2 causado por essas mudanças conformacionais após o primeiro evento de complexação. Portanto a saturação dos quatro sítios da Hb por moléculas de O_2 seria um processo sequencial envolvendo a existência de estados intermediários.

Apesar dos modelos MWC e KNF explicarem de forma simplificada e precisa o comportamento cooperativo da ligação O_2 –Hb, dados mais recentes evidenciaram peculiaridades estruturais e dinâmicas da Hb em diferentes condições, apontando para a necessidade da elaboração modelos mais detalhados para explicar a cooperatividade ¹¹⁷.

ANEXO B – Medidas de relaxação e parâmetros dinâmicos da Cdc25B

Medidas de relaxação e parâmetros dinâmicos dos *spins* de ^{15}N da cadeia principal da construção *C1* da Cdc25B. Os parâmetros dinâmicos (S^2 e R_{ex}) foram ajustados através do formalismo de Lipari-Szabo, utilizando diferentes modelos dinâmicos (Tabela 3).

Res	$R_1(\text{s}^{-1})$	$R_2(\text{s}^{-1})$	NOE	Modelo	S^2	$R_{ex}(\text{s}^{-1})$
371	1.24±0.01	6.13±0.25	0.32±0.02	5	0.22±0.01	
372	1.24±0.01	6.01±0.57	0.29±0.02	5	0.21±0.03	
375	0.96±0.03	8.16±0.27	0.19±0.09	5	0.36±0.01	
377	1.02±0.05	12.26±0.35	0.47±0.11			
389	0.73±0.02	18.06±0.45	0.65±0.04	5	0.84±0.02	
391	0.75±0.01	19.00±0.38	0.86±0.04	1	0.96±0.01	
392	0.78±0.01	14.92±0.28	0.73±0.02	5	0.73±0.01	
393	0.65±0.01	14.70±0.10	0.67±0.03	5	0.70±0.01	
394	0.66±0.01	19.54±0.23	0.72±0.04	2	0.89±0.01	
398	0.71±0.01	21.13±0.63	0.78±0.02	1	0.96±0.01	
399	0.64±0.02	18.74±0.25	0.82±0.06	1	0.89±0.01	
400	0.72±0.01	20.97±0.76	0.79±0.04	1	0.99±0.01	
401	0.68±0.01	17.98±0.14	0.85±0.04	1	0.88±0.01	
405	0.73±0.01	22.66±0.20	0.87±0.03	1	1.00±0.00	
406	0.79±0.01	23.06±0.02	0.89±0.04	1	1.00±0.00	
407	0.78±0.01	20.57±0.15	0.80±0.03	2	0.96±0.01	
408	0.80±0.01	22.16±1.17	0.87±0.03	3	1.00±0.01	0.00±0.74
412	0.72±0.01	17.57±0.65	0.74±0.03	5	0.83±0.03	
413	0.75±0.01	27.59±0.53	0.78±0.03	4	0.96±0.01	7.81±0.58

Tabela 11 – Continuação

Resíduo	$R_1(\text{s}^{-1})$	$R_2(\text{s}^{-1})$	NOE	Modelo	S^2	$R_{ex}(\text{s}^{-1})$
418	0.81±0.01	19.24±1.06	0.59±0.03	2	0.90±0.01	
419	0.79±0.04	21.80±4.24	0.62±0.08	2	0.91±0.03	
422	0.70±0.03	21.16±0.79	0.89±0.08	1	0.98±0.02	
423	0.70±0.01	18.75±0.80	0.83±0.03	1	0.98±0.01	
424	0.69±0.01	19.25±0.65	0.81±0.04	1	0.93±0.02	
425	0.70±0.01	19.64±0.75	0.81±0.04	1	0.96±0.01	
426	0.70±0.01	22.27±1.92	0.82±0.05	1	0.93±0.02	
430	0.74±0.01	20.46±0.02	0.79±0.05	1	0.96±0.00	
431	0.78±0.01	23.50±2.83	0.77±0.04	1	1.00±0.01	
432	0.70±0.01	23.55±1.21	0.90±0.05	3	0.93±0.02	4.25±1.30
437	0.74±0.03	20.25±2.03	0.87±0.07	1	1.00±0.02	
439	0.61±0.02	20.44±0.06	0.83±0.07	3	0.81±0.03	3.60±0.66
440	0.85±0.02	25.89±2.38	0.79±0.05	1	1.00±0.01	
441	0.73±0.01	21.73±0.70	0.92±0.04	1	1.00±0.01	
446	0.71±0.02	20.81±0.73	0.84±0.06	1	0.97±0.02	
448	0.73±0.01	21.79±1.90	0.76±0.04	1	0.95±0.01	
449	0.77±0.01	20.55±0.96	0.70±0.03	2	0.95±0.01	
452	0.73±0.01	11.38±0.09	0.80±0.04			
453	0.69±0.01	18.61±0.93	0.82±0.05	1	0.93±0.02	
455	0.67±0.01	21.26±0.27	0.71±0.04	4	0.83±0.01	4.57±0.36
456	0.70±0.01	13.78±0.09	0.60±0.03	5	0.66±0.01	
469	0.73±0.01	23.87±1.33	0.86±0.04	1	1.00±0.01	
470	0.72±0.01	20.21±0.99	0.79±0.03	1	1.00±0.01	
471	0.69±0.02	21.22±1.70	0.83±0.05	1	0.94±0.02	
483	0.75±0.01	23.06±0.72	0.78±0.04	4	0.95±0.02	4.08±0.78
484	0.71±0.01	22.58±0.54	0.79±0.04	3	0.93±0.02	3.59±0.63
486	0.75±0.01	18.09±0.53	0.83±0.03	1	0.96±0.01	

