**RESSONÂNCIA MAGNÉTICA NUCLEAR NA DETERMINAÇÃO DE
ESTRUTURA DE PROTEÍNAS: APLICAÇÃO À MUTANTE HIS15ALA DE
HPR DE STAPHYLOCOCCUS AUREUS, E AO PEPTÍDEO-C DA
PROINSULINA HUMANA**

CLAUDIA ELISABETH MUNTE

Tese apresentada ao Instituto de
Física de São Carlos, Universidade de
São Paulo, para obtenção do título de
Doutor em Ciências "Física Aplicada"

Orientador: Prof. Dr. Richard Charles Garratt

ou

USP/FSC/SBI

8-2-001403

**São Carlos – São Paulo
2001**

Munte, Claudia Elisabeth

"Ressonância Magnética Nuclear na Determinação de Estrutura de Proteínas: Aplicação à Mutante His15Ala de HPr de Staphylococcus aureus, e ao Peptídeo-C da Proinsulina humana"/
Claudia Elisabeth Munte – São Carlos, 2001

Tese (Doutorado) – Instituto de Física de São Carlos da Universidade de São Paulo, 2001 –
páginas: 142

Área: Física Aplicada

Prof. Dr. Richard Charles Garratt

1. RMN, 2. Staphylococcus aureus, 3. Determinação de Proteínas

I. Título

**IFSC UNIVERSIDADE
DE SÃO PAULO**
Instituto de Física de São Carlos

Caixa Postal 369
13560-970 São Carlos, SP
Av. Trabalhador São-carlense, 400,
13566-590 São Carlos, SP

Fone/Fax 16 273 9777
www.if.sc.usp.br
wladerez@if.sc.usp.br

**MEMBROS DA COMISSÃO JULGADORA DA TESE DE DOUTORADO DE CLÁUDIA
ELISABETH MUNTE APRESENTADA AO INSTITUTO DE FÍSICA DE SÃO CARLOS, DA
UNIVERSIDADE DE SÃO PAULO, EM 04 DE MAIO DE 2001.**

COMISSÃO JULGADORA:

Prof. Dr. Richard Charles Garratt/IFSC-USP

Prof. Dr. Glaucius Oliva/IFSC-USP

Prof. Dr. Otávio Henrique Thiemann/IFSC-USP

Prof. Dr. Fábio C. Almeida/UFRJ

Dr. Luiz Alberto Colnago/EMBRAPA

Agradecimentos

Ao Richard, pelo apoio, paciência e principalmente confiança em aceitar orientar esse trabalho de doutorado, enriquecendo-o com seu imenso conhecimento.

Ao Hans-Robert, por me introduzir nesse fascinante mundo de NMR em proteínas, e por sempre me receber de braços abertos em seu grupo.

À Maria Helena e à Jaciara, pela amizade, pela ajuda e paciência.

Ao Benny, pelo apoio, pelas sugestões, pela compreensão.

Ao Marco, “Kako-véio”, pela amizade querida, pela estabilidade, pelas flores e poesia.

Ao Papi, sempre curioso por meu trabalho; à Christine, pelo orgulho que só uma mãe pode ter...

À Mami, sempre presente em meus pensamentos e coração.

À Didi e sua “Bolinha”, ao Ralf e Peter e meus lindos sobrinhos, ao Kiko, Marcos, Aninha, Gabi.

À Nelma, “pra não dizer que não falei das flores”...

À “Bel-chata”, pelo carinho e papos-relax; ao Beto, pela força (principalmente nas mãos!).

Ao PH, pela amizade e pela super-super-super ajuda (minha dívida em chocolates cresce a cada dia...)

Ao Single, Stéfano, Diet, Carlos, Bianca e todos os demais companheiros de grupo, tanto na USP quanto na Universidade Regensburg, pelo delicioso ambiente de trabalho.

À FAPESP, pelo apoio financeiro, indispensável à realização desse trabalho.

Resumo

A técnica de espectroscopia por Ressonância Magnética Nuclear (NMR) de alta resolução foi utilizada para estudos estruturais em duas biomoléculas: a proteína HPr da bactéria *Staphylococcus aureus*, e o peptídeo-C da insulina humana. Ambas estão relacionadas com a regulação da absorção de glicose pelas células, no primeiro caso em procariontes, e no segundo em organismos superiores.

A proteína HPr (“Histidine-containing protein”) de *Staphylococcus aureus* é uma das componentes centrais do sistema PTS (fosfoenolpiruvato:açúcar-fosfotransferase) de translocação grupal, responsável pelo transporte ativo de açúcar para o interior da célula bacteriana. Nesse processo, a His15 do sítio ativo de HPr é fosforilada pela enzima EI, transferindo, a seguir, o grupo fosfato para a enzima EIIA. A mutação His15→Ala interrompe a transferência do grupo fosfato; apesar disso, a afinidade entre HPr(H15A) e as enzimas EI/EIIA se mostrou semelhante à da nativa. Utilizando técnicas de NMR bidimensionais (COSY, TOCSY, NOESY, HSQC) e tridimensionais (HNCA, HNCO, NOESY-HSQC) foi determinada a estrutura da mutante His15→Ala de HPr de *S. aureus*. Sua estrutura consiste de um sanduíche-aberto, composto de 3 hélices- α paralelamente empacotadas contra uma folha formada por 4 fitas- β anti-paralelas. Esse padrão é encontrado em todas as proteínas HPr já determinadas em diversas espécies, divergindo, porém, significativamente da estrutura previamente publicada para a proteína nativa de *S. aureus* com relação à orientação relativa de alguns elementos de estrutura secundária. Através de uma análise detalhada dos espectros NOESY das proteínas HPr mutante e nativa puderam ser encontradas diferenças conformacionais na região em torno do sítio-ativo. Uma comparação com as outras estruturas de HPr já publicadas revelou uma maior semelhança entre a proteína mutante de *S. aureus* e a proteína no complexo HPr/EI de *E. coli*, fornecendo evidências de que a estrutura encontrada para a mutante represente a conformação assumida pela proteína HPr no momento de sua interação com a enzima EI, assim explicando a sua afinidade inalterada.

O peptídeo-C da proinsulina é importante para a biosíntese da insulina, tendo sido considerado, por muito tempo, biologicamente inerte. Estudos recentes em pacientes diabéticos retomaram a discussão quanto a sua possível atividade reguladora. Utilizando a técnica de espectroscopia de NMR bidimensional (COSY, TOCSY, NOESY), foram realizados estudos estruturais no peptídeo-C da proinsulina humana. Quando dissolvido em 50%/50% água e TFE, o peptídeo-C apresentou 3 regiões centrais (9-12, 15-18, 22-25) com tendência à formação de dobras, uma região N-terminal (2-5) com 2 conformações em voltas- β tipo I e I', e uma região C-terminal (26-31), de conformação extremamente bem definida, incluindo uma volta- β tipo III' (27-30). Em estudos descritos na literatura já foi demonstrada a atividade do pentapeptídeo C-terminal (EGSLQ), na forma de interações quirais com um receptor ainda desconhecido. Estudos anteriores por NMR prevêem a existência de uma estrutura na região C-terminal, a qual foi denominada de “CA-Knuckle”. Nossa proposta é que a estrutura aqui obtida para o pentapeptídeo C-terminal seja justamente o “CA-Knuckle”, representando o sítio-ativo do peptídeo-C da proinsulina humana.

Abstract

High resolution Nuclear Magnetic Resonance spectroscopy has been used for structural studies on two biological macromolecules; the HPr protein from the bacterium *Staphylococcus aureus*, and the C-peptide from human proinsulin. Both are related to the regulation of glucose absorption by cells, the former case in prokaryotes and the latter in higher organisms.

The HPr protein (Histidine Containing Protein) from *S. aureus* is one of the central components of the PTS (Phosphoenolpyruvate;sugar-phosphotransferase) system responsible for the active transport of sugars into the bacterial cell. During this process, His15 of the HPr active site is phosphorylated by enzyme I (EI), and then subsequently transfers this phosphate onto enzyme IIA (EIIA). The His15Ala mutant of HPr, whilst unable to participate in phosphate transfer, nevertheless retains similar affinities for both EI and EIIA. Using two-dimensional (COSY, TOCSY, NOESY, HSQC) and three-dimensional (HNCA, HNCO, NOESY-HSQC) NMR techniques, the structure of the His15Ala mutant of the HPr protein from *S. aureus* was determined. Its structure consists of an open β -sandwich, composed of three α -helices packed against a four-stranded anti-parallel β -sheet. This pattern has been seen in all other HPr proteins from other species so far determined but is markedly different from the previously published native structure from *S. aureus* with respect to the relative orientations of some of the elements of secondary structure. A detailed comparison of the native and mutant structures revealed differences in the conformation of the active site loop. The latter assumes a conformation similar to that seen in the structure of the complex between *E. coli* HPr and EI. This may explain the normal affinities of the mutant protein for EI and EIIA despite the absence of the active site histidine.

The C-peptide of proinsulin is important for the biosynthesis of insulin but has been considered for a long time to be biologically inert. Recent studies in diabetic patients have stimulated a new debate concerning its possible regulatory role. Structural studies of the C-peptide were performed using two-dimensional NMR spectroscopy (COSY, TOCSY and NOESY). In the presence of 50% TFE three central regions of the molecule (residues 9-12, 15-18 and 22-25) showed tendencies to form β -bends. The N-terminal region (residues 2 to 5) was present in the form of either a type I or I' β -turn, whilst the C-terminal region (26-31) presented the most well-defined structure of the whole molecule which included a type III' β -turn. The C-terminal pentapeptide (EGSLQ) has been described in the literature as being responsible for chiral interactions with an as yet uncharacterized receptor. Previous NMR studies have predicted the existence of a well-defined structure at the C-terminus of the C-peptide, known as the CA-knuckle. We propose that the structure described here for the C-terminal pentapeptide is the CA-knuckle and represents the active site of the C-peptide of human proinsulin.

Índice

1. INTRODUÇÃO	1
1.1. NMR na análise de biomoléculas	1
1.2. A Proteína HPr	2
1.3. A Insulina	7
2. ESPECTROSCOPIA POR RMN EM BIOMOLÉCULAS	11
2.1. Noções Básicas	11
2.1.1. Equações de Bloch	14
2.1.2. O FID	19
2.1.3. Largura de Linha	21
2.1.4. Deslocamento Químico	21
2.1.5. Acoplamento Escalar ou Spin-Spin	22
2.1.6. Efeito Overhauser Nuclear	25
2.1.7. NMR Multidimensional	28
2.1.8. Experimentos de NMR utilizados na determinação estrutural	30
2.2. RMN no Estudo de Proteínas e Peptídeos	32
2.2.1. Sistemas de Spin e Assinalamento Seqüencial	32
2.2.2. Distâncias Interprotônicas	41
2.2.3. Ângulos Torcionais	43
2.2.4. Análise Rotamérica	44
2.2.5. Peptídeos e o uso de TFE	47
2.3. Cálculo de Estruturas Tridimensionais	49
2.3.1. Função para a Energia	50
2.3.2. Dinâmica Molecular	52
2.3.3. Minimização de Energia	52
2.3.4. Protocolos de Anelamento Simulado	52
2.4. Análise das Estruturas Obtidas	54

3. PROTEÍNA HPr(H15A) DE <i>Staphylococcus aureus</i>	57
3.1. Materiais e Métodos	57
3.1.1. Preparação das Amostras	57
3.1.2. Espectros de RMN	57
3.1.3. Estratégia de Assinalamento	62
3.1.4. Restrições Experimentais	62
3.1.5. Cálculo de Estruturas	64
3.2. Resultados e Discussão	65
3.2.1. Assinalamento Seqüencial e Estrutura Secundária	65
3.2.2. Estrutura Tridimensional	68
3.2.3. Comparação com a proteína nativa HPr(wt) de <i>S. aureus</i>	77
3.2.4. Discussão e Conclusões	84
 4. PEPTÍDEO C DA PRÓ-INSULINA HUMANA	 88
4.1. Materiais e Métodos	88
4.1.1. Preparação das Amostras	88
4.1.2. Espectros de RMN	89
4.1.3. Estratégia de Assinalamento	92
4.1.4. Restrições Experimentais	93
4.1.5. Cálculo de Estruturas	93
4.2. Resultados e Discussão	94
4.2.1. Assinalamento Seqüencial e Estrutura Secundária	94
4.2.2. Elementos Estruturais	97
4.2.3. Discussão e Conclusões	103
 ABREVIÇÕES E SIGLAS	 I
LISTA DE FIGURAS E TABELAS	IV
BIBLIOGRAFIA	VII
APÊNDICE A: HPr(H15A) de <i>S. aureus</i>: Deslocamentos Químicos ^1H, ^{13}C e ^{15}N e	
Análise Rotamérica	XX
APÊNDICE B: HPr(wt) de <i>S. aureus</i>: Deslocamentos Químicos ^1H	XXVIII
APÊNDICE C: Peptídeo C Insulina Humana: Deslocamentos Químicos ^1H	XXXI

1. Introdução

1.1. NMR NA ANÁLISE DE BIOMOLÉCULAS

Os concomitantes avanços em biologia molecular e em espectroscopia por Ressonância Magnética Nuclear (RMN ou NMR) multidimensional tiveram como reflexo um aumento explosivo na utilização da espectroscopia por NMR a fim de obter informações estruturais e dinâmicas em macromoléculas biológicas de pequeno e médio porte. Técnicas eficientes em biologia molecular na expressão de proteínas enriquecidas com os isótopos ^{13}C e ^{15}N , detectáveis por NMR, resultaram em inúmeros avanços na implementação de técnicas de espectroscopia heteronuclear multidimensional. Como consequência, o limite no tamanho de proteínas passíveis de serem estudadas por NMR aumentou de ~ 10 kDa, usando NMR ^1H homonuclear, para ~ 30 kDa, usando NMR ^{13}C e ^{15}N heteronuclear, ou mesmo para ~ 50 kDa com o uso de NMR heteronuclear combinado a enriquecimento ^2H parcial [1]. Técnicas de transcrição *in vivo* permitiram a aplicação de NMR ^{13}C e ^{15}N heteronuclear a moléculas de RNA e DNA.

A espectroscopia por NMR em solução e a cristalografia de Raios-X são, essencialmente, as únicas técnicas experimentais capazes de fornecer a estrutura tridimensional de uma macromolécula com resolução atômica. O banco de dados de proteínas PDB [2] possui atualmente ~ 15.000 estruturas depositadas, sendo que mais de 97% são resultado de uma dessas técnicas. As demais estruturas provêm, essencialmente, de modelos teóricos. O número de estruturas resolvidas por cristalografia de Raios-X excede em ~ 5 vezes as resolvidas por espectroscopia de NMR, em grande parte devido ao limite no tamanho da proteína tratável por NMR e do fato da técnica ser bastante recente (~ 15 anos). Além disso, os custos de amostras isotopicamente enriquecidas são altos, e a quantidade de tempo a ser investido na resolução de uma estrutura por NMR é grande. Em contrapartida, a cristalografia de Raios-X é limitada, principalmente, pela obtenção de monocristais através de um procedimento em geral não previsível.

Em adição, a espectroscopia por NMR é uma técnica poderosa na investigação de fenômenos dependentes do tempo, como cinética de reação e dinâmica intramolecular em macromoléculas. Historicamente a espectroscopia de NMR em macromoléculas era limitada pela inerente baixa sensibilidade da técnica, e pela complexidade dos espectros. O desenvolvimento de magnetos mais

potentes (atualmente ~22,2 T) e de espectrômetros mais estáveis resultou em um significativo aumento da sensibilidade, e a complexidade dos espectros foi essencialmente convertida em vantagem graças aos avanços teóricos e experimentais na espectroscopia por NMR multidimensional.

Na presente tese de doutorado será utilizada a técnica de espectroscopia por NMR de alta resolução, para estudos estruturais em duas biomoléculas: a proteína HPr da bactéria *Staphylococcus aureus*, e o peptídeo-C da insulina humana. Ambas estão relacionadas com a regulação da absorção de glicose pelas células, no primeiro caso em procariontes, e no segundo em organismos superiores.

1.2. A PROTEÍNA HPR

A proteína HPr (“Histidine-containing protein”) de *Staphylococcus aureus* possui 88 aminoácidos e um peso molecular de 9,496kDa, sendo parte do sistema PTS (fosfoenolpiruvato:açúcar-fosfotransferase) de translocação em grupo [3-5]. O sistema PTS é responsável pelo transporte ativo de açúcar para o interior da célula bacteriana, essencial para suprir suas necessidades energéticas: o açúcar é fosforilado, transformando-se em açúcar-fosfato durante o processo. A membrana celular é impermeável ao açúcar-fosfato, resultando em seu acúmulo no interior da célula. Além disso, o sistema PTS possui propriedades regulatórias.

O sistema PTS consiste de 3 a 5 proteínas: a enzima EI e a proteína HPr são as principais componentes, sendo comuns em todos os organismos. O sistema enzima permease EII é específico para cada organismo e para cada carboidrato, podendo consistir de um único polipeptídeo (IIABC) ou de duas (IIA e IIB/C) ou três (IIA, IIB, IIC) cadeias polipeptídicas [6]. IIA é o domínio citossólico, possuindo uma histidina no sítio ativo; IIB/C são os domínios integrados na membrana, sendo B a região hidrofílica com uma cisteína em seu centro ativo, e C a região hidrofóbica, que possui o sítio de ligação do açúcar. Na nomenclatura antiga o domínio citossólico IIA era chamado de EI^{II} e o domínio integrado à membrana IIB/C de EI^I [4]. A seguir é descrito o processo de translocação de açúcar via sistema PTS (figura 1):

- PEP (fosfoenolpiruvato) é desfosforilado por ação da enzima EI: o grupo fosfato é transferido ao átomo N^{ε2} do anel imidazol da histidina de EI;
- EI interage com HPr, transferindo o grupo fosfato ao átomo N^{δ1} da histidina-15 de HPr;
- HPr transfere o grupo fosfato ao átomo N^{ε2} da histidina de EI^{II} (ou EI^{III});
- EI^{II} transfere o grupo fosfato para a cisteína de EI^IB/C (ou EI^I);
- EI^IB/C interage com o açúcar específico, transferindo-lhe o grupo fosfato e permitindo sua entrada na célula bacteriana.

O efeito da fosforilação do açúcar é duplo: o monossacarídeo simples fosforilado apresenta, em pH fisiológico, duas cargas negativas, tornando-se como íon membrano-impermeável; além disso o potencial de difusão do carboidrato não-fosforilado ao redor da membrana permanece alto, e o açúcar pode penetrar passivamente a célula. Como resultado, o transporte passivo de volta para o meio é reduzido, e o açúcar permanece dentro da bactéria.

Figura 1. Sistema fosfotransferase (PTS), interações e regulação por HPr.

O sistema PTS é também responsável pela regulação da concentração de açúcar intracelular, por intermédio de ativação da repressão catabólica [7]; a proteína HPr participa nesse processo regulatório [8-11]. Em bactérias Gram-positivas, como *Staphylococcus aureus*, HPr possui dois sítios de fosforilação altamente conservados: a histidina-15 do sítio ativo, e a serina-46. A fosforilação da Ser-46 pela HPr-quinase/fosfatase ATP-dependente inibe a interação de HPr com EI, e assim a fosforilação de His15, de forma que a proteína duplamente fosforilada não venha a existir [12, 13]. A desfosforilação de Ser46 ocorre igualmente via HPr-quinase/fosfatase ATP-dependente; sua fosforilação é favorecida por frutose-1,6-bisfosfato e ATP, os quais são indicadores da alta energia na célula e funcionam como ativadores positivos. Dessa forma, na presença de altas concentrações de ATP e frutose-1,6-bisfosfato o sistema PTS é desligado. A

fosforilação de Ser46 permite a formação do complexo entre HPr e proteína de controle catabólico A (CcpA), o qual leva, na presença de frutose-1,6-bifosfato, à repressão catabólica.

Tabela 1. Estruturas de HPr resolvidas até o momento (março de 2001).

Espécie	NMR ou Raios-X	Qualidade	código PDB	Referência
<i>Escherichia coli</i>				
HPr	Raios-X	2.0 Å	1POH	[14]
HPr	NMR	^1H , ^{13}C , ^{15}N alta	1HDN	[15]
P-HPr	NMR	^1H , ^{13}C , ^{15}N alta	1PFH	[16]
HPr Ser46Asp	Raios-X	1.5 Å	1OPD	[17]
Jel42 Fab / HPr	Raios-X	2.5 Å	2JEL	[18]
EIN / HPr	NMR	^1H , ^{13}C , ^{15}N alta	3EZB	[19]
HPr His15Asp	Raios-X	1.8 Å	1CM2	[20]
<i>Bacillus subtilis</i>				
HPr M51Val, Ser83Cys	Raios-X	2.0 Å	2HPR	[21]
HPr Ser46Asp	Raios-X	2.0 Å	1SPH	[22]
HPr Met51Val	NMR	^1H , ^{13}C , ^{15}N alta	2HID	[8]
P-HPr Met51Val	NMR	^1H , ^{13}C , ^{15}N alta	1JEM	[8]
<i>Staphylococcus aureus</i>				
HPr	NMR	^1H baixa	1ZER	[23]
<i>Staphylococcus carnosus</i>				
HPr	NMR	^1H , ^{13}C , ^{15}N alta	1QR5	[24]
<i>Enterococcus faecalis</i>				
HPr	Raios-X	1.6 Å	1PTF	[25]
HPr	NMR	^1H , ^{13}C , ^{15}N alta	1QFR	[26]
<i>Mycoplasma capricolum</i>				
HPr	Raios-X	1.8 Å	1PCH	[27]

A proteína HPr interage, portanto, com EI, EIIA, com uma quinase/fosfatase ATP-dependente e, na forma Ser-P-HPr com CcpA. Essa característica a torna bastante interessante para estudos de interação proteína-proteína, o que explica a grande quantidade de estruturas resolvidas para diversos organismos, tanto de bactérias Gram-positivas quanto de Gram-negativas, por cristalografia de Raios-X e por espectroscopia de NMR (tabela 1). Apesar de uma limitada similaridade entre as seqüências (tabela 2), todas as proteínas HPr possuem o mesmo padrão estrutural: um sanduíche-aberto formado de três hélices- α empacotadas contra uma folha composta por quatro fitas- β anti-paralelas. Na figura 2 são mostradas as estruturas de HPr de

Escherichia coli (Gram-negativa), resolvida por raios-X [14], e de *Enterococcus faecalis* (Gram-positiva), resolvida por NMR multidimensional de alta resolução [26].

Tabela 2. Alinhamento de proteínas HPr de algumas espécies. Em vermelho estão destacados os resíduos conservados em todas as espécies (negrito), e conservados apenas em espécies Gram-positivas (normal).

<i>S. aureus</i>	1	M	E	Q	N	S	Y	V	I	I	D	11	E	T	G	I	H	A	R	P	A	T	21	M	L	V	Q	T	A	S	K	F	D
<i>E. faecalis</i>		M	E	K	K	E	F	H	I	V	A		E	T	G	I	H	A	R	P	A	T		L	L	V	Q	T	A	S	K	F	N
<i>B. subtilis</i>		M	A	Q	K	T	F	K	V	T	A		D	S	G	I	H	A	R	P	A	T		V	L	V	Q	T	A	S	K	Y	D
<i>S. carnosus</i>		M	E	Q	Q	S	Y	T	I	I	D		E	T	G	I	H	A	R	P	A	T		M	L	V	Q	T	A	S	K	F	D
<i>E. coli</i>		M	F	Q	Q	E	V	T	I	T	A		P	N	G	L	H	T	R	P	A	A		Q	F	V	K	E	A	K	G	F	T

<i>S. aureus</i>	31	S	D	I	Q	L	E	Y	N	G	K	41	K	V	N	L	K	S	I	M	G	V	51	M	S	L	G	V	G	K	D	A	E
<i>E. faecalis</i>		S	D	I	N	L	E	Y	K	G	K		S	V	N	L	K	S	I	M	G	V		M	S	L	G	V	G	Q	G	A	D
<i>B. subtilis</i>		A	D	V	N	L	E	Y	N	G	K		T	V	N	L	K	S	I	M	G	V		M	S	L	G	I	A	K	G	A	E
<i>S. carnosus</i>		S	D	I	Q	L	E	Y	N	G	K		K	V	N	L	K	S	I	M	G	V		M	S	L	G	V	G	K	D	A	E
<i>E. coli</i>		S	E	I	T	V	T	S	N	G	K		S	A	S	A	K	S	L	F	K	L		Q	T	L	G	L	T	Q	G	T	V

<i>S. aureus</i>	61	I	T	I	Y	A	D	G	S	D	E	71	S	D	A	I	Q	A	I	S	D	V	81	L	S	K	E	G	L	T	K	
<i>E. faecalis</i>		V	T	I	T	V	D	G	A	D	E		A	E	G	M	A	A	I	V	E	T		L	Q	K	E	G	L	A	E	
<i>B. subtilis</i>		I	T	I	S	A	S	G	A	D	E		N	D	A	L	N	A	L	E	E	T		M	K	S	E	G	L	G	E	
<i>S. carnosus</i>		I	T	I	Y	A	D	G	S	D	E		A	D	A	I	Q	A	I	T	D	V		L	S	K	E	G	L	T	E	
<i>E. coli</i>		V	T	I	S	A	E	G	E	D	E		Q	K	A	V	E	H	L	V	K	L		M	A	E	L	E

(a)

(b)

Figura 2. Padrão estrutural em proteínas HPr. (a) Estrutura por raios-X da proteína HPr de *E. coli* [14] a 2Å de resolução. (b) Estrutura de HPr de *E. faecalis*, resolvida por NMR de alta resolução [26]; apenas a primeira das estruturas depositadas é mostrada.

Muitos detalhes estruturais são bem conservados em todas as estruturas de HPr. Apesar disso, HPrs não são completamente intercambiáveis em sistemas heterólogos, indicando que aspectos específicos de cada espécie são críticos para a interação proteína-proteína que acompanha a fosfotransferência [28]. Estruturas tridimensionais bem definidas são, portanto, necessárias para identificar as diferenças sutis, críticas, entre HPrs de diferentes espécies, as quais resultam em sua especificidade. Na última década foram publicados diversos estudos em proteínas HPr, visando uma melhor compreensão dessas diferenças entre espécies e da forma de interação de HPr com as outras proteínas do sistema PTS [29-41]. Em todas as proteínas HPr os resíduos His15, Arg17 e Pro18 são conservados, sendo parte de um “loop” bem ordenado que constitui o sítio ativo para a transferência do grupo fosfato [31]. Na maioria das estruturas o anel da cadeia lateral da His15 se alinha perpendicularmente ao eixo da primeira hélice (resíduos 18-27). O sítio regulatório, constituído dos resíduos Lys45 e Ser46, é também conservado e inicia a pequena hélice- α regulatória (resíduos 46-53). Detalhes estruturais importantes envolvendo ambos os sítios variam de estrutura para estrutura, levando a diferentes interpretações mecanísticas dos resultados em cada caso. Essas discrepâncias podem ser, em parte, decorrentes de um equilíbrio dinâmico entre estados conformacionais da proteína, cujas populações dependem das condições experimentais (por ex. força iônica, pH, temperatura, estrutura cristalina ou em solução) [24]. A existência de um equilíbrio conformacional entre dois estados de populações aproximadas pôde ser demonstrada para a hélice regulatória em HPr de *E. faecalis* [41]. Experimentos de NMR de alta pressão mostram que, provavelmente, tal equilíbrio exista entre HPr de *S. carnosus* e compreenda o sítio ativo e a hélice regulatória (Görler *et al.*, não publicado).

A proteína HPr de *Staphylococcus aureus* não possui estrutura resolvida por Raios-X, e a estrutura resolvida por NMR (^1H , 500 MHz) é de baixa resolução, o que justifica sua re-determinação em alta resolução (^1H , ^{13}C , ^{15}N). Experimentos realizados com uma mutante His15→Ala (H15A) mostraram uma igual afinidade para com as enzimas I e II se comparada à da proteína nativa (Hengstenberg *et al.*, não publicado). O presente trabalho abordará a determinação da estrutura de HPr(H15A) de *Staphylococcus aureus* e uma comparação detalhada com a proteína nativa, utilizando a técnica de espectroscopia por Ressonância Magnética Nuclear de alta resolução. Semelhanças entre as estruturas de HPr de *S. aureus* e de outras espécies serão discutidas.

1.3. A INSULINA

A insulina é um dos mais importantes hormônios, cuja função principal é a regulação dos níveis de açúcar no soro sanguíneo. É uma proteína pequena, movendo-se rapidamente através do sangue, sendo facilmente capturada por receptores em superfícies celulares para transmitir sua mensagem. Produzida no pâncreas, é liberada ao sangue após as refeições, quando os níveis de açúcar estão altos. A consequência nos tecidos-alvo (por ex. fígado, músculos e células adiposas) é um aumento da entrada de glicose na célula e o seu armazenamento na forma de glicogênio ou gordura.

A insulina é gerada a partir de seu precursor, a preproinsulina, e sintetizada nas células- β das Ilhotas de Langerhans. A preproinsulina é rapidamente convertida em proinsulina, através da remoção do peptídeo sinal, ainda no retículo endoplasmático rugoso. Com a passagem pelo complexo de Golgi se inicia a clivagem proteolítica para gerar insulina, composta de duas cadeias A e B, e peptídeo-C, os quais são armazenados em grânulos. Finalizado o processo de clivagem os grânulos maduros são secretados por exocitose [42-44]. A clivagem da proinsulina é realizada por duas endopeptidases cálcio-dependentes: pró-hormônio-convertase 2 e 3 (PC2 e PC3), as quais clivam especificamente em dois sítios marcados por pares de aminoácidos dibásicos (figura 3) [45]. A endopeptidase tipo 1 (PC3) cliva no sítio Arg-Arg da junção B/C; a segunda endopeptidase (PC2) cliva no sítio Lys-Arg da junção C/A. Por ação da carboxil peptidase H ocorre a perda dos aminoácidos básicos Arg e Lys. Duas pontes de dissulfeto mantêm unidas as cadeias B e A da insulina.

Figura 3. Processamento da proinsulina para insulina. Os círculos na junção B/C (em azul) e C/A (em vermelho) simbolizam, respectivamente, os pares de aminoácidos dibásicos Arg-Arg e Lys-Arg. Os traços (amarelos) unindo as cadeias B e A simbolizam as pontes de dissulfetos. Uma terceira ponte de dissulfeto pode ser vista na cadeia A (também em amarelo).

Desde a descoberta do modo de biosíntese da insulina, em 1967 [42-43], o peptídeo-C foi considerado apenas como auxiliar no correto enovelamento da proinsulina, destituído de outras atividades biológicas. A insulina é filogeneticamente antiga, sendo encontrada em mamíferos, répteis, aves e peixes; sua estrutura primária foi determinada e uma alta similaridade entre as espécies foi demonstrada. Em contraste, peptídeos-C entre as espécies possuem um maior grau de variabilidade, conforme pode ser visto no alinhamento da tabela 2, podendo indicar a ausência de um sítio ativo associado a uma função específica [46]. Estudos recentes em pacientes diabéticos retomaram a discussão quanto a uma possível atividade reguladora do peptídeo-C. Paralelamente, experimentos de desenovelamento e enovelamento realizados sobre a insulina contestam a função inicialmente atribuída ao peptídeo-C.

Tabela 3. Alinhamento entre peptídeos-C de algumas espécies. Em vermelho estão destacados os resíduos conservados em todas as espécies (negrito), e conservados exceto por deleção em alguma(s) espécie(s) (normal). Ratos produzem dois distintos peptídeos-C, aqui representados por *I* e *II*.

	1				6						11				16						21					26					31
Humano	E	A	E	D	L	Q	V	G	Q	V	E	L	G	G	G	P	G	A	G	S	L	Q	P	L	A	L	E	G	S	L	Q
Macaco	E	A	E	D	P	Q	V	G	Q	V	E	L	G	G	G	P	G	A	G	S	L	Q	P	L	A	L	E	G	S	L	Q
Porquinho da Índia	E	L	E	D	P	Q	V	E	Q	T	E	L	G	M	G	L	G	A	G	G	L	Q	P	-	-	L	Q	G	A	L	Q
Rato I	E	V	E	D	P	Q	V	P	Q	L	E	L	G	G	G	P	E	A	G	D	L	Q	T	L	A	L	E	V	A	R	Q
Rato II	E	V	E	D	P	Q	V	A	Q	L	E	L	G	G	G	P	G	A	G	D	L	Q	T	L	A	L	E	V	A	R	Q
Cavalo	E	A	E	D	P	Q	V	G	E	V	E	L	G	G	G	P	G	L	G	G	L	Q	P	L	A	L	A	G	P	Q	Q
Porco	E	A	E	N	P	Q	A	G	A	V	E	L	G	G	G	L	G	-	G	-	L	Q	A	L	A	L	E	G	P	P	Q
Bovino	E	V	E	G	P	Q	V	G	A	L	E	L	A	G	G	P	G	A	G	G	-	-	-	-	-	L	E	G	P	P	Q
Ovelha	E	V	E	G	P	Q	V	G	A	L	E	L	A	G	G	P	G	A	G	G	-	-	-	-	-	L	E	G	P	P	Q
Cachorro	-	-	-	-	-	-	-	-	D	V	E	L	A	G	A	P	G	E	G	G	L	Q	P	L	A	L	E	G	A	L	Q

A diabetes tipo I é decorrente da incapacidade do pâncreas em produzir proinsulina, e portanto insulina e peptídeo-C. Até o momento, a forma de tratamento consiste em administrar, periodicamente, insulina nos pacientes; o peptídeo-C não é suprido. Diversos estudos mostram que a administração de peptídeo-C em animais com diabetes experimental e em pacientes com diabetes tipo I é acompanhada de melhora da função renal [47-49], aumento de utilização de glicose [47,48,50], aumento da circulação sanguínea nos músculos e pele [47,48,51] e melhora da função nervosa autônoma [52]. Estudos *in vitro* confirmaram que o peptídeo-C estimula o transporte de glicose em músculos esqueléticos, e que esse efeito é mediado por caminhos diferentes do que o receptor para a insulina [53]. O peptídeo-C estimula atividades de Na⁺, K⁺-ATPase e eNOs (sintase NO endotelial); em segmentos renais tubulares esse estímulo é compatível com a ativação de um receptor acoplado a uma proteína-G, com subsequente ativação de um caminho de sinalização intracelular Ca²⁺-dependente [51]. Mesmo assim, até o momento nenhum receptor para peptídeo-C pôde ser demonstrado, levando à suposição de que seus efeitos pudessem ser mediados por interações não-quirais com membranas [54]. Essa suposição

foi, porém, rebatida [55-58] por experimentos demonstrando que a transdução de sinal do peptídeo-C segue as regras normais da bioquímica de ligantes/receptores: a ligação é estereoespecífica e ocorre em intervalos de concentração fisiológicos. Há portanto, atualmente na literatura, discussões a respeito da natureza molecular da interação peptídeo-receptor responsável pelos efeitos fisiológicos e farmacológicos do peptídeo-C.

Enquanto que a estrutura tridimensional da insulina já foi determinada para diversos organismos, a estrutura da proinsulina ainda não foi resolvida. Um modelo para a proinsulina foi proposto em 1975 [59], partindo da suposição de que haja semelhança estrutural entre a insulina e as regiões B e A da proinsulina: usando a estrutura tridimensional da insulina foi modelado o peptídeo-C, por previsão empírica a partir de sua seqüência primária. A proinsulina humana possui 86 aminoácidos; a clivagem após Arg31-Arg32 e Lys64-Arg65 e a perda dos aminoácidos básicos resulta em um peptídeo-C com 31 aminoácidos, denominados C1 a C31, e uma insulina com 30 (cadeia B) e 21 (cadeia A) aminoácidos, denominados respectivamente B1 a B30 e A1 a A21. A insulina é estabilizada por 3 pontes de dissulfeto: A6-A11, A7-B7 e A20-B19 (figura 4).

Figura 4. Estrutura da insulina humana em 20% de ácido acético, livre de zinco, resolvida por NMR [60]. Apenas a primeira das 20 estruturas depositadas é mostrada (código pdb: 2HIU). As 3 ligações de dissulfeto estão indicadas. O final da cadeia B (B30) e o início da cadeia A (A1), pontos de junção com o peptídeo-C na proinsulina, estão destacadas em negrito.

A semelhança entre a insulina e as regiões B e A da proinsulina foi confirmada por experimentos de NMR [61] e FTIR [62]; entretanto, divergências tem surgido quanto à conformação do peptídeo-C. Por análise empírica das seqüências de peptídeos-C em diversos organismos, foi prevista uma estrutura helicoidal no mínimo entre as posições C6 e C11, seguida de uma região não-helicoidal entre C13 e C19. É bastante provável que a seqüência Gly-Pro-Gly-X (em algumas espécies Gly-Pro-Pro-Gln) que ocorre na região central (C15 a C18) forme um grampo- β , permitindo que o peptídeo se dobre e preencha o espaço entre a região C-terminal da cadeia B e a região N-terminal da cadeia A da proinsulina. Uma segunda região helicoidal entre posições C21 e C27 foi prevista para a maioria das espécies [59,63,64]. A tendência helicoidal nos primeiros resíduos e não-helicoidal da região central foi reforçada por experimentos recentemente realizados (CD e NMR) no peptídeo-C humano, em 95% de TFE (1,83mM, T=313K, pH=3,6): em água, em membranas artificiais ou na proinsulina o peptídeo-C se mostrou não-estruturado [57].

Dados a importância e o crescente interesse no peptídeo-C, suas propriedades fisiológicas e relevância farmacológica, aliados ao fato de não existir, até o presente momento, nenhum estudo na literatura quanto à sua determinação estrutural completa, essa questão foi abordada na presente tese usando a técnica de espectroscopia de NMR de alta resolução. Espera-se encontrar elementos estruturais estáveis em solução, os quais possam ser relacionados às atividades conferidas ao peptídeo-C.

2. Espectroscopia por NMR em Biomoléculas

2.1. NOÇÕES BÁSICAS

A Ressonância Magnética Nuclear (NMR) foi detectada pela primeira vez em 1946, por Felix Bloch e Edward Purcell [65,66]. Magnetismo nuclear e espectroscopia de NMR são manifestações do momento angular de spin nuclear, uma quantidade vetorial $\mathbf{I}$ de magnitude dada por

$$|\mathbf{I}|^2 = \mathbf{I} \cdot \mathbf{I} = \hbar^2 I(I+1) , \quad (1)$$

onde I é o número quântico do momento magnético nuclear e $\hbar$ a constante de Planck dividida por 2π . Devido às restrições impostas pela mecânica quântica, apenas uma das 3 componentes cartesianas de $\mathbf{I}$ pode ser determinada simultaneamente com $\mathbf{I}^2$; por convenção, o valor da componente z é especificada pela equação

$$I_z = \hbar m , \quad (2)$$

na qual $m=(-I, -I+1, \dots, I-1, I)$ é o número quântico magnético. Núcleos atômicos com momento de spin angular $\mathbf{I}$ não-nulo possuem, também, um momento magnético nuclear $\boldsymbol{\mu}$ definido por

$$\boldsymbol{\mu} = \gamma \mathbf{I}; \quad \mu_z = \gamma I_z = \gamma \hbar m , \quad (3)$$

onde a constante giromagnética γ é característica de cada núcleo. A quantização do momento de spin angular resulta, portanto na quantização de seu momento magnético. Na presença de um campo magnético externo $\mathbf{B}$, os estados de spin do núcleo terão energias dadas por

$$E = -\boldsymbol{\mu} \cdot \mathbf{B} . \quad (4)$$

Em espectrômetros de alto campo de NMR, o campo magnético estático é alinhado com o eixo z do sistema de coordenadas de laboratório, $\mathbf{B} = B_0 \mathbf{k}$. Nessa geometria, (4) se reduz simplesmente a

$$E_m = -\gamma I_z B_0 = -m \hbar \gamma B_0 . \quad (5)$$

Ou seja, na presença de um campo magnético estático as projeções do momento angular dos núcleos no eixo z do referencial de laboratório resultam em $2I+1$ níveis de energia igualmente espaçados entre si, conhecidos como os níveis de Zeeman. Os diferentes níveis de energia são desigualmente populados no equilíbrio, uma vez que orientações de mais baixa energia do vetor de dipolo magnético são mais prováveis. A população relativa de um estado é dada pela distribuição de Boltzmann:

$$\begin{aligned} \frac{N_m}{N} &= \exp\left(\frac{-E_m}{k_B T}\right) \bigg/ \sum_{m=-I}^I \exp\left(\frac{-E_m}{k_B T}\right) \\ &= \exp\left(\frac{m\hbar\gamma B_0}{k_B T}\right) \bigg/ \sum_{m=-I}^I \exp\left(\frac{m\hbar\gamma B_0}{k_B T}\right) , \end{aligned} \quad (6)$$

onde N_m é o número de núcleos no m -ésimo estado e N o número total de spins, T é a temperatura absoluta e k_B a constante de Boltzmann. Próximo à temperatura ambiente, $m\hbar\gamma B_0/k_B T \ll 1$ e a função exponencial em (6) pode ser expandida em primeira ordem pela Série de Taylor:

$$\begin{aligned} \frac{N_m}{N} &\approx \left(1 + \frac{m\hbar\gamma B_0}{k_B T}\right) \bigg/ \sum_{m=-I}^I \left(1 + \frac{m\hbar\gamma B_0}{k_B T}\right) \\ &\approx \left(1 + \frac{m\hbar\gamma B_0}{k_B T}\right) \bigg/ (2I+1) \end{aligned} \quad (7)$$

O momento magnético $\mathbf{M}$ e o momento angular $\mathbf{J}$ de uma amostra macroscópica são dados pela soma vetorial das correspondentes quantidades para núcleos individuais, μ e $\mathbf{I}$. No equilíbrio térmico as componentes transversais x e y de μ e $\mathbf{I}$ para os diferentes núcleos na amostra mediam a zero. O pequeno excesso populacional no nível de mais baixa energia resulta em uma magnetização $\mathbf{M} = M_0 \mathbf{k}$ da amostra, paralela (longitudinal) ao campo magnético estático. O valor de M_0 pode ser obtido combinando-se as equações (3) e (7):

$$\begin{aligned} M_0 &= \sum_{m=-I}^I \mu_z N_m \\ &= \gamma \hbar \sum_{m=-I}^I m N_m \\ &\approx N \gamma \hbar \sum_{m=-I}^I m \left(1 + \frac{m\hbar\gamma B_0}{k_B T}\right) \bigg/ (2I+1) , \\ &\approx \left[N \gamma^2 \hbar^2 B_0 / k_B T (2I+1) \right] \sum_{m=-I}^I m^2 \\ &\approx N \gamma^2 \hbar^2 B_0 I(I+1) / 3 k_B T \end{aligned} \quad (8)$$

Em analogia com outras áreas de espectroscopia, transições entre os níveis de energia podem ser estimuladas com a aplicação de uma radiação eletromagnética. A regra de seleção para transições dipolares magnéticas é dada por

$$\Delta m = \pm 1, \quad (9)$$

e portanto a energia para excitar uma transição entre os estados Zeeman m e $m+1$ será dada por

$$\Delta E = \hbar \gamma B_0. \quad (10)$$

Pela Lei de Planck, $E = \hbar \omega = h \nu$ e portanto a frequência da radiação eletromagnética aplicada deverá ser igual a

$$\begin{aligned} \omega_0 &= \gamma B_0 \\ \nu_0 &= \omega_0 / 2\pi = \gamma B_0 / 2\pi \end{aligned} \quad (11)$$

Em espectroscopia de NMR a frequência de ressonância é tipicamente da ordem de MHz, ou seja, na região de rádio-frequência (rf). Nas aplicações em biomoléculas, os núcleos mais importantes com $I = 1/2$ são ^1H (abundância natural de 99,98%), os isótopos ^{13}C e ^{15}N (abundâncias naturais respectivas de 1,11 e 0,36%) e o ^{31}P (abundância natural de 100%); o mais importante núcleo com $I = 1$ é o deutério (^2H). Nesses núcleos, com a aplicação de um campo magnético externo de 11,74 T a frequência de ressonância será de 500 MHz para o núcleo ^1H ; 125,75 MHz e 50,69 MHz respectivamente para os núcleos ^{13}C e ^{15}N ; 202,41 MHz para o núcleo ^{31}P . O deutério é utilizado para manter a homogeneidade do campo, ressonando em 76,76 MHz. Espectrômetros de NMR são, em geral, denominados pela frequência de ressonância do próton e não pelo campo magnético externo.

Proposto por Bloch em 1950, utilizam-se hoje em dia técnicas de RMN pulsada. Escolhendo adequadamente a largura de banda de um pulso de RF todos os núcleos de um mesmo isótopo (ex. ^1H) são excitados simultaneamente independentemente das suas exatas frequências de ressonância. O sinal observado imediatamente após o pulso, o “*free-induction decay*” ou FID, contém as frequências correspondentes a todos os núcleos da amostra. Este sinal, no domínio do tempo, é processado por computador usando o formalismo da transformada (discreta) de Fourier. Nessa transformação, analisa-se quais são as frequências e suas intensidades contidos no sinal temporal; o resultado é um espectro no domínio das frequências. Nas modernas técnicas de NMR não se utiliza apenas um único pulso de excitação, mas uma seqüências de

pulsos, que servem para manipular os spins com o propósito de obter novas informações. Existe uma enorme variedade de seqüências de pulsos específicas para investigar os mais diferentes fenômenos, como por exemplo o Efeito Overhauser Nuclear.

2.1.1. Equações de Bloch

Bloch formulou um modelo vetorial semi-clássico bastante simples para descrever o comportamento de um conjunto de spins nucleares $I=1/2$ que não interagem entre si, quando sujeitos a um campo magnético estático [67]. A evolução temporal da magnetização (momento magnético) é central no formalismo de Bloch. Na presença de um campo magnético $\mathbf{B}(t)$, o qual pode incluir componentes adicionais ao campo estático, a magnetização $\mathbf{M}(t)$ fica sujeita a um torque igual à derivada temporal do momento angular $\mathbf{J}(t)$:

$$\frac{d\mathbf{J}(t)}{dt} = \mathbf{M}(t) \times \mathbf{B}(t) \quad . \quad (12)$$

Multiplicando ambos os lados da equação por γ teremos

$$\frac{d\mathbf{M}(t)}{dt} = \mathbf{M}(t) \times \gamma \mathbf{B}(t) \quad . \quad (13)$$

Utilizando-se um sistema de referência girando com velocidade angular ω em relação ao referencial de laboratório, inicialmente sobreposto a este, a equação de movimento para $\mathbf{M}(t)$ nos dois sistema de coordenadas é relacionada por

$$\begin{aligned} \left(\frac{d\mathbf{M}(t)}{dt} \right)_{rot} &= \left(\frac{d\mathbf{M}(t)}{dt} \right)_{lab} + \mathbf{M}(t) \times \omega \\ &= \mathbf{M}(t) \times (\gamma \mathbf{B}(t) + \omega) \quad . \\ &= \mathbf{M}(t) \times \gamma \mathbf{B}_{eff}(t) \end{aligned} \quad (14)$$

A equação de movimento para a magnetização no referencial giratório tem a mesma forma que no referencial de laboratório, sendo que o campo magnético é substituído por um campo efetivo $\mathbf{B}_{eff}$. Com a escolha conveniente de $\omega = -\gamma \mathbf{B}(t)$ esse campo efetivo se anula: no referencial giratório a magnetização se torna independente do tempo; no referencial de laboratório $\mathbf{M}(t)$ precessa em torno de $\mathbf{B}(t)$ com uma frequência $\omega = -\gamma B$.

No caso particular em que $\mathbf{B}(t)$ consiste apenas de um campo estático $\mathbf{B}_0$, a frequência de precessão, ou frequência de Larmor, será dada por

$$\omega_0 = -\gamma B_0 , \quad (15)$$

ou seja, a magnetização precessa ao redor do eixo de um campo estático principal (definido como direção z) com a frequência de Larmor. Essa frequência de precessão é idêntica à frequência da radiação eletromagnética necessária para excitar transições entre os níveis Zeeman; essa identidade é a razão pela qual uma descrição clássica (limitada) da espectroscopia de NMR é válida para sistemas de spin isolados.

A precessão do momento magnético ao redor do campo magnético estático constitui um campo magnético dependente do tempo, o qual, segundo a Lei de Faraday, deverá induzir uma força eletromotriz em uma bobina localizada próxima à amostra. A equação (14) sugere que a magnetização possa ser detectada dessa forma. No equilíbrio térmico, porém, a magnetização é colinear com o campo estático, não produzindo sinal na bobina. Uma detecção somente será possível após a perturbação do equilíbrio térmico dos spins, obtida com a aplicação de uma radiação eletromagnética de rádio-frequência (rf).

A componente magnética de um campo de rádio-frequência, linearmente polarizado ao longo da direção x do referencial de laboratório, é dada por

$$\mathbf{B}_{rf}(t) = 2B_1 \cos(\omega_{rf}t + \phi) \mathbf{i} , \quad (16)$$

onde B_1 e ω_{rf} e ϕ são, respectivamente, a amplitude, a frequência angular e a fase do campo aplicado. O campo de rf pode ser decomposto em duas componentes, circularmente polarizadas, girando em direções opostas ao redor do eixo z :

$$\begin{aligned} \mathbf{B}_{rf}(t) = B_1 & \left[\cos(\omega_{rf}t + \phi) \mathbf{i} + \sin(\omega_{rf}t + \phi) \mathbf{j} \right] \\ & + B_1 \left[\cos(\omega_{rf}t + \phi) \mathbf{i} - \sin(\omega_{rf}t + \phi) \mathbf{j} \right] . \end{aligned} \quad (17)$$

Somente a componente que gira no mesmo sentido que o momento magnético interage significativamente com este; a componente oposta, não-ressonante, influencia os spins com um termo de ordem $(B_1/2B_0)^2$, em geral muito pequeno, conhecido como o deslocamento de Bloch-Siegert [68]. Ignorando o termo não-ressonante, o campo de rf pode ser simplesmente escrito como

$$\mathbf{B}_{rf}(t) = B_1 \left[\cos(\omega_{rf}t + \phi) \mathbf{i} + \sin(\omega_{rf}t + \phi) \mathbf{j} \right] . \quad (18)$$

A solução da equação (13) para um campo de rf, em adição a campo estático B_0 , pode ser obtida com a transformação a um sistema de referência giratório que faça com que o campo perturbativo $B_1(t)$ se torne independente do tempo. Esse novo sistema de referência giratório, conhecido como *coordenadas rotativas* ou “*rotating frame*”, gira com velocidade angular ω_f em torno do eixo z; a equação de movimento da magnetização $\mathbf{M}'(t)$ é dada por

$$\frac{d\mathbf{M}'(t)}{dt} = \mathbf{M}'(t) \times \gamma \mathbf{B}' \quad , \quad (19)$$

onde $\mathbf{B}'$ é o campo efetivo no referencial giratório:

$$\begin{aligned} \mathbf{B}' &= \mathbf{B}_1 + \mathbf{B}_0 + \frac{\omega_f}{\gamma} \mathbf{k} \\ &= B_1 \cos \phi \mathbf{i}' + B_1 \sin \phi \mathbf{j}' + \left(B_0 + \frac{\omega_f}{\gamma} \right) \mathbf{k}' \quad . \\ &= B_1 \cos \phi \mathbf{i}' + B_1 \sin \phi \mathbf{j}' + \Delta B_0 \mathbf{k}' \\ &= B_1 \cos \phi \mathbf{i}' + B_1 \sin \phi \mathbf{j}' - \frac{\Omega}{\gamma} \mathbf{k}' \end{aligned} \quad (20)$$

A variável $\Delta B_0 = -(\omega_0 - \omega_f)/\gamma$ é conhecida como o *campo estático reduzido*, sendo muitas vezes substituída pelo “*offset*” $\Omega = -\gamma \Delta B_0$. A magnitude do campo efetivo é dada por

$$B' = \sqrt{(B_1)^2 + (\Delta B_0)^2} \quad ; \quad (21)$$

sua direção é definida por θ e ϕ (figura 5a), sendo θ o ângulo formado com o eixo z:

$$\tan \theta = \frac{B_1}{\Delta B_0} \quad . \quad (22)$$

O campo $B_1(t)$ é, em geral, denominado simplesmente de *campo* B_1 . Com a aplicação do campo B_1 , $\mathbf{M}'(t)$ precessa em torno do campo efetivo $\mathbf{B}'$ com frequência angular

$$\omega' = -\gamma B' \quad . \quad (23)$$

Se o campo de rf for aplicado somente por um tempo τ_p o ângulo de rotação efetiva α é dado por

$$\alpha = \omega' \tau_p = -\gamma B' \tau_p \quad . \quad (24)$$

O tempo τ_p é conhecido como o *tempo do pulso*, e o ângulo α como *ângulo de excitação* (“flip angle”).

Figura 5. Orientações dos campos magnéticos e comportamento da magnetização no referencial giratório “rotating frame”. **(a)** Orientações de ΔB_0 , B_1 e B' : os ângulos θ e ϕ são definidos pelas equações (20) e (22). **(b)** Com a aplicação de um campo de rf aplicado com fase y, o campo efetivo B' localizar-se-á no plano y-z do “rotating frame”; o vetor de magnetização $M'(t)$ precessará ao redor de B' . **(c)** Se o campo de rf for aplicado em ressonância, B' será colinear com o eixo y e a magnetização rotacionará no plano x-z.

Para o caso particular em que $B_1 \gg \Delta B_0$ na equação (22), o que ocorre quando a frequência do campo de rf for próximo à frequência de Larmor ($\omega_f \approx \omega_0$), o campo estático reduzido ΔB_0 se anula e o campo efetivo dado pela equação (20) se resume a

$$\mathbf{B}' = B_1 \cos \phi \mathbf{i}' + B_1 \sin \phi \mathbf{j}' . \quad (25)$$

Nessa condição, denominada *em ressonância*, o campo efetivo será colinear com o campo B_1 no “rotating frame”: a influência do campo estático B_0 terá sido removida. A magnetização $M'(t)$ irá precessar ao redor do eixo definido pelo campo B_1 (figuras 5b,c) com frequência:

$$\omega' = -\gamma B' = -\gamma B^1 \equiv \omega_1 . \quad (26)$$

Após a aplicação do pulso de rf, a magnetização retoma sua rotação em torno de z, com frequência de Larmor ω_0 . Ao contrário, porém, da situação de equilíbrio, a magnetização após a perturbação (pulso) não mais será colinear com o campo estático, podendo ser detectada por uma bobina convenientemente localizada no plano x-y. A rotação da magnetização no plano transversal, e portanto sua detecção, dar-se-ia indefinidamente, não fosse a existência de

mecanismos de relaxação, decorrentes do retorno do sistema ao equilíbrio. Segundo Bloch, dois processos contribuem ao decaimento observado para o sinal de NMR: a relaxação longitudinal e a transversal.

Relaxação longitudinal (ou spin-rede). O retorno da diferença de população Zeeman ao equilíbrio, dado pela distribuição de Boltzman, é conhecido como relaxação longitudinal, ou spin-rede. Esse mecanismo de relaxação reflete a restauração da componente M_z da magnetização, sendo caracterizado, em primeira ordem, pela expressão

$$\frac{dM_z(t)}{dt} = R_1 [M_0 - M_z(t)] , \quad (27)$$

cuja solução é

$$M_z(t) = M_0 - [M_0 - M_z(0)] \exp(-R_1 t) . \quad (28)$$

A constante R_1 é chamada taxa de relaxação spin-rede, em geral expressa na forma do tempo de relaxação spin-rede $T_1 = 1/R_1$.

Relaxação transversal (ou spin-spin). O decaimento da magnetização transversal no plano x-y, conhecido como relaxação transversal, ou spin-spin, é caracterizado, em primeira ordem, por

$$\begin{aligned} \frac{dM_x(t)}{dt} &= -R_2 M_x(t) \\ \frac{dM_y(t)}{dt} &= -R_2 M_y(t) \end{aligned} , \quad (29)$$

com soluções

$$\begin{aligned} M_x(t) &= M_x(0) \exp(-R_2 t) \\ M_y(t) &= M_y(0) \exp(-R_2 t) \end{aligned} . \quad (30)$$

Analogamente a R_1 , R_2 é chamada de taxa de relaxação spin-spin, definindo o tempo de relaxação spin-spin $T_2 = 1/R_2$.

A equação (13) não inclui os mecanismos de relaxação do sistema: as equações (27) e (29) devem ser adicionadas, levando às famosas equações de Bloch para o comportamento da magnetização no referencial de laboratório

$$\begin{aligned}
\frac{dM_x(t)}{dt} &= [\mathbf{M}(t) \times \gamma \mathbf{B}(t)]_x - R_2 M_x(t) \\
\frac{dM_y(t)}{dt} &= [\mathbf{M}(t) \times \gamma \mathbf{B}(t)]_y - R_2 M_y(t) \\
\frac{dM_z(t)}{dt} &= [\mathbf{M}(t) \times \gamma \mathbf{B}(t)]_z - R_1 [M_z(t) - M_0]
\end{aligned} \quad , \quad (31)$$

ou

$$\begin{aligned}
\frac{dM_x(t)}{dt} &= \gamma [M_y(t) B_z - M_z(t) B_y] - R_2 M_x(t) \\
\frac{dM_y(t)}{dt} &= \gamma [M_z(t) B_x - M_x(t) B_z] - R_2 M_y(t) \\
\frac{dM_z(t)}{dt} &= \gamma [M_x(t) B_y - M_y(t) B_x] - R_1 [M_z(t) - M_0]
\end{aligned} \quad . \quad (32)$$

As equações de Bloch no referencial sistema giram de coordenadas são obtidas a partir das equações (19) e (20), com a adição de (27) e (29):

$$\begin{aligned}
\frac{dM_x(t)}{dt} &= \gamma [M_y(t) \Delta B_0 - M_z(t) B_1 \sin \phi] - R_2 M_x(t) \\
\frac{dM_y(t)}{dt} &= \gamma [M_z(t) B_1 \cos \phi - M_x(t) \Delta B_0] - R_2 M_y(t) \\
\frac{dM_z(t)}{dt} &= \gamma [M_x(t) B_1 \sin \phi - M_y(t) B_1 \cos \phi] - R_1 [M_z(t) - M_0]
\end{aligned} \quad . \quad (33)$$

Na ausência do campo de rf, $B_1=0$; essa condição é chamada de *precessão livre*, e as equações de Bloch se tornam simplesmente

$$\begin{aligned}
\frac{dM_x(t)}{dt} &= -\Omega M_y(t) - R_2 M_x(t) \\
\frac{dM_y(t)}{dt} &= \Omega M_x(t) - R_2 M_y(t) \\
\frac{dM_z(t)}{dt} &= R_1 [M_0 - M_z(t)]
\end{aligned} \quad . \quad (34)$$

2.1.2. O FID

Em um experimento simples de NMR aplica-se um pulso de rf em ressonância, de duração τ_p e amplitude B_1 , sobre a magnetização inicialmente em equilíbrio. Supondo que tal pulso seja aplicado ao

longo da direção y do referencial giratório, $\phi=90^\circ$ na equação (33). O efeito do pulso será uma rotação α da magnetização no plano x - z :

$$M(\tau_p) = M_0 \sin \alpha \mathbf{i} + M_0 \cos \alpha \mathbf{k} . \quad (35)$$

Para $\alpha=90^\circ$ a magnetização no plano transversal alcança seu valor máximo, igualando as populações dos dois estados Zeeman: tal pulso é denominado de *pulso de 90°*. A precessão da magnetização, após a aplicação do pulso, gera o sinal detectado pelo espectrômetro de NMR: o “*free-induction decay*”, ou *FID*. Esse FID pode ser descrito em termos de suas duas componentes transversais, resolvendo-se as equações de Bloch para a livre precessão no referencial giratório (34) com a condição dada pela equação (35):

$$\begin{aligned} M_x(t) &= M_0 \sin \alpha \cos(\Omega t) \exp(-R_2 t) \\ M_y(t) &= M_0 \sin \alpha \sin(\Omega t) \exp(-R_2 t) \end{aligned} \quad (36)$$

Combinando-se as duas equações em uma notação complexa, teremos

$$M^+(t) = M_x(t) + iM_y(t) = M_0 \sin \alpha \exp(i\Omega t - R_2 t) . \quad (37)$$

A equação (37) mostra que, após o pulso, as componentes do vetor de magnetização que precessam no plano transversal serão amortecidas pelo fator exponencial $\exp(-R_2 t)$. O sinal, ou FID, é detectado pelo espectrômetro como $s^+(t) = \lambda M^+(t)$, sendo λ uma constante de proporcionalidade experimental. O tempo de relaxação T_2 limita, portanto, o tempo de existência do FID, enquanto que T_1 governa o tempo requerido para que o equilíbrio seja recuperado. Submetendo-se o sinal complexo no domínio temporal $s^+(t)$ a uma transformada de Fourier obtém-se o espectro no domínio da frequência

$$\begin{aligned} S(\omega) &= \int_0^\infty s^+(t) \exp(-i\omega t) dt \\ &= v(\omega) - iu(\omega) \end{aligned} \quad (38)$$

onde

$$\begin{aligned} v(\omega) &= \lambda M_0 \frac{R_2}{R_2^2 + (\Omega - \omega)^2} \\ u(\omega) &= \lambda M_0 \frac{\Omega - \omega}{R_2^2 + (\Omega - \omega)^2} \end{aligned} \quad (39)$$

A parte real $u(\omega)$ do espectro complexo representa um sinal Lorentziano de absorção, sendo normalmente mostrado como espectro de NMR; a parte imaginária $u(\omega)$ fornece um sinal Lorentziano dispersivo.

2.1.3. Largura de Linha

A largura de linha, definida como largura à meia altura da linha de ressonância, afeta tanto a resolução quanto a relação sinal-ruído de um espectro de NMR. A largura de linha *homogênea* do sinal $u(\nu)$ da equação (39) é dado por

$$\Delta \nu = \frac{R_2}{\pi} . \quad (40)$$

A falta de homogeneidade no campo magnético resulta em um alargamento do sinal de NMR. Essa falta de homogeneidade pode ser incorporada em um fator R_{inom} ; a largura de linha *não-homogênea* é representada por

$$\Delta \nu = \frac{R_2 + R_{inom}}{\pi} = \frac{R_2^*}{\pi} . \quad (41)$$

Nos modernos espectrômetros de NMR a contribuição não-homogênea é da ordem de 1 Hz. A contribuição homogênea é aproximadamente proporcional ao tempo de correlação rotacional total da molécula medida, dependendo, portanto, da massa e da forma molecular, além da viscosidade do solvente. Larguras de linhas significativamente maiores que as esperadas são indícios de um aumento do tempo de correlação rotacional, resultante da agregação da amostra. Efeitos de troca química também podem provocar um aumento da largura de linha.

2.1.4. Deslocamento Químico

A condição de ressonância dada em (9) é somente válida para núcleos isolados no espaço livre. Porém, na realidade, núcleos estão situados na matéria como partes de moléculas e são rodeadas por uma nuvem eletrônica. O movimento desses elétrons, induzido pelo campo magnético externo B_0 , gera campos magnéticos secundários em um efeito denominado *blindagem eletrônica*. Esses campos secundários são subtraídos ao campo externo, causando sua diminuição. A distribuição eletrônica em uma molécula é, em geral, anisotrópica; em soluções líquidas

isotrópicas, porém, colisões levam a rápidas reorientações da molécula, e o efeito da blindagem pode ser expresso por um campo magnético efetivo B' :

$$B' = B_0(1 - \sigma) , \quad (42)$$

onde σ é a constante de blindagem média isotrópica eletrônica. Esse novo campo magnético deve ser considerado na condição de ressonância:

$$\omega = \gamma B_0(1 - \sigma) . \quad (43)$$

Uma vez que o deslocamento na frequência de ressonância devido à blindagem depende do meio químico em que o núcleo está imerso, ele é denominado de deslocamento químico. Esse deslocamento químico, δ , é usualmente medido em ppm (partes por milhão) em relação a uma substância padrão (ω_{ref}):

$$\delta(ppm) = \frac{\omega_{rf} - \omega_{ref}}{\omega_0} 10^6 . \quad (44)$$

Para núcleos ^1H esse deslocamento é da ordem de 10 ppm; para ^{13}C e ^{31}P é de algumas centenas de ppm. Mesmo sendo pequeno, o deslocamento químico é o mais importante parâmetro em NMR de alta resolução, permitindo que se observe seletivamente núcleos isolados e grupos de núcleos e que conclusões possam ser tiradas quanto ao meio químico.

2.1.5. Acoplamento Escalar ou Spin-Spin

O acoplamento dipolar entre momentos de dipolo magnéticos é uma quantidade anisotrópica que, em soluções isotrópicas, é mediada a zero (em primeira ordem). Apesar disso, espectros de alta resolução em líquidos revelam uma estrutura fina devido à interações entre núcleos. Ramsey e Purcell sugeriram, em 1952, ser a interação mediada pelos elétrons que formam as ligações químicas entre os núcleos [69]. Essa interação é conhecida por *acoplamento escalar*, ou *spin-spin*, sendo representada pela constante de acoplamento $^nJ_{ab}$, onde n é o número de ligações covalentes separando os núcleos a e b . A intensidade dessa interação spin-spin é dependente do ângulo torcional entre os núcleos que interagem, trazendo, portanto, informações importantes sobre a conformação local da molécula.

O acoplamento escalar modifica os níveis de energia do sistema, e portanto os espectros de NMR. O exemplo mais simples é dado por dois núcleos de spin- $1/2$, designados por I e S , com

números quânticos magnéticos m_I e m_S e frequências de ressonância ω_I e ω_S . Sendo J_{IS} a constante de acoplamento escalar entre os núcleos, a energia para cada um de seus estados é obtida através da equação

$$E(m_I, m_S) = -m_I \omega_I - m_S \omega_S + 2\pi m_I m_S J_{IS} , \quad 2\pi J_{IS} / |\omega_I - \omega_S| \ll 1 . \quad (45)$$

A condição para J_{IS} em (45) é denominada *regime de acoplamento fraco*. Os níveis de energia e seu números quânticos magnéticos totais serão

$$\begin{aligned} E_{\alpha\alpha} &= -\frac{1}{2}\hbar\omega_I - \frac{1}{2}\hbar\omega_S + \frac{1}{2}\hbar\pi J_{IS} & m_{\alpha\alpha} &= +\frac{1}{2} + \frac{1}{2} = +1 \\ E_{\alpha\beta} &= -\frac{1}{2}\hbar\omega_I + \frac{1}{2}\hbar\omega_S - \frac{1}{2}\hbar\pi J_{IS} & m_{\alpha\beta} &= +\frac{1}{2} - \frac{1}{2} = 0 \\ E_{\beta\alpha} &= +\frac{1}{2}\hbar\omega_I - \frac{1}{2}\hbar\omega_S - \frac{1}{2}\hbar\pi J_{IS} & m_{\beta\alpha} &= -\frac{1}{2} + \frac{1}{2} = 0 \\ E_{\beta\beta} &= +\frac{1}{2}\hbar\omega_I + \frac{1}{2}\hbar\omega_S + \frac{1}{2}\hbar\pi J_{IS} & m_{\beta\beta} &= -\frac{1}{2} - \frac{1}{2} = -1 \end{aligned} . \quad (46)$$

As transições observáveis obedecem a regra de seleção $\Delta m = \pm 1$, ocorrendo entre os estados $\alpha\alpha$ e $\alpha\beta$, $\beta\alpha$ e $\beta\beta$ (mudança de estado do spin I) e entre os estados $\alpha\alpha$ e $\beta\alpha$, $\alpha\beta$ e $\beta\beta$ (mudança de estado do spin S).

Na ausência de acoplamento escalar, $J_{IS}=0$ e as frequências de transição serão dadas por

$$\begin{aligned} \omega_{\alpha\alpha-\alpha\beta} &= \omega_{\beta\alpha-\beta\beta} = \omega_S \\ \omega_{\alpha\alpha-\beta\alpha} &= \omega_{\alpha\beta-\beta\beta} = \omega_I \end{aligned} , \quad (47)$$

resultando em um espectro de NMR com apenas duas linhas de ressonância, uma em ω_S e outra em ω_I . Introduzindo-se o acoplamento escalar J_{IS} as frequências de transição serão modificadas para

$$\begin{aligned} \omega_{\alpha\alpha-\alpha\beta} &= \omega_S - \pi J_{IS} \\ \omega_{\beta\alpha-\beta\beta} &= \omega_S + \pi J_{IS} \\ \omega_{\alpha\alpha-\beta\alpha} &= \omega_I - \pi J_{IS} \\ \omega_{\alpha\beta-\beta\beta} &= \omega_I + \pi J_{IS} \end{aligned} , \quad (48)$$

e o resultado será um espectro de NMR com quatro linhas de ressonância: duas centradas em torno da frequência ω_S do spin S , separadas por $2\pi J_{IS}$, e duas centradas em torno da frequência ω_I

do spin I , com a mesma separação $2\pi J_{IS}$. Um sistema de dois spins fracamente acoplados é usualmente denominado de sistema AX , enquanto que no caso de acoplamento forte o sistema é denominado de AB . Os diagramas de níveis de energia para um sistema AX , sem e com acoplamento escalar, podem ser vistos na figura 6.

Figura 6. Níveis de energia e espectros de NMR esperados para um sistema de spin AX , na ausência (a) e na presença (b) de interação de acoplamento escalar entre os spins. As transições permitidas estão indicadas.

O modelo de Bloch prevê que, em um sistema de dois spins, dois sinais de NMR serão obtidos. Como mostrado por (48), se os spins compartilham de um acoplamento escalar não-nulo, quatro sinais de ressonância aparecerão. O modelo básico de Bloch pode ser estendido a fim de descrever a evolução de um sistema acoplado escalarmente, bastando para isso que cada linha de ressonância resultante da interação de acoplamento escalar seja tratada como um vetor de magnetização independente no referencial giratório. Nesse novo modelo, porém, alguns problemas poderão surgir:

- os efeitos resultantes do intenso acoplamento que ocorre quando $J_{IS} \approx |\omega_I - \omega_S|/2\pi$ não poderão ser descritos;
- os resultados da aplicação de pulsos não-seletivos à magnetização transversal em um sistema homonuclear acoplado não poderão ser descritos sem certas suposições adicionais;
- a transferência da magnetização via transições proibidas, quando o sistema não se encontra no equilíbrio térmico, não poderá ser explicada.

A descrição desses fenômenos requer a introdução de um formalismo mais rigoroso, quântico, com uso de operadores de densidade. A teoria completa de NMR pode ser encontrada nas monografias de A. Abragam (1961), C.P. Slichter (1978), R. R. Ernst *et al.* (1987) e M. Goldman (1988).

2.1.6. Efeito Overhauser Nuclear

O acoplamento dipolar reflete a interação entre os campos magnéticos de dipolos, sendo, portanto, dependente da orientação e da distância entre eles. Essa interação magnética direta é muito mais intensa do que o acoplamento escalar, podendo provocar variações nas frequências de ressonância de vários kHz. Em soluções isotrópicas, porém, o rápido movimento molecular promedia a interação dipolar sobre todas as possíveis orientações da molécula, resultando em uma contribuição nula. Essa rápida reorientação das interações dipolares resulta em uma fonte de campos flutuantes capazes de estimular relaxações longitudinais. Sendo o acoplamento dipolar uma conexão direta entre os núcleos envolvidos, relaxações com transferência de magnetização via transições proibidas podem ocorrer. Tal transferência é denominada *relaxação cruzada*, originando o Efeito Overhauser Nuclear (NOE). O NOE é, portanto, um aspecto da relaxação nuclear: ele provém da mudança de intensidade de uma determinada ressonância quando transições de uma outra são perturbadas de alguma forma. Para uma análise completa desse efeito, mecanismos de relaxação em sistemas multi-spin que interagem entre si devem ser estudados. Como já dito anteriormente, uma descrição detalhada do fenômeno não pode ser obtida pelo modelo de Bloch, requerendo a introdução de um formalismo mais rigoroso. No contexto dessa tese, o Efeito Overhauser Nuclear será descrito de uma forma conceitual.

Supondo um sistema de spin- $\frac{1}{2}$ homonucleares (mesmo γ), composto de $2N$ spins I e $2N$ spins S , não acoplados escalarmente, as energias e os números quânticos magnéticos totais dos 4 estados possíveis serão dadas, conforme a equação (46), por

$$\begin{aligned}
 E_{\alpha\alpha} &= -\frac{1}{2}\hbar\omega_I - \frac{1}{2}\hbar\omega_S & m_{\alpha\alpha} &= +\frac{1}{2} + \frac{1}{2} = +1 \\
 E_{\alpha\beta} &= -\frac{1}{2}\hbar\omega_I + \frac{1}{2}\hbar\omega_S & m_{\alpha\beta} &= +\frac{1}{2} - \frac{1}{2} = 0 \\
 E_{\beta\alpha} &= +\frac{1}{2}\hbar\omega_I - \frac{1}{2}\hbar\omega_S & m_{\beta\alpha} &= -\frac{1}{2} + \frac{1}{2} = 0 \\
 E_{\beta\beta} &= +\frac{1}{2}\hbar\omega_I + \frac{1}{2}\hbar\omega_S & m_{\beta\beta} &= -\frac{1}{2} - \frac{1}{2} = -1
 \end{aligned} \tag{49}$$

As transições entre os estados podem ser divididas em quatro categorias:

- transições S ($\alpha\alpha \rightarrow \alpha\beta$ e $\beta\alpha \rightarrow \beta\beta$);
- transições I ($\alpha\alpha \rightarrow \beta\alpha$ e $\alpha\beta \rightarrow \beta\beta$);
- transições $\Delta m=0$ ($\beta\alpha \rightarrow \alpha\beta$);
- transições $\Delta m=2$ ($\alpha\alpha \rightarrow \beta\beta$).

Enquanto que as duas primeiras transições são permitidas pelas regras de seleção ($\Delta m = \pm 1$), as duas últimas não são observadas sob condições normais. Caso ocorra uma perturbação no equilíbrio populacional do sistema, este deverá reagir procurando recuperá-lo de qualquer maneira, de forma que, em princípio, as quatro transições podem ser utilizadas como caminhos para a relaxação (figura 7a). As taxas em que os quatro mecanismos podem ocorrer são dadas por W_1^S , W_1^I , W_0 , e W_2 , respectivamente para transições S , I , $\Delta m=0$ e $\Delta m=2$.

Figura 7. Relaxação em um sistema de spin- $\frac{1}{2}$ homonucleares, composto de $2N$ spins I e $2N$ spins S , não acoplados escalarmente. **(a)** caminhos para a relaxação após perturbação do sistema; **(b)** população dos níveis de energia no equilíbrio; **(c)** população dos níveis de energia após saturação das transições S .

As populações nos níveis de energia obedecem à distribuição de Boltzmann. Sendo os spins homonucleares, a diferença de energia entre os níveis $\alpha\beta$ e $\beta\alpha$ é desprezível, de forma que, no equilíbrio, os dois níveis podem ser considerados igualmente populados. O nível $\alpha\alpha$, por ser de mais baixa energia, possui um excesso de população δ , enquanto que o nível $\beta\beta$ possui uma deficiência populacional idêntica (figura 7b). Dessa forma, as diferenças populacionais entre os estados, no equilíbrio, serão:

- δ (para as transições S)
- δ (para as transições I)
- *nula* (para as transições $\Delta m=0$)
- 2δ (para as transições $\Delta m=2$)

Em um experimento NOE, promove-se a *saturação* (saturar = igualar as populações) de uma das transições permitidas, observando-se o efeito sobre a outra transição permitida. Saturando-se, por exemplo, as transições *I* (figura 7c), as novas diferenças populacionais serão:

- *nula* (para as transições *I*)
- δ (para as transições *S*)
- δ (para as transições $\Delta m=0$)
- δ (para as transições $\Delta m=2$)

O sistema não se encontra mais em equilíbrio e deverá reagir na tentativa de recuperá-lo. O caminho de relaxação via W_1^I será irrelevante pois a transição *I* se encontra sob saturação. A diferença populacional para as transições *S* foi mantida, o que descarta W_1^S . A única forma de recuperar o equilíbrio será através de W_0 ou W_2 . Duas situações devem ser consideradas:

O caminho de relaxação W_0 é dominante. Nesse caso, W_0 tentará promover uma transferência de $\delta/2$ spins do nível $\beta\alpha$ para o nível $\alpha\beta$, a fim de recuperar a diferença de população *nula* para as transições $\Delta m=0$. Como consequência as diferenças populacionais para as transições *S* tenderão a cair para $\delta/2$, provocando uma diminuição da intensidade do sinal obtido para o spin *S*. Essa tendência é contrabalanceada por W_1^S , uma vez que, de seu ponto de vista, as transições *S* já se encontravam em equilíbrio. O resultado final sobre a intensidade do sinal do spin *S* dependerá do balanço entre W_0 e W_1^S . Dessa forma, se W_0 for o caminho de relaxação dominante, a saturação de *I* provocará uma diminuição do sinal de *S*, resultando em um *NOE negativo em S devido a I*.

O caminho de relaxação W_2 é dominante. Devido a W_2 ocorrerá uma tentativa de transferência de $\delta/2$ spins do nível $\beta\beta$ para o nível $\alpha\alpha$, a fim de recuperar a diferença de população 2δ para as transições $\Delta m=2$. As diferenças populacionais para as transições *S* tenderão a subir para $3\delta/2$, aumentando a intensidade do sinal obtido para o spin *S*. Novamente, W_1^S tentará manter o equilíbrio das transições *S* já existente, de forma que o resultado final dependerá do balanço entre W_2 e W_1^S . Assim, se W_2 for o caminho de relaxação dominante, a saturação de *I* provocará um aumento do sinal de *S*, ou seja, um *NOE positivo em S devido a I*.

Sendo consequência indireta do acoplamento dipolar, o NOE deverá depender do movimento molecular e das distâncias interatômicas. O movimento molecular é representado pelo *tempo de correlação rotacional*, τ_c , de forma que a intensidade do NOE é dada pela equação

$$NOE \propto \frac{\tau_c}{r^6} \quad (50)$$

O tempo de correlação rotacional depende do tamanho da molécula e da viscosidade da solução. Em moléculas pequenas em solução não viscosa ocorrem NOEs positivos, enquanto que em macromoléculas, ou soluções muito viscosas, os NOEs são negativos.

2.1.7. NMR Multidimensional

No experimento mais simples de NMR, a magnetização transversal excitada por um pulso de rádio-freqüência é detectada e armazenada em intervalos de tempo regulares durante o *período de detecção*, gerando uma representação digital do FID. A transformada de Fourier do sinal digitalizado leva ao espectro unidimensional convencional, trazendo como informações o deslocamento químicos e o acoplamento escalar. Em experimentos unidimensionais mais complexos submete-se o sistema de spin à perturbações durante o período de aquisição ou durante um *período de preparação* que antecede o período de aquisição. A comparação dos espectros obtidos na presença e na ausência da perturbação fornecerá informações sobre os sistemas de spin afetados. Um exemplo são os experimentos NOE-diff, nos quais a saturação seletiva de um spin particular, durante o período de preparação, altera as intensidades de spins próximos (acoplados dipolarmente). Técnicas de NMR unidimensionais são, de forma geral, bastante aplicadas a moléculas pequenas, sendo porém inviáveis nos estudos em macromoléculas biológicas, dado o alto grau de complexidade e de sobreposição que ocorre nos espectros destas.

Em experimentos bidimensionais, dois novos elementos são introduzidos entre preparação e a aquisição: o *período de evolução* e o *período de mistura*. O esquema geral utilizado para a obtenção de espectros bidimensionais pode ser dividido em 4 partes distintas, ilustradas na figura 8.

Figura 8. Esquema geral para espectroscopia de NMR bidimensional. Os experimentos de NMR 2D podem ser divididos em quatro componentes: preparação, evolução, mistura e aquisição.

O período de evolução consiste de um intervalo de tempo variável, começando com um valor inicial que é incrementado por m passos (em geral iguais) até um valor final. A cada incremento a mesma seqüência de pulsos é executada e um FID consistido de k pontos de dados digitalizados é adquirido. O FID, por sua vez, consiste de p repetições somadas para mediação do sinal ou ciclagem de fase. Assim, um total de $m \times p$ sinais são adquiridos durante um experimento 2D, os quais são armazenados na forma de uma matriz $m \times k$ como função dos dois períodos de

tempo independentes. O tempo de aquisição (o tempo em que o o sinal é detectado) é denominado de t_2 , e o tempo de evolução indireto é denominado de t_1 . O sinal adquirido durante t_2 é modulado por eventos ocorridos durante o tempo de evolução t_1 , de forma que uma correlação exista entre os dois domínios. A transformada de Fourier dos dados adquiridos em t_2 (FID) fornece a dimensão direta F_2 , horizontal (linhas). O gráfico da amplitude do sinal de ressonância obtido em F_2 como função de t_1 forma um interferograma de aparência similar ao FID (figura 9). A transformada de Fourier do interferograma, com relação a t_1 , fornece a dimensão indireta, F_1 , vertical (colunas). O espectro obtido dessa forma será bidimensional, constando de picos nas posições em que $F_1=F_2$, ou seja, diagonais e reproduzindo o espectro 1D. Para que informações adicionais sejam fornecidas é necessário o aparecimento de picos fora da diagonal ($F_1 \neq F_2$); tais picos resultarão da *transferência de coerência* entre os spins, ocorrida durante o período de mistura. O período de mistura estabelece o tipo de correlação entre as duas dimensões, ditando a informação contida no espectro.

Figura 9. Dependência do sinal de NMR com o período de evolução, t_1 . **(a)** O sinal detectado durante o período de aquisição t_2 é modulado como função do período de evolução t_1 . **(b)** Após a transformada de Fourier dos dados adquiridos, com relação a t_2 , a amplitude do sinal de ressonância varia periodicamente como função de t_1 .

O princípio de NMR bidimensional pode ser expandido a n dimensões através da repetição evolução-mistura $n-1$ vezes. A transformada de Fourier da dimensão direta, t_n , e das dimensões indiretas $t_{(n-1)}, \dots, t_1$ produz um espectro n -dimensional com n dimensões de frequência independentes: $F_n, F_{(n-1)}, \dots, F_1$.

2.1.8. Experimentos de NMR utilizados na determinação estrutural

A seguir serão brevemente introduzidos os experimentos básicos utilizados na determinação de estrutura de proteínas. Uma descrição rigorosa pode ser encontrada nas monografias de Ernst *et al.* (1987) e Cavanagh *et al.* (1995). As monografias de T.D.W. Claridge (1999), H. Friebolin (1992) e R.S. Macomber (1998), apesar de voltadas para NMR em química orgânica, são bastante úteis para os aspectos experimentais e aplicações.

Em proteínas a grande quantidade de prótons resulta em uma intensa sobreposição dos picos nos espectros 1D, tornando sua interpretação praticamente impossível. Uma forma de resolver esse problema é com a utilização de espectroscopia bidimensional, a qual abrange uma grande quantidade de seqüências de pulsos específicas. Para proteínas maiores do que 10kDa, porém, espectros 2D também apresentam alta sobreposição de picos: técnicas 3D são necessárias. A forma mais simples de espectroscopia 3D consiste em modificar as seqüências 2D tradicionais, de forma a separar os picos sobrepostos nos espectros 2D em uma terceira dimensão, em geral com relação aos heteronúcleos ^{13}C ou ^{15}N de proteínas isotopicamente enriquecidas.

Experimentos homonucleares bidimensionais

COSY. Entre os experimentos 2D o COSY (“COrelated SpectroscopY”) foi o primeiro a ser implementado, permanecendo até hoje bastante útil em NMR de proteínas. Picos cruzados em espectros COSY se originam da transferência de coerência entre spins acoplados escalarmente (item 2.1.5). Em proteínas, essa transferência é limitada a prótons separados entre si por duas ou três ligações químicas, de forma a tornar os aminoácidos sistemas de spins isolados. Experimentos COSY poderiam ser utilizados para o assinalamento completo de cada sistema de spin, não fosse pela grande degenerescência de deslocamentos químicos na região espectral de mais alta frequência (3 a 0 ppm). A seqüência de pulsos consiste de dois pulsos, separados pelo período de evolução. Existe uma grande quantidade de variantes dessa técnica, entre elas DQF-COSY (“Double-Quantum-Filtered” COSY).

TOCSY. Também conhecido por HOHAHA (“HOmonuclear HArtmann-HAhn”), os experimentos TOCSY (“TOtal COrelated SpectroscopY”) utilizam o tempo de mistura para a transferência de magnetização entre spins via acoplamento escalar forte. Como resultado a magnetização pode ser transferida por vários acoplamentos, promovendo picos cruzados entre todas as ressonâncias de um sistema de spin, o que facilita o assinalamento se comparado ao COSY.

NOESY. Ao contrário dos experimentos TOCSY e COSY, que utilizam o acoplamento escalar como forma de transferência de magnetização, no experimento NOESY (“Nuclear

Overhauser-Effect Spectroscopy”) essa transferência ocorre via acoplamento dipolar durante o tempo de mistura. No espectro aparecerão ressonâncias entre prótons a uma distância máxima de 5Å entre si: além dos picos do espectro TOCSY, aparecerão picos representando conexões espaciais entre os sistemas de spin, de grande utilidade no assinalamento seqüencial.

Experimentos heteronucleares bidimensionais

HSQC. Em experimentos HSQC (Heteronuclear Single Quantum Coherence) é realizada a transferência de coerência entre o spin heteronuclear e o spin do próton. Experimentos de HSQC são utilizados para o assinalamento das ressonâncias ^{13}C (^{13}C -HSQC) ou ^{15}N (^{15}N -HSQC) a partir dos deslocamentos químicos dos prótons: nos espectros aparecerão picos entre prótons e ^{13}C , respectivamente ^{15}N , conectados entre si por uma ligação química.

Experimentos heteronucleares tridimensionais

TOCSY-HSQC e NOESY-HSQC. São experimentos heteronucleares editados, representando a mais simples forma de utilização de spins heteronucleares para facilitar a espectroscopia de NMR em proteínas maiores. Onde. Picos cruzados entre prótons do correspondente experimento 2D (TOCSY, NOESY) são distribuídos ao longo de planos correspondentes aos deslocamentos químicos do heteronúcleo (^{13}C , ^{15}N) diretamente ligado aos spins ^1H . Esses experimentos consistem simplesmente de uma composição das seqüências homonucleares (TOCSY, NOESY) com HSQCs.

HCCH-TOCSY e HCCH-COSY. Os experimentos HCCH-COSY (^1H - ^{13}C - ^{13}C - ^1H “correlation spectroscopy”) e HCCH-TOCSY (^1H - ^{13}C - ^{13}C - ^1H “total correlation spectroscopy”) são utilizados no assinalamento de ressonâncias ^1H e ^{13}C alifáticos. Esses experimentos permitem a dispersão dos picos dos espectro COSY e TOCSY ao longo de uma terceira dimensão ^{13}C , utilizando 3 passos de transferência de magnetização por acoplamento escalar: primeiro do ^1H ao ^{13}C a ele ligado, depois do ^{13}C ao seu próximo vizinho ^{13}C , por fim do ^{13}C ao ^1H a ele ligado. O que permite essas transferências específicas são as diferenças nas constante de acoplamento J em cada passo: $^1J_{\text{CH}}$ (acoplamento entre carbono e hidrogênio diretamente ligados) é da ordem de 140Hz, enquanto que $^1J_{\text{CC}}$ (acoplamento entre carbonos diretamente ligados entre si) varia entre 32 e 40Hz.

HNCO, HNCA, HN(CO)CA e outros. Experimentos de tripla ressonância correlacionando spins $^1\text{H}_{\text{N}}$, ^{15}N , $^1\text{H}_{\alpha}$, $^{13}\text{C}_{\alpha}$ e ^{13}CO através da interação de acoplamento escalar de uma ou duas ligações. A nomenclatura dos experimentos indica os caminhos de correlação: HN, N, HA, CA e CO representam, respectivamente, os spins $^1\text{H}_{\text{N}}$, ^{15}N , $^1\text{H}_{\alpha}$, $^{13}\text{C}_{\alpha}$ e ^{13}CO . Os spins que aparecerão nas dimensões indiretas (F1, F2) e na direta (F3) são indicados no nome do experimento; spins

através dos quais a coerência será apenas transferida são indicados entre parêntesis. No caso de transferência entre $^1\text{H}_\text{N}$ e ^{15}N , a nomenclatura de $^1\text{H}_\text{N}$ é abreviada para H (ao invés de HN). Dessa forma, o experimento HNCO irá correlacionar os spins $^1\text{H}_\text{N}$ e ^{15}N de um aminoácido com os spins ^{13}CO do aminoácido anterior. O experimento HNCA irá correlacionar os spins $^1\text{H}_\text{N}$, ^{15}N e $^{13}\text{C}_\alpha$ de um mesmo aminoácido (pico forte) e os spins $^1\text{H}_\text{N}$, ^{15}N de um aminoácido com o spin $^{13}\text{C}_\alpha$ do aminoácido anterior (pico fraco). O experimento HN(CO)CA irá correlacionar os spins $^1\text{H}_\text{N}$, ^{15}N de um aminoácido com o spin $^{13}\text{C}_\alpha$ do aminoácido anterior. Uma infinidade de experimentos podem ser criados dessa forma, utilizando os mais variados caminhos de correlação. Tais experimentos são muito úteis no assinalamento seqüencial de proteínas.

2.2. NMR NO ESTUDO DE PROTEÍNAS E PEPTÍDEOS

A metodologia para a determinação de estrutura de proteínas por espectroscopia de NMR de alta resolução estabeleceu-se nos últimos 15 anos, graças às contribuições essenciais dos grupos de Ernst (1987), no desenvolvimento de seqüências de pulsos, e Wüthrich (1986), nas estratégias de análise de dados. Uma vez preparada a amostra e obtidos os espectros de NMR convenientes, a determinação estrutural segue basicamente 4 passos:

- identificação dos sistemas de spin e assinalamento seqüencial;
- assinalamento de restrições experimentais, como distâncias interprotônicas e ângulos torcionais;
- determinação, através de algoritmos de dinâmica molecular, de um conjunto de estruturas que satisfaçam as restrições experimentais;
- refinamento das estruturas via cálculo retroativo de espectros e/ou análise das estruturas de mais baixa energia.

Experimentos adicionais como titulação e troca ($\text{H}_2\text{O} \rightarrow \text{D}_2\text{O}$) fornecem informações sobre a carga e a dinâmica da molécula analisada.

2.2.1. Sistemas de Spin e Assinalamento Seqüencial

No estágio inicial de qualquer estudo de proteínas por espectroscopia de NMR, cada ressonância magnética nuclear deve ser associada a um núcleo específico na molécula a ser investigada. Essa associação deve ser específica à seqüência, ou seja, cada ressonância deve ser associada ao spin de um determinado resíduo de aminoácido na seqüência da proteína. A espectroscopia por NMR fornece 3

tipos de informações úteis para o assinalamento espectral: a interação por ligações (via acoplamento escalar), a interação através do espaço (via acoplamento dipolar) e o ambiente químico (via deslocamento químico). Conforme o tipo de amostra e equipamento disponíveis, estratégias homonucleares (^1H) ou heteronucleares (^{13}C e/ou ^{15}N) serão utilizadas no assinalamento sequencial.

Em proteínas as correlações resultantes do acoplamento escalar ^1H - ^1H são, de forma geral, observadas apenas entre prótons separados por 2 ou 3 ligações químicas. Os picos cruzados em espectros 2D de NMR de correlação escalar homonuclear ^1H ocorrem, portanto, apenas entre prótons de um mesmo aminoácido (ou sistema de spin) e nunca entre prótons inter-residuais. Dessa forma, experimentos 2D baseados em correlação escalar, como o COSY e o TOCSY, poderão ser utilizados no assinalamento dos aminoácidos.

Inicialmente as ressonâncias ^1H são classificadas de acordo com seus deslocamentos químicos. Na grande maioria dos resíduos de uma proteína, a ressonância dos prótons segue um determinado padrão:

- entre 10,0 e 7,0 ppm para os prótons amídicos (HN) da cadeia principal;
- entre 8,0 e 6,0 ppm para os prótons amídicos (HN) das cadeias laterais, e para os prótons aromáticos;
- entre 6,0 e 3,5 ppm para os prótons alfa ($\text{H}\alpha$);
- entre 3,5 e 1,0 ppm para prótons alifáticos ($\text{H}\beta$, $\text{H}\gamma$, ...) da cadeia lateral;
- abaixo de 1,5 ppm para prótons metílicos.

O procedimento para o assinalamento das ressonâncias, conhecido como *estratégia de assinalamento sequencial*, foi desenvolvido por Wüthrich e colaboradores [70]. Em uma primeira etapa, conjuntos de ressonâncias NH, $\text{H}\alpha$ e de prótons alifáticos pertencentes a um mesmo aminoácido são identificados, utilizando-se o fato de que picos cruzados conectam apenas prótons de um mesmo sistema de spin. Uma proteína com N resíduos dará origem a $(N-E)$ sistemas de spin do tipo HN- $\text{H}\alpha$ -alifáticos, onde E significam as exceções dadas pelas Prolinas (que não possuem o próton amídico) e o aminoácido N-terminal do tipo NH_3 . Nesses últimos os sistemas de spin terão a forma $\text{H}\alpha$ -alifáticos. Os prótons aromáticos dos aminoácidos Tyr, Phe, Trp, His, e os prótons amídicos das cadeias laterais dos aminoácidos Glu e Asp não são acoplados escalarmente ao resto das respectivas cadeias laterais. A conexão entre os prótons nesses sistemas de spin só poderá ser concluída através de sua conexão espacial, dada pelo acoplamento dipolar (NOE). A seguir são ilustrados (figuras 10a-e) e brevemente descritos os sistemas de spin característicos de cada aminoácido, tipicamente encontrados em espectros TOCSY, e suas peculiaridades:

- GLICINA (sistema AX) – Único aminoácido que possui dois sinais HN-H α 1 e HN-H α 2 que, quando não degenerados, são indicativo de que o aminoácido se encontra em uma região molecular de baixa mobilidade, ou seja, estruturada;
- ALANINA (sistema A3X) – A Alanina possui um pico cruzado HN-H β * (onde * indicam os 3 prótons do grupo metil), intenso e estreito, o que a torna muito fácil de ser identificada;
- TREONINA (sistema A3MX) – O pico HN-H γ 2* é encontrado na mesma região que o HN-H β * da Alanina, sendo porém menos intenso. A presença de picos cruzados com H γ 1 é indicativo de que esse átomo forma ponte de hidrogênio;
- VALINA (sistema A3B3MX) – Os dois picos HN-H γ 1* e HN-H γ 2* são relativamente intensos, porém, quando degenerados, podem levar à confusão com a Treonina;

Figura 10a. Espectros da Glicina, Alanina, Treonina e Valina. Os círculos pretos indicam picos encontrados em espectros COSY e TOCSY, os círculos cinzas são apenas encontrados nos espectros TOCSY. Quadrados mostram as conexões espaciais encontradas apenas nos espectros NOESY.

- LEUCINA (sistema A3B3MPTX) – A longa cadeia da Leucina resulta, em geral, em um sistema de spin incompleto (incompleta transferência da magnetização), o que dificulta bastante a sua identificação;
- ISOLEUCINA (sistema A3MPT(B3)X) – Com padrão semelhante ao da Leucina, é difícil de ser identificada;

Figura 10b. Espectros da Leucina e Isoleucina (conforme 3a).

- SERINA (sistema AMX) – Os deslocamentos químicos dos prótons $H\beta_1$ e $H\beta_2$ são bastante próximos ao do próton $H\alpha$, podendo levar, no caso de degenerescência $H\beta_1/H\beta_2$, a um padrão semelhante ao de uma Glicina não degenerada. A inspeção do espectro COSY permite uma identificação clara. Assim como na Treonina, a presença de picos cruzados com $H\gamma$ é indicativo de que esse átomo forma ponte de hidrogênio;
- CISTEÍNA (sistema AMX) – Com sistema idêntico ao do Aspartato, somente pode ser diferenciada após o assinalamento seqüencial;
- ASPARTATO (sistema AMX) – Com sistema idêntico ao da Cisteína, somente pode ser diferenciado após o assinalamento seqüencial;
- ASPARAGINA (sistema AMX) – Através do espectro NOESY encontra-se a conexão espacial entre o sistema principal e os prótons amídicos $H\delta_{21}$ e $H\delta_{22}$, o que permite diferenciar a Asparagina do Aspartato e da Cisteína;

Figura 10c. Espectros da Serina, Cisteína, Aspartato e Asparagina (conforme 3a).

- HISTIDINA (sistema AMX) – Sistema semelhante ao da Asparagina, sendo que a conexão espacial via NOESY é somente vista entre o sistema principal e o próton aromático H δ 2;
- TIROSINA (sistema AMX) – O intenso pico entre os prótons aromáticos H δ 1/H δ 2 e H ϵ 1/H ϵ 2 permite diferenciá-la da histidina;
- FENILALANINA (sistema AMX) – O padrão único formado pelos prótons aromáticos pode ser conectado ao sistema principal via picos NOESY entre H β 1, H β 2 e H δ 1/H δ 2;
- TRIPTOFANO (sistema AMX) – O duplo anel do triptofano forma dois sistemas aromáticos independentes, conectados entre si via pico NOESY entre H ζ 2 e H ϵ 1. O incomum valor para o deslocamento químico do próton H ϵ 1 torna esse sistema facilmente identificável. A conexão entre anéis e o sistema principal se dá via picos NOESY entre os prótons H β 1, H β 2 e H δ 1, H ϵ 3;

Figura 10d. Espectros da Histidina, Tirosina, Fenilalanina, Triptofano (conforme 3a).

- **METIONINA** (sistema AM(PT)X) – O intenso pico metílico $H\epsilon^*$ é facilmente identificável no espectro 1D e sua correlação com o resto do aminoácido se dá apenas espacialmente. A semelhança entre os deslocamentos químicos dos prótons $H\beta_1$, $H\beta_2$ e $H\epsilon^*$ dificulta a correlação espacial, tornando o sistema idêntico ao do Glutamato. A identificação exata somente pode ser obtida após o assinalamento sequencial;
- **GLUTAMATO** (sistema AM(PT)X) – Sistema idêntico ao da Metionina, exceto pelo pico metílico $H\epsilon^*$. Somente pode ser diferenciado após o assinalamento sequencial;
- **GLUTAMINA** (sistema AM(PT)X) - Através do espectro NOESY encontra-se a conexão espacial entre o sistema principal e os prótons amídicos $H\epsilon_{21}$ e $H\epsilon_{22}$, o que permite diferenciá-la do Glutamato e da Metionina.

Figura 10e. Espectros da Metionina, Glutamato, Glutamina, Prolina, Arginina e Lisina (conforme 3a).

- PROLINA (sistema A2T2(MP)X) – O padrão característico formado pelos picos cruzados entre H δ 1 e H δ 2 e os prótons restantes, associados com a ausência do próton amídico torna a Prolina facilmente identificável;
- ARGININA (sistema A2T2(MP)X) – Com padrão semelhante ao da Prolina, distingue-se desta pela (em geral) degenerescência dos prótons H δ 1/H δ 2 e pela presença de picos cruzados com o próton amídico;
- LISINA (sistema A2(F2T2)MPX) – A presença do pico cruzado entre prótons H δ 1/H δ 2 e H ϵ 1/H ϵ 2 é bastante intenso no espectro COSY, o que auxilia a diferenciá-la da Arginina;

O envolvimento do aminoácido em elementos de estrutura secundária e o ambiente químico em que o mesmo se encontra pode provocar desvios significativos nos deslocamentos químicos de seus prótons, de forma que um assinalamento baseado puramente nos valores tabelados não seja possível.

Figura 11. Espectros fictícios do tetrapeptídeo Met-Gly-Ala-Glu. **(a)** Espectro TOCSY, mostrando os picos típicos dos 4 sistemas de spin: Metionina (vermelhos), Glicina (azuis), Alanina (verdes) e Glutamato (amarelos). A região “fingerprint” dos picos entre prótons NH e H α está destacada. **(b)** Região “fingerprint” do espectro NOESY. A seta indica o início do “noe-pass”, originado do valor de H α da Met1. Quadrados vazios simbolizam picos NOEs entre o próton H α de um aminoácido e o próton amídico HN do aminoácido seguinte ($d_{H\alpha(i)-HN(i+1)}$). Quadrados com círculo inscrito representam NOEs intraresiduais; os círculos provem do espectro TOCSY.

No segundo passo do processo de assinalamento, cada sistema de spin identificado via correlações escalares é associado a um particular resíduo na cadeia polipeptídica, utilizando-se a conexão espacial dada pela interação de acoplamento dipolar (NOE). O procedimento consiste em identificar, na região NH-H α (“fingerprint region”) do espectro NOESY, os intensos picos $d_{H\alpha(i)-HN(i+1)}$ entre o próton H α de um aminoácido e o próton amídico HN do aminoácido seguinte. Por serem decorrentes de contatos inter-residuais, tais picos não são encontrados nos espectros TOCSY. A identificação de uma série de interações NOE seqüenciais permite que n sistemas de spin sejam colocados na ordem $i, i+1, \dots, i+n-1$. Conforme cresce n , a seqüência de sistemas de spin se ajusta a uma única secção da seqüência de aminoácidos da proteína, conduzindo ao assinalamento seqüencial. A figura 11 traz um exemplo de assinalamento em um tetrapeptídeo Met-Gly-Ala-Glu; o mapa gerado é denominado “NOE-pass”.

Nos últimos anos, diversos algoritmos tem sido desenvolvidos visando o assinalamento automatizado dos sistemas de spin. Uma automação parcial é possível, porém efeitos como degenerescência de ressonâncias, sobreposição de picos e correlações incompletas impedem uma automação completa.

Dispondo-se de proteínas enriquecidas com ^{13}C e/ou ^{15}N , experimentos de correlação heteronucleares podem ser utilizados. Os deslocamentos químicos ^{13}C e ^{15}N para cada aminoácido podem ser obtidos através dos experimentos bidimensionais ^{13}C -HSQC e ^{15}N -HSQC, nos quais as ressonâncias de ^{13}C e ^{15}N são correlacionadas com as ressonâncias protônicas. No experimento 3D ^1H - ^{15}N TOCSY-HSQC (ou NOESY-HSQC) obtém-se os mesmos picos como no espectro TOCSY (ou NOESY) bidimensional, distribuídos em diferentes planos de acordo com os deslocamentos químicos dos núcleos ^{15}N de cada aminoácido; o mesmo pode ser feito para os núcleos ^{13}C . A estratégia básica do assinalamento em experimentos bidimensionais é mantida, com a vantagem de diminuir significativamente a sobreposição de picos.

Alternativamente pode-se utilizar experimentos HCCH-COSY e HCCH-TOCSY, combinados à experimentos de tripla ressonância, como o HNCA. Analogamente ao TOCSY-HSQC, os experimentos HCCH-COSY e HCCH-TOCSY distribuem os picos dos respectivos experimentos COSY e TOCSY bidimensionais em planos de acordo com os deslocamentos químicos ^{13}C , desde que esses picos provenham de prótons correlacionados entre si via ligação ^1H - ^{13}C - ^{13}C - ^1H . Os demais prótons não geram sinais nos espectros HCCH; a correlação com o próton amídico é perdida. Experimentos de tripla ressonância recuperam essa correlação: por

exemplo, no HNCA, picos fortes são vistos entre prótons amídicos e carbonos- α , intraresiduais, separados na terceira dimensão do nitrogênio amídico. Um segundo conjunto de picos, mais fracos, pode ser observado para cada plano ^{15}N ; esses picos resultam da conexão inter-residual entre o próton amídico do aminoácido e o carbono- α do aminoácido anterior, de forma que o assinalamento seqüencial também pode ser obtido via experimento HNCA. Diversos outros experimentos de tripla ressonância podem ser utilizados tanto para correlações intra-residuais quanto para correlações seqüenciais, com uma clara vantagem sobre os experimentos NOESY 2D e 3D: nesses últimos os picos seqüenciais na região “fingerprint” podem ser confundidos com picos não-seqüenciais, resultantes da proximidade estrutural dos átomos na proteína, levando a erros no assinalamento seqüencial.

Uma primeira informação estrutural já pode ser obtida após o assinalamento seqüencial, comparando-se os deslocamentos químicos dos prótons da cadeia principal com o valores tabelados para os 20 aminoácidos em estrutura randômica. O ambiente químico em torno do próton amídico e do próton $\text{H}\alpha$ é bastante sensível à existência de elementos de estrutura secundária. Quando envolvidos em hélices, a blindagem devido à nuvem eletrônica é diminuída e esses prótons ressoam em mais baixas frequências, levando a um desvio negativo do deslocamento químico. Já em fitas- β a blindagem aos prótons é aumentada, resultando em ressonâncias a mais altas frequências, e portanto um desvio positivo no deslocamento químico. Em proteínas isotopicamente enriquecidas os deslocamentos químicos ^{13}C e ^{15}N também podem ser comparados com os valores para o hexapeptídeo Gly-Gly-*X-Ala*-Gly-Gly e Gly-Gly-*X-Pro*-Gly-Gly (apêndice A) [71], fornecendo igualmente informações estruturais. Nitrogênios amídicos da cadeia principal se comportam da mesma forma que os prótons amídicos $\text{H}\alpha$; já para os carbonos da cadeia principal o comportamento é oposto.

2.2.2. Distâncias Interprotônicas

A determinação da estrutura tridimensional de macromoléculas por NMR de alta resolução é baseada, principalmente, em informações de distâncias interprotônicas. Em experimentos do tipo NOESY obtém-se picos cruzados (NOEs) entre ressonâncias de prótons próximos no espaço; o volume V_{NOE} desse sinal está relacionado com a distância interprotônica d_{NOE} através da relação

$$V_{\text{NOE}} \propto d_{\text{NOE}}^{-6} . \quad (51)$$

Essa dependência com o inverso da sexta potência da distância leva ao rápido decaimento na intensidade do NOE, limitando a distância máxima observável em $\sim 5\text{\AA}$. Para calibração dos volumes utilizam-se referências conhecidas, por exemplo a distância fixa de $1,8\text{\AA}$ entre prótons metilênicos ou também a distância de $2,5\text{\AA}$ entre prótons vizinhos no anel benzênico.

Uma variedade de diferentes métodos para a integração de volumes de picos em espectros NOESY tem sido apresentada nos últimos anos, substituindo o método simplificado (e possivelmente ainda mais utilizado) de estimar o volume a partir da amplitude do pico. Amplitudes de picos podem ser obtidas automaticamente por procedimentos simples de “peak picking” ou de uma inspeção visual dos dados contando-se as curvas de níveis. Como extensão desse método simplificado, pode-se efetuar a multiplicação das intensidades por suas meias-larguras em cada dimensão. Melhores resultados podem ser obtidos por simulação dos picos com curvas Lorentzianas ou de outras formas, de cujos parâmetros podem ser calculados os respectivos volumes. Uma alternativa a todos esses métodos é a integração numérica direta, possuindo a vantagem de ser independente da forma de curva assumida para o pico. Em sua forma básica, consiste de dois passos sucessivos: primeiro, a definição de todos os pontos de dados que são parte de um mesmo pico, e segundo, a adição das correspondentes intensidades para se obter a integral do pico. Entre os vários algoritmos numéricos desenvolvidos nos últimos anos destaca-se a integração automática de picos em espectros de NMR multidimensional, baseado em um procedimento de segmentação interativa otimizada [72]. Esse algoritmo foi desenvolvido no Max-Planck-Institut für Medizinische Forschung em colaboração com a Bruker, e incorporado ao programa AURELIA [73].

Os NOEs assinalados podem ser classificados em 4 categorias: *intra-residuais* (entre prótons de um mesmo aminoácido i), *seqüenciais* (entre prótons de um aminoácido i e seu primeiro vizinho $i+1$), *médios* (entre prótons de um aminoácido i e seu segundo $i+2$, terceiro $i+3$ e quarto vizinho $i+4$) e *longos* (os demais). NOEs médios são característicos de estruturas em hélice, enquanto que os longos estabilizam folhas- β e a estrutura terciária. Os contatos típicos encontrados em estruturas secundárias estão listados na tabela 4. Esses contatos, quando associados à seqüência de aminoácidos, geram o mapa de contatos NOE (“NOE-contact map”).

Tabela 4. Distâncias inter-protônicas sequenciais típicas (em Å) dos elementos de estrutura secundária. O símbolo $d_{H\alpha(i)-HN(i+1)}$, por exemplo, representa a distância entre o próton $H\alpha$ do aminoácido i e o próton amídico de $i+1$.

	Hélice- α $\varphi = -57^\circ$ $\psi = -47^\circ$	Hélice 3_{10} $\varphi = -60^\circ$ $\psi = -30^\circ$	β -antiparalela $\varphi = -139^\circ$ $\psi = +135^\circ$	β -paralela $\varphi = -119^\circ$ $\psi = +113^\circ$
$d_{H\alpha(i)-HN(i)}$	2,7	2,7	2,8	2,8
$d_{H\alpha(i)-H\beta(i)}$	2.2–2.9	2.2–2.9	2.2–2.9	2.2–2.9
$d_{H\beta(i)-HN(i)}$	2.0–3.4	2.0–3.4	2.4–3.7	2.6–3.8
$d_{H\alpha(i)-HN(i+1)}$	3,5	3,4	2,2	2,2
$d_{HN(i)-HN(i+1)}$	2,8	2,6	4,3	4,2
$d_{H\beta(i)-HN(i+1)}$	2.5–3.8	2.9–3.0	3.2–4.2	3.7–4.4
$d_{H\alpha(i)-HN(i+2)}$	4,4	3,8		
$d_{HN(i)-HN(i+2)}$	4,2	4,1		
$d_{H\alpha(i)-HN(i+3)}$	3,4	3,3		
$d_{H\alpha(i)-H\beta(i+3)}$	2,5–4,4	3,1–5,1		
$d_{H\alpha(i)-HN(i+4)}$	4,2			

2.2.3. Ângulos Torcionais

Informações complementares, como ângulos torcionais, podem ser bastante úteis na determinação da estrutura tridimensional de uma proteína, restringindo o espaço conformacional. Um ângulo torcional θ da cadeia peptídica está relacionado ao acoplamento 3J vicinal pela equação de Karplus [74]:

$$^3J(\theta) = A \cos^2 \theta + B \cos \theta + C \quad . \quad (52)$$

Os parâmetros A , B e C variam conforme os ângulos torcionais e foram determinados empiricamente por diversos grupos [75–77]. O ângulo ω é fixo em $\sim 180^\circ$ devido ao caráter de ligação dupla parcial da ligação peptídica CN ; o ângulo ψ só pode ser determinado por experimentos heteronucleares (^{15}N - ou ^{13}C -/ 1H -NMR) porém, em proteínas, somente certas combinações de pares φ, ψ são possíveis. Os ângulos φ e χ^1 podem ser determinados via constantes de acoplamento $^3J_{NH\alpha}$ e $^3J_{\alpha\beta}$ pela equação de Karplus, valendo os seguintes parâmetros:

- $A=6,4$; $B=-1,4$; $C=1,9$ para o ângulo φ , exceto Glicina e Prolina [77];
- $A=-9,4$; $B=-1,1$; $C=14,9$ para o ângulo φ (Glicina) [75];
- $A=9,5$; $B=-1,6$; $C=1,8$ para o ângulo χ^1 [76];
- Devido à estrutura em anel para com o Nitrogênio, a Prolina compõe um sistema próprio que não pode ser representado por essa equação.

Os valores das constantes de acoplamento J são obtidos por simulação de curvas Lorentzianas (dubletos) sobre os picos de espectros COSY. A equação de Karplus fornece até 4 soluções para o ângulo φ , sendo que geralmente as excedentes podem ser descartadas devido a impedimentos estéricos ou pela análise de elementos de estrutura secundária (que fixa os ângulos em torno de valores característicos). Para o ângulo χ^1 (3 soluções possíveis) a conformação exata somente pode ser determinada por análise rotamérica, após identificação estereo-específica dos prótons de metileno não degenerados (exceto Alanina e Treonina). Para Valina e Isoleucina, que possuem somente um próton $H\beta$, devem ser levados em conta os contatos NOE aos grupos metílicos $H\gamma_1$, $H\gamma_2$.

De modo geral, acoplamentos J pequenos ($<6\text{Hz}$) são indicativos de estruturas helicoidais, enquanto que acoplamentos grandes ($>8\text{Hz}$) indicam a presença de uma fita- β .

2.2.4. Análise Rotamérica

Na identificação dos sistemas de spin pelos métodos anteriormente descritos, a quebra de degenerescência dos prótons de grupos metileno e metil leva a uma ambigüidade no assinalamento estereoespecífico, com conseqüente incerteza nas restrições de distâncias. Uma análise detalhada da estrutura fina em espectros COSY sensíveis à fase, em adição à observação de picos específicos em espectros NOESY, podem possibilitar essa identificação estereoespecífica. Esse procedimento foi descrito por Basus, em 1989 [78].

Assinalamento estereoespecífico de prótons β Para esses prótons o assinalamento é centrado na definição do ângulo de rotação χ^1 ao redor da ligação $H\alpha-H\beta$. As configurações energeticamente mais favoráveis são aquelas com os átomos ao redor dessa ligação em uma conformação empilhada: os prótons $H\beta$ podem estar *gauche* (*g*) ou *trans* (*t*) em relação ao próton $H\alpha$. Existem 3 possíveis rotâmeros: g^2g^3 , g^2t^3 e t^2g^3 , onde os índices 2 e 3 indicam os prótons $H\beta_2$ e $H\beta_3$ definidos de acordo com a convenção IUB-IUPAC. A identificação dos rotâmeros e o assinalamento estereoespecífico se dá a partir da análise das constantes acoplamento $^3J_{\alpha\beta_2}$ e $^3J_{\alpha\beta_3}$ (acoplamento J do próton $H\alpha$ para prótons $H\beta_2$ e $H\beta_3$), dos NOE(α,β_2) e NOE(α,β_3) (do próton $H\alpha$ para prótons $H\beta_2$ e $H\beta_3$) e dos

NOE(NH, β^2) e NOE(NH, β^3) (do próton NH para prótons H β^2 e H β^3). Um esquema pode ser visto na tabela 5. O caso g^2g^3 pode ser facilmente identificado a partir das constantes de $^3J_{\alpha\beta^2}$ e $^3J_{\alpha\beta^3}$, uma vez que ambas deverão ser pequenas. Os NOE($\alpha\beta^2$) e NOE($\alpha\beta^3$) serão idênticos, pois ambos os prótons H β estão aproximadamente equidistantes do próton H α . Porém o assinalamento estereoespecífico dos prótons H β^2 e H β^3 só será possível a partir da análise dos NOE(NH, β^2) e NOE(NH, β^3): uma rotação do ângulo torcional ϕ em torno dos intervalos possíveis fornece distâncias de 3.5 e 4.0Å (NOE forte-médio) entre prótons NH e H β^2 , e distâncias de 2.5 a 3.4Å (NOEs fracos) entre prótons NH e H β^3 . Os outros dois rotâmeros, g^2t^3 e t^2g^3 , não podem ser diferenciados um do outro a partir das constantes de acoplamento, em ambos os casos distintas. A diferenciação se dá a partir dos NOE(NH, β^2) e NOE(NH, β^3): no caso g^2t^3 os NOEs serão respectivamente forte-médio e forte (distância de 2.2 a 3.1Å), e no caso t^2g^3 serão respectivamente forte e fraco. Uma vez obtida essa diferenciação, o assinalamento estereoespecífico pode ser feito com base no acoplamento J .

Tabela 5. Caracterização das três conformações possíveis em torno da ligação C α -C β . As orientações *gauche* e *trans* estão indicadas como *g* e *t*, respectivamente. As posições β^2 e β^3 são definidas de acordo com a convenção IUB-IUPAC. Uma projeção Newman é mostrada para cada uma das conformações e o ângulo χ^1 é indicado. Os intervalos de distâncias indicados são baseados na rotação do ângulo torcional ϕ através de todos os intervalos de distâncias permitidos em resíduos distintos de glicinas. As constantes de acoplamentos são baseadas na equação empírica $J(\theta) = (9.5 \pm 0.3) - (1.6 \pm 0.2) \cos \theta + (1.8 \pm 0.6) \text{ Hz}$.

Conformação	g^2g^3	g^2t^3	t^2g^3
χ^1	60°	180°	-60°
$^3J_{\alpha\beta^2}$ (Hz)	2.6-5.1	2.6-5.1	11.8-14.0
$^3J_{\alpha\beta^3}$ (Hz)	2.6-5.1	11.8-14.0	2.6-5.1
NOE (α , β^2)	Forte	Forte	Fraco
NOE (α , β^3)	Forte	Fraco	Forte
NOE (NH, β^2)	Fraco	Forte-Médio	Forte
d_{NH,β^2} (Å)	3.5-4.0	2.5-3.4	2.2-3.1
NOE (NH, β^3)	Forte-Médio	Forte	Fraco
d_{NH,β^3} (Å)	2.5-3.4	2.2-3.1	3.5-4.0

Assinalamento estereoespecífico de prótons metil de valinas. Em determinados casos pode-se obter o assinalamento estereoespecífico de grupos metil de valinas. Para tal, observam-se a constante acoplamento $^3J_{\alpha\beta}$ (acoplamento J do próton $H\alpha$ para o próton $H\beta$), os NOE($\alpha,\gamma1$) e NOE($\alpha,\gamma2$) (do próton $H\alpha$ para os grupos metil $H\gamma1$ e $H\gamma2$) e os NOE(NH, $\gamma1$) e NOE(NH, $\gamma2$) (do próton NH para os grupos metil $H\gamma1$ e $H\gamma2$). Os rotâmeros (g^+ , g^- e t) possíveis ao redor do ângulo $C\alpha-C\beta$ são mostrados na tabela 6. Destes, a conformação g^+ é a predominante conforme demonstrado em uma análise estatística de cadeias laterais de valinas em estruturas resolvidas por Raios-X. Neste caso, tem-se uma constante de acoplamento $^3J_{\alpha\beta}$ em torno de 12 Hz, sendo muito distinta daquelas apresentadas pelos outros rotâmeros g^- e t . O assinalamento estereoespecífico pode ser feito analisando-se os NOE(NH, $H\gamma1$) e NOE(NH, $H\gamma2$): o primeiro é forte-médio e o segundo fraco. Os outros dois rotâmeros possuem constante de acoplamento $^3J_{\alpha\beta}$ menor que 5 Hz, e sua diferenciação se dá através dos NOE(NH, $H\gamma1$) e NOE(NH, $H\gamma2$): para a conformação g^- , o primeiro é forte e o segundo forte-médio; para a conformação t , o primeiro é fraco e o segundo é forte. Constantes de acoplamento $^3J_{\alpha\beta}$ em torno de 7-8 Hz são frequentemente observadas. Em aminoácidos livres, a cadeia lateral é livre para rotacionar rapidamente, mediando a constante de acoplamento para 7 Hz. Medidas de constantes de acoplamento de 7-8 Hz em valinas são indicativos para rotação livre, apesar de também poderem ser consequência de conformações energeticamente não favoráveis.

Tabela 6. Caracterização dos três rotâmeros possíveis em torno de $C\alpha - C\beta$ ligados em valinas. Uma projeção de Newman é mostrada para cada conformação. As constantes de acoplamentos foram calculadas como descrito na tabela 4. Os prótons metil $\gamma1$ e $\gamma2$ são definidos de acordo com a convenção IUB-IUPAC.

Conformação	g^-	t	g^+

$^3J_{\alpha\beta}$ (Hz)	2.6-5.1	2.6-5.1	11.8-14.0
NOE (α, γ^1)	Fraco	Forte	Forte
NOE (α, γ^2)	Forte	Fraco	Forte
NOE (NH, γ^1)	Forte	Fraco	Forte-Médio
NOE (NH, γ^2)	Forte-Médio	Forte	Fraco

Assinalamento estereoespecífico em outros resíduos. Em geral, o assinalamento estereoespecífico pode ser realizado com base em NOEs entre resíduos distantes sequencialmente na proteína (*long range* NOE), apesar de não haver relatos desse tipo de assinalamento na literatura até então.

2.2.5. Peptídeos e o uso de TFE

Proteínas e peptídeos, quando em solução, são macromoléculas dinâmicas oscilando em torno de uma estrutura média em diversas escalas de tempo, variando desde picosegundos até segundos. Enquanto que em proteínas a presença de uma grande quantidade de interações levam à estabilização diversos elementos de estrutura secundária, e ao empacotamento em uma estrutura terciária bem definida, peptídeos tendem a assumir uma variedade de estados conformacionais. A existência de compostos que estabilizem estruturas secundárias em peptídeos, como SDS (dodecil sulfato de sódio) e diversos álcoois fracamente polares, tem se mostrado bastante útil: tais compostos possibilitam a obtenção de informações a respeito de uma possível tendência estrutural de peptídeos em solução.

O TFE, ou 2,2,2-trifluoretanol ($\text{CF}_3\text{CH}_2\text{OH}$), é um solvente orgânico, cujo efeito sobre a estrutura de peptídeos tem sido estudada há mais de 3 décadas. Na maioria dos casos, a adição de TFE (10-50%) potencializa a helicidade de peptídeos que já se encontram helicoidais em solução, ou induz a helicidade em peptídeos essencialmente desordenados. Já foram reportados alguns casos de estabilização de folhas- β e voltas- β pela adição de TFE, e até mesmo exemplos de alteração conformacional em peptídeos (estrutura- β para hélice) [79]. De uma forma geral, TFE é utilizado como um solvente mimetizador de membrana, uma vez que induz estruturas de uma forma similar à produzida pelo ambiente da membrana. Por outro lado, estudos em proteínas globulares em soluções de TFE/ H_2O mostram que, em geral, parte da estrutura terciária é perdida, sendo mantida a estrutura secundária. Já foram relatados, inclusive, casos em que a adição de TFE estabiliza estruturas secundárias não-nativas. Com relação à estruturas quaternárias, o TFE pode tanto aumentar quanto interromper as interações.

Ainda que o mecanismo exato pelo qual o TFE produz essa gama de mudanças na estrutura de proteínas e peptídeos ainda não seja conhecido, algumas de suas propriedades já puderam ser explicadas a partir das características da molécula (figura 9a). A presença do grupo trifluormetil (CF_3) cria uma região hidrofóbica, permitindo que moléculas de TFE

interajam com os resíduos em peptídeos desenovelados. Por sugestão de Rajam et al. [79] deve ocorrer uma substituição da camada de hidratação em torno do peptídeo; sendo a molécula de TFE ~9 vezes maior que a de H_2O , a substituição das moléculas de água pelas do trifluoretanol representam uma contribuição entrópica favorável à energia livre total de interação. Dada a característica de bom doador e mal aceitor do grupo OH da molécula de TFE, a expulsão das moléculas de água facilita a formação de pontes de hidrogênio intramoleculares entre grupos amídicos peptídicos. A água é conhecida por romper, ou enfraquecer, as ligações de hidrogênio entre $C=O$ e $N-H$, desestabilizando hélices [80]; o mesmo não ocorre com a presença de TFE, dada a habilidade do grupo $C=O$ em formar ligações de hidrogênio bifurcadas (figura 12b).

Figura 12. Características da molécula de TFE: (a) suas diferentes regiões; (b) um modelo de sua interação com uma ligação de hidrogênio peptídica intramolecular, quando comparada à da água.

Em proteínas globulares, os efeitos do TFE devem decorrer da incorporação das moléculas de trifluorálcool no interior da proteína. A habilidade do TFE em interagir especificamente com resíduos aromáticos [79] deve ter uma importância especial nos diversos efeitos observados. A entrada de moléculas de TFE deve resultar em uma alteração no empacotamento das cadeias laterais internas de maior mobilidade, resultando em aparente perda de estrutura terciária da proteína.

2.3. CÁLCULO DE ESTRUTURAS TRIDIMENSIONAIS

A determinação da estrutura tridimensional de macromoléculas por NMR é feita a partir de dados estruturais inferidos, como distâncias inter-protônicas (a partir dos NOEs) e ângulos torcionais (acoplamento J). Tais dados servem como *restrições experimentais* em protocolos de “simulated annealing” (SA), baseados em programas de dinâmica molecular (MD) e de minimização de energia (EM), a serem aplicados sobre a estrutura primária da proteína. Devido ao baixo limite de distâncias possíveis de serem medidas ($d_{\text{NOE}} \leq 0,5$ nm) e à dubiedade de muitos dos picos NOE, uma grande variedade de estruturas tridimensionais pode ser coerente com os dados experimentais. Essa família de estruturas serve como indicativo da boa convergência do programa: setores rígidos e bem definidos da macromolécula se sobrepõem; setores flexíveis, ou não suficientemente bem definidos, divergem.

A acurácia das estruturas obtidas depende diretamente do número de restrições experimentais obtidas (distâncias, ângulos torcionais). Quanto mais restrições puderem ser obtidas, mais acurada será a família de estruturas, desde que não hajam erros de assinalamento. Essa acurácia atinge um valor assintótico a partir de 10 a 14 NOEs por resíduo. Além disso, é essencial a coerência entre modelo obtido e espectros medidos. Através de programas de cálculo retroativo, como o RELAX [81], são gerados espectros NOESY teoricamente esperados para o modelo. A análise desses espectros teóricos possibilita, além da identificação de assinalamentos incorretos, a introdução de NONOEs como restrições experimentais. A presença de um pico no espectro teórico e não no experimental mostra que, no modelo obtido, átomos se aproximaram de indevidamente. NONOEs representam distâncias interprotônicas maiores do que 4,5Å e quando introduzidas como restrições adicionais levam ao afastamento desses átomos.

Existem diversos pacotes de programas para determinação estrutural, baseados em protocolos de “simulated annealing”. O mais conhecido é o CNS - Crystallography and NMR System - sucessor do XPLOR, desenvolvido por Brunger *et al.* [82]. O CNS é utilizado tanto para a determinação estrutural de macromoléculas por espectroscopia de NMR quanto para o refinamento das estruturas obtidas por cristalografia de raios-X. O programa se baseia na minimização de uma função de energia aproximada: combinações arbitrárias de termos empíricos, geométricos e efetivos, descrevendo dados experimentais, podem ser utilizados. Outro pacote bastante útil é o DYANA – Dynamics Algorithm for NMR Applications – desenvolvido por Güntert *et al.* [83]. Baseado em dinâmica de ângulo torcional, possui como maior vantagem sobre o CNS uma convergência muito mais rápida das estruturas, em decorrência da diminuição dos graus de liberdade do sistema. Parâmetros como ângulos e tamanhos de ligações são fixos, restringindo o espaço conformacional, o que leva a uma excelente,

porém artificial, sobreposição das estruturas geradas. Estruturas calculadas podem ser facilmente visualizadas e analisadas com auxílio do programa MOLMOL [84].

2.3.1. Função para a Energia

Os termos para a energia potencial E_{pot} de uma molécula podem ser subdivididos em duas classes:

$$E_{pot} = E_{emp} + E_{exp} \quad . \quad (53)$$

O termo E_{emp} descreve a energia da molécula segundo uma equação empírica, função de suas coordenadas; o termo E_{exp} compreende os termos de energia provenientes de informações experimentais.

Energia empírica. A função de energia empírica contém termos de energia de interação conformacionais e não-ligantes envolvendo 2, 3 ou 4 átomos:

$$E_{emp} = \sum_{p=1}^N w_{bond}^p E_{bond} + w_{angle}^p E_{angle} + w_{dihe}^p E_{dihe} + w_{impr}^p E_{impr} + w_{vdw}^p E_{vdw} + w_{elec}^p E_{elec} \quad , \quad (54)$$

onde E_{bond} é a energia das ligações covalentes r , E_{angle} a energia dos ângulos θ de valências, E_{dihe} a energia dos ângulos ϕ torcionais, E_{impr} a energia dos ângulos torcionais rígidos φ (quiralidade, planaridade), E_{vdw} a energia da interação de van der Waals e E_{elec} a energia da interação eletrostática. Os diversos w^p são fatores de peso. Os primeiros 4 termos são conformacionais e descrevem a geometria da ligação covalente da molécula. São definidos por potenciais harmônicos, com exceção dos ângulos torcionais *dihe* que, por sua periodicidade, podem gerar vários mínimos de potencial, sendo melhor representados por uma função coseno:

$$\begin{aligned} E_{bond} &= \frac{1}{2} \sum_{bonds} k_b (r - r_0)^2 \\ E_{angle} &= \frac{1}{2} \sum_{angles} k_\theta (\theta - \theta_0)^2 \\ E_{dihe} &= \frac{1}{2} \sum_{torsion} k_\phi [1 + \cos(n\phi + \delta)], \quad n > 1 \\ E_{impr} &= \frac{1}{2} \sum_{impropers} k_\varphi (\varphi - \varphi_0)^2 \end{aligned} \quad . \quad (55)$$

As grandezas de equilíbrio r_0 , θ_0 e ϕ_0 foram determinadas por difração de raios-X; δ é um ângulo de fase. As constantes de força k são otimizadas de modo que a manter a estabilidade local do sistema. Os dois últimos termos da energia são não-ligantes. A energia de van der Waals entre dois átomos i e j separados por uma distância r_{ij} é descrita pelo potencial de Lennard-Jones:

$$E_{vdw}(r_{ij}) = \frac{1}{2} \sum_{pairs(ij)} k_{vdw} 4\varepsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right] , \quad (56)$$

onde ε é definida pela energia mínima ($E_{min} = -\varepsilon$) e σ pelo raio de van der Waals mínimo ($r_{min} = 2^{1/6} \sigma$). O termo de interação eletrostática é dado pelo potencial Coulombiano:

$$E_{elec}(r_{ij}) = - \sum_{pairs(ij)} k_{elec} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} , \quad (57)$$

sendo q a carga dos átomos e ϵ_0 a constante dielétrica do vácuo. Esses dois termos são, em geral, substituídos por um único termo simples e puramente repulsivo:

$$E_{vdw}(r_{ij}) = C_{rep} [\max(0, k^{rep} R_{min})^{irexp} - R^{irexp}]^{rexp} ; \quad E_{elec} = 0 \quad (58)$$

com R_{min} sendo o raio de van der Waals mínimo, $irexp$ e $rexp$ coeficientes ajustados.

Energia experimental. As restrições experimentais obtidas dos dados de NMR podem ser descritas pela função

$$E_{exp} = E_{noe} + E_{cdih} + E_{coup} + E_{shifts} , \quad (59)$$

representando respectivamente as restrições experimentais de distância (E_{noe}), de ângulos diedrais φ obtidos a partir da constante de acoplamento $^3J_{NH\alpha}$ (E_{cdih}), de valores para constantes de acoplamento J (E_{coup}) e de valores para os deslocamentos químicos comparados a um banco de dados estrutural (E_{shifts}). Para esses pseudo-potenciais existem, no CNS, diversas alternativas que podem ser escolhidas pelo usuário de acordo com a forma da molécula e da confiabilidade dos parâmetros estruturais. Os melhores resultados são, em geral, obtidos com os potenciais bi-harmônicos:

$$E_{noe}(r_{ij}) = \begin{cases} k_{noe} (r_{ij} - r_{max})^2 & r_{ij} > r_{max} \\ 0 & r_{min} < r_{ij} < r_{max} \\ k_{noe} (r_{ij} - r_{min})^2 & r_{ij} < r_{min} \end{cases} . \quad (60)$$

Os valores r_{max} e r_{min} são, respectivamente, o limite superior e inferior para as distâncias obtidas a partir dos NOEs. O parâmetro k_{noe} pode ser colocado como dependente da temperatura e dos limites dos erros. Os outros termos podem ser definidos analogamente.

2.3.2. Dinâmica Molecular

Em algoritmos de dinâmica molecular calcula-se a trajetória de uma molécula de N partículas segundo as equações clássicas de movimento. Pela equação de Newton, a partícula i de massa m_i e posição $\mathbf{r}_i$ estará sujeita, no tempo t , a uma força $\mathbf{F}_i$:

$$\mathbf{F}_i(t) = m \frac{\partial^2 \mathbf{r}_i(t)}{\partial t^2} = -\nabla_i E_{pot} \quad (61)$$

A solução numérica das equações de movimento é obtida computacionalmente a partir de uma escolha de incremento de tempo, Δt . O cálculo das coordenadas da partícula i no tempo $(t + \Delta t)$ é obtido das coordenadas nos dois tempos anteriores (t) e $(t - \Delta t)$, pela soma das duas séries de Taylor para $(t + \Delta t)$ e $(t - \Delta t)$ até terceira ordem:

$$\mathbf{r}_i(t + \Delta t) = 2\mathbf{r}_i(t) - \mathbf{r}_i(t - \Delta t) + \frac{\partial^2 \mathbf{r}_i(t)}{\partial t^2} (\Delta t)^2 \quad (62)$$

Esse método, denominado de Verlet (1967) leva, pela Série de Taylor, a uma invariância de posição e velocidade da partícula que pode, porém, ser desprezada nos intervalos de tempo característicos ($\Delta t = 10^{-15}$ s).

2.3.3. Minimização de Energia

O método de minimização de energia utiliza o método de Powell de gradiente conjugado, minimizando a energia total E_{tot} utilizando as coordenadas cartesianas como variáveis. Somente coordenadas de átomos livres (não-fixos) são modificadas durante o processo. Esse método costuma ser utilizado ao final de uma dinâmica molecular, a fim de regularizar a estrutura.

2.3.4. Protocolos de “simulated annealing”

O objetivo dos cálculos de dinâmica molecular é encontrar uma conformação da molécula que se aproxime do mínimo global de energia total. A fim de evitar que a conformação da molécula caia em

um mínimo local são criados protocolos nos quais a dinâmica molecular seja direcionada com diferentes fatores de escala para os diversos termos de energia. Um desses protocolos, denominado de “Simulated Annealing” (SA) [85-88] é bastante conveniente por abranger grande parte do espaço conformacional. Segundo esse protocolo, o sistema pode alcançar valores de energia superiores de acordo com a distribuição de Boltzmann:

$$P(E) = e^{\frac{-E}{k_B T}} . \quad (63)$$

A “temperatura efetiva” T do sistema é obtida do valor médio para a energia cinética

$$\langle E_c \rangle = \frac{1}{2} \sum_{i=1}^N m_i \langle v_i^2 \rangle = \frac{3}{2} N k_B T , \quad (64)$$

podendo assim ser utilizada como fator de escala para o direcionamento do processo de “resfriamento” do sistema. A “temperatura efetiva” T não simboliza propriamente uma temperatura física, e sim serve como parâmetro para determinar o tamanho da barreira de potencial a ser vencida. A energia cinética é lentamente diminuída durante os cálculos através da redução da velocidade v_i com a temperatura. Para tal, as equações de movimento são acrescidas de um fator de amortecimento, a fim de evitar oscilações de alta frequência dos átomos que poderiam levar a uma dissociação da molécula:

$$\mathbf{F}_i(t) = m \frac{\partial^2 \mathbf{r}_i(t)}{\partial t^2} = -\nabla_i E_{pot} - \gamma_i m_i \mathbf{v}_i \left(1 - \frac{T}{T_{curr}}\right) . \quad (65)$$

A velocidade será, portanto, alterada de acordo com a temperatura desejada T , a temperatura atual T_{curr} e do coeficiente de fricção γ . Paralelamente ao resfriamento do sistema são aumentados gradativamente os pesos dos termos de energia experimentais, de modo a restringir as estruturas calculadas ao espaço conformacional que mais se adapte às restrições por NMR.

Como regra utiliza-se esse procedimento sobre diversas estruturas iniciais geradas randomicamente, que deverão, ao final da dinâmica, convergir a uma só estrutura. Para isso é necessário que hajam suficientes restrições experimentais e que estas sejam, em sua maioria, corretas. Parâmetros para avaliar a qualidade da estrutura são as energias e seus desvios e o número de violações de restrições experimentais. As subunidades flexíveis da molécula serão vistas, por NMR, em uma conformação média-temporal que em geral não corresponde a um mínimo de energia, não convergindo a uma única conformação final.

2.4. ANÁLISE DAS ESTRUTURAS OBTIDAS

Uma vez obtida uma família de estruturas por NMR em solução, surge a questão de como avaliar sua qualidade e sua validade. Em cristalografia, por exemplo, parâmetros como resolução e fator-R fornecem uma medida da acurácia a ser esperada para o modelo correspondente, enquanto que fatores-B atômicos e ocupação podem dar indicações das regiões mais ordenadas e desordenadas na estrutura. Em NMR não existe um parâmetro simples para a avaliação da qualidade das estruturas obtidas. A precisão de uma estrutura resolvida por NMR é, em geral, expressa em função do desvio quadrático médio (RMSD – “root-mean-square deviation”) das coordenadas pareadas ao longo dos modelos finais, ou relativo a um modelo médio calculado. Esse desvio quadrático médio pode ser calculado sobre toda a molécula ou sobre os segmentos que possuem conformações bem definidas. A grande maioria das estruturas publicadas utilizam esse critério como sendo o mais importante, trazendo, em geral, resultados muito precisos porém não necessariamente acurados. Valores RMSD podem ser artificialmente baixos, dependendo dos diversos parâmetros utilizados durante os passos de refinamento, além do fato de não existir um critério bem definido na seleção das melhores estruturas a serem incluídas entre os modelos finais.

Os programas AQUA – Analyzing the QUALity – e PROCHECK-NMR foram desenvolvidos por Laskowski *et al.* [89], visando uma forma de validar estruturas geradas por NMR e desenvolver uma sistemática nos critérios de análise de estruturas publicadas. Os programas são capazes de calcular o nível de completude de NOEs baseado na estrutura obtida (AQUA) e analisar o grau de coerência entre estruturas geradas e restrições experimentais, além da qualidade de suas propriedades geométricas (PROCHECK-NMR).

Restrições experimentais. Um dos mais importantes indicativos da qualidade de estruturas determinadas por NMR é a coerência dos modelos finais com as restrições experimentais, dada em geral na forma de número e tamanho de restrições violadas. Um fator importante a ser considerado é a indeterminação na relação entre intensidade do NOE e distância atribuída. Erros randômicos e difusão de spin podem, em parte, ser corrigidos por procedimentos de inversão da matriz de relaxação. Outras importantes fontes de indeterminação são erros sistemáticos introduzidos por diferentes seqüências e formas de pulso, e principalmente o desconhecimento da função de densidade espectral $J(\omega)$, presente na relação

$$NOE \approx \frac{const}{r^6} J(\omega) . \quad (66)$$

Em geral assume-se que a proteína é rígida e descrita por um único tempo de correlação, porém já foi demonstrado que os erros nas distâncias, introduzidos por desprezar-se os movimentos internos na molécula, podem ser superiores a 50% [90]. Torna-se, portanto, necessária a introdução de intervalos razoáveis de erro nas distâncias obtidas. Esses intervalos, mesmo que menores que 1 Å, farão com que a acurácia das estruturas calculadas não seja superior a 1 Å de resolução; valores melhores que esse serão fictícios, provenientes de artefatos na análise de NOEs e/ou cálculo estrutural [91]. Restrições nos ângulos torcionais obtidos de constantes de acoplamento e informações de deslocamentos químicos são bastante efetivas em fornecer estruturas mais precisas. Essas restrições são, de uma forma geral, associadas a valores típicos em estruturas secundárias. Deve-se, porém, ter o cuidado de não restringir muito o intervalo de erro desses ângulos ($<30^\circ$), o que levaria a estruturas “precisas”, adotando conformações secundárias típicas, porém não necessariamente corretas.

Propriedades geométricas. A qualidade estereoquímica de uma proteína fornece outra medida da qualidade do modelo, sendo, para certos parâmetros, independente dos procedimentos de refinamento utilizados. São analisadas propriedades geométricas da proteína (comprimentos e ângulos de ligação, ângulos diedrais e impróprios, etc.), com base em dados obtidos por comparação das várias estruturas resolvidas por raios-X, de alta resolução, depositadas no PDB. O exemplo mais comum é o gráfico de Ramachandran [92], que define quais combinações dos ângulos diedrais ϕ, ψ são favoráveis, desfavoráveis e proibidas. Em estruturas cristalográfica de mais baixa resolução é encontrado um maior número de combinações ϕ, ψ desfavoráveis e proibidas. A porcentagem de resíduos nas regiões mais favoráveis pode ser graficada como função da resolução, fornecendo uma curva típica para as estruturas de raios-X depositadas. Dessa forma, a partir da quantidade de resíduos com ϕ, ψ favoráveis nas estruturas de NMR obtidas pode ser encontrada uma *resolução equivalente* para o modelo. De forma análoga são obtidas resoluções equivalentes para as outras propriedades geométricas.

Estimativas da precisão local da família de estrutura, em geral dadas em termos de RMSD, podem ser muito melhor caracterizadas através da *variância circular* (CV – circular variance) em conjunto com *acessibilidade ao solvente*. Enquanto que o RMSD é calculado em função das coordenadas atômicas, a CV é calculada com base no desvio angular. Dado um ângulo diedral θ , sua variância é definida por

$$Var(\theta) = 1 - R_{av} \quad , \quad (67)$$

sendo $R_{av}=R/n$, R definido por

$$R^2 = \left(\sum_{i=1}^n \cos \theta_i \right)^2 + \left(\sum_{i=1}^n \sin \theta_i \right)^2 \quad (68)$$

e n sendo o número de membros na família de estruturas. O valor da CV varia entre 0 e 1, sendo que valores baixos indicam maior proximidade em torno de um único valor θ . Em uma região flexível, como por exemplo em um loop de alta acessibilidade ao solvente, valores de CV altos seriam esperados. Outro parâmetro útil para a análise da qualidade das estruturas obtidas é o *fator-G*, que fornece uma medida da “normalidade” para uma dada propriedade estereoquímica. O *fator-G* é essencialmente um “log-odds score” baseado na distribuição observada para uma determinada propriedade em estruturas de alta resolução, obtidas por raios-X. O banco de dados utiliza 163 cadeias de estruturas resolvidas por raios-X, de resolução 2.0Å ou melhor, *fator-R* não maior que 20% e homologia não superior a 35%; todos os átomos de ocupação zero são excluídos. Quando aplicado a um dado resíduo, um *fator-G* baixo indica uma conformação de baixa probabilidade. Assim, um resíduo poderá ter uma CV quase nula, indicando alta população em uma determinada conformação (e portanto grande “precisão”), mas um *fator-G* alto, indicando ser essa conformação bastante desfavorável e portanto pouco provável. Esse seria um forte indício de assinalamento incorreto de NOEs e/ou ângulos diedrais.

3. Proteína HPr(H15A) de *S. aureus*

3.1. MATERIAIS E MÉTODOS

Como parte principal do presente trabalho de doutorado foi resolvida a estrutura tridimensional da proteína HPr de *Staphylococcus aureus*, com mutação His15→Ala15 em seu sítio ativo – HPr(H15A) –, e comparada com a estrutura da proteína nativa – HPr(wt). Todos os dados experimentais foram obtidos na Alemanha, em um trabalho-conjunto com o Prof. Dr. Hans-Robert Kalbitzer, do Instituto de Biofísica e Bioquímica Física, Universidade de Regensburg, e com o Prof. Dr. Wolfgang Hengstenberg, Departamento de Biologia da Universidade de Bochum-Ruhr.

3.1.1. Preparação das Amostras

A expressão e a purificação das proteínas HPr(H15A) e HPr(wt) de *Staphylococcus aureus* foi realizada pelo Prof. Dr. Wolfgang Hengstenberg, de forma análoga à descrita para *Staphylococcus carnosus* [24]. Aproximadamente 75mg de proteína não-enriquecida e enriquecida com ^{13}C - ^{15}N foram colocadas à disposição. Para as amostras de HPr(H15A) utilizadas na determinação estrutural foram dissolvidas ~15 mg de proteína em 0,5 ml de uma solução aquosa contendo DSS 0,05mM (referência), EDTA 0,5mM (para prevenir proteases cálcio-dependentes), NaN_3 1mM (para prevenir fungos) e 5% de D_2O (para o sinal de “lock”). Após a dissolução o pH foi ajustado em 7,0 através da adição de alguns microlitros de NaOH e/ou HCl. As medidas foram realizadas a uma temperatura de 303K.

Para os estudos comparativos entre HPr(wt) e HPr(H15A) foram preparadas simultaneamente amostras em condições idênticas de concentração, solvente, temperatura e pH.

3.1.2. Espectros de NMR

Os espectros de NMR foram obtidos em espectrômetros Bruker DMX-500, DRX-600 e DMX 800, operando respectivamente em frequências de próton de 500, 600 e 800 MHz. O sinal da água foi suprimido através de pré-saturação seletiva; todos os espectros foram coletados no modo sensível à fase usando TPPI. Para calibração dos deslocamentos químicos

^1H utilizou-se DSS como referência interna. O software XWINNMR (Bruker, Karlsruhe, Alemanha) foi utilizado para aquisição e processados dos dados, em estações Silicon Graphics. A seguir são descritos os espectros utilizados para a análise e seus parâmetros de aquisição e processamento.

Figura 13. Espectro TOCSY da proteína HPr(H15A) de *Staphylococcus aureus*, em água, obtido no espectrômetro Bruker de 800 MHz. O tempo de mistura foi de 60 ms. T= 303K; pH=7,0.

TOCSY e NOESY. Os espectros TOCSY e NOESY bidimensionais foram adquiridos com 4K×1K pontos de dados no domínio temporal, no espectrômetro de 800 MHz. Para o TOCSY o tempo de trava de spin foi de 60ms, utilizando a sequência MLEV-17. No NOESY o tempo de

mistura foi de 100ms. O processamento foi realizado com $2K \times 2K$ pontos de dados no domínio da frequência, utilizando-se filtros de multiplicação Gaussiana em ambas as dimensões ($LB=-6\text{Hz}$ e $GB=0.12$ em t_2 ; $LB=-8\text{Hz}$ e $GB=0.1$ em t_1). Os espectros TOCSY e NOESY da proteína HPr(H15A) podem ser vistos nas figuras 13 e 14.

HPr(H15A) - *S. aureus* - NOESY (800 MHz, pH7, 303K)

Figura 14. Espectro NOESY da proteína HPr(H15A) de *Staphylococcus aureus*, em água, obtido no espectrômetro de 800 MHz. O tempo de mistura foi de 100 ms. T=303K; pH=7,0.

COSY. O espectro COSY, variante DQF, foi adquirido com $16K \times 1K$ pontos de dados no domínio temporal, no espectrômetro de 800 MHz. O processamento foi realizado com $4K \times 2K$,

utilizando-se filtros idênticos aos do TOCSY e NOESY (figura 15). Para a análise do acoplamento J foi realizada uma “strip transform” da região NH-H α , com 16K $\times$ 1K e um filtro de multiplicação exponencial em t_2 (LB=0.1) e senoidal (SSB=0) em t_1 .

Figura 15. Espectro DQF-COSY da proteína HPr(H15A) de *Staphylococcus aureus*, em água, obtido no espectrômetro Bruker de 800 MHz. T= 303K; pH=7,0.

HSQC. O espectro 2D heteronuclear ^1H - ^{15}N HSQC