Tabela 11 – Continuação

Resíduo	$R_1(\text{s}^{-1})$	$R_2(\text{s}^{-1})$	NOE	Modelo	S^2	$R_{ex}(\text{s}^{-1})$
487	0.71±0.01	19.88±0.71	0.90±0.05	1	0.94±0.02	
491	0.76±0.01	20.29±0.57	0.94±0.05	1	1.00±0.01	
492	0.78±0.04	20.51±6.92	0.88±0.08	1	1.00±0.03	
493	0.72±0.01	22.61±0.69	0.80±0.02	4	0.94±0.01	3.14±0.75
494	0.72±0.03	17.27±1.00	0.66±0.08	2	0.88±0.03	
495	0.72±0.01	20.26±0.33	0.85±0.04	1	0.97±0.01	
496	0.70±0.01	19.76±0.96	0.78±0.04	1	0.95±0.01	
497	0.69±0.01	18.84±0.51	0.82±0.04	1	0.92±0.01	
505	0.67±0.03	18.65±2.46	0.99±0.10	1	0.92±0.04	
506	0.70±0.03	16.66±3.49	0.77±0.07	1	0.99±0.02	
507	0.68±0.02	20.93±0.05	0.81±0.05	3	0.91±0.02	1.91±0.48
508	0.75±0.01	21.93±0.04	0.83±0.03	3	0.98±0.01	1.86±0.27
510	0.81±0.02	21.22±1.49	0.88±0.05	1	1.00±0.01	
514	0.73±0.01	18.21±0.79	0.80±0.03	1	0.97±0.01	
515	0.75±0.01	21.95±0.25	0.82±0.03	3	0.97±0.01	2.46±0.34
518	0.70±0.01	18.14±1.14	0.78±0.04	2	0.90±0.01	
521	0.78±0.01	17.64±0.60	0.86±0.03	3	1.00±0.01	0.00±0.39
523	0.70±0.01	21.70±0.55	0.79±0.03	4	0.90±0.01	3.74±0.59
524	0.74±0.01	22.07±0.47	0.81±0.04	1	1.00±0.01	
527	0.68±0.01	21.79±1.40	0.85±0.03	3	0.89±0.01	3.82±1.43
528	0.62±0.01	20.07±0.99	0.81±0.03	3	0.80±0.01	3.94±0.97
531	0.69±0.02	19.49±0.80	0.68±0.04	2	0.90±0.02	
533	0.69±0.01	22.06±0.73	0.74±0.04	4	0.89±0.02	3.69±0.81
534	0.78±0.02	36.64±0.15	0.70±0.04	4	0.95±0.02	17.43±0.35
537	0.76±0.01	23.32±1.49	0.79±0.03	1	0.99±0.01	
538	0.76±0.01	22.39±0.40	0.73±0.03	4	0.96±0.01	2.23±0.40
543	0.82±0.01	20.85±0.65	0.74±0.03	1	1.00±0.01	

Tabela 11 – Continuação

Resíduo	$R_1(\text{s}^{-1})$	$R_2(\text{s}^{-1})$	NOE	Modelo	S^2	$R_{ex}(\text{s}^{-1})$
544	0.81 ± 0.01	22.25 ± 0.81	0.79 ± 0.04	1	1.00 ± 0.01	
545	0.82 ± 0.01	21.92 ± 0.61	0.72 ± 0.02	4	0.92 ± 0.01	2.24 ± 0.65
547	0.82 ± 0.02	21.88 ± 1.52	0.62 ± 0.04	4	0.91 ± 0.01	3.20 ± 1.52
548	0.91 ± 0.02	17.14 ± 0.76	0.67 ± 0.04	2	0.86 ± 0.01	
551	1.19 ± 0.03	11.20 ± 0.47	0.46 ± 0.04	5	0.47 ± 0.02	
553	1.20 ± 0.01	9.37 ± 0.52	0.36 ± 0.02	5	0.38 ± 0.03	
554	1.18 ± 0.02	9.69 ± 0.60	0.33 ± 0.02	5	0.40 ± 0.03	
558	1.20 ± 0.01	7.89 ± 0.14	0.30 ± 0.03	5	0.34 ± 0.01	
563	1.28 ± 0.01	4.47 ± 0.17	0.13 ± 0.02	5	0.15 ± 0.01	

ANEXO C – Comparação entre os deslocamentos químicos previstos e experimentais

Comparação entre os deslocamentos químicos calculados das trajetórias de DM concatenadas ($DM1+2+3$) e os previstos da estrutura cristalográfica com uso do programa SPARTA+. A comparação é feita entre a magnitude da diferença dos deslocamentos químicos experimentais e previstos da estrutura cristalográfica ($|\delta_{\text{PDB}} - \delta_{\text{Exp}}|$) ou da trajetória $DM1+2+3$ ($|\delta_{\text{DM}} - \delta_{\text{Exp}}|$). Valores de $|\delta_{\text{PDB}} - \delta_{\text{Exp}}| - |\delta_{\text{DM}} - \delta_{\text{Exp}}| > 0$ indicam maior com as simulações de DM e valores de $|\delta_{\text{PDB}} - \delta_{\text{Exp}}| - |\delta_{\text{DM}} - \delta_{\text{Exp}}| < 0$ indicam maior acordo com a estrutura cristalográfica. A precisão do programa SPARTA+ associado à previsão dos deslocamentos químicos de cada tipo atômico estão representadas por linhas tracejadas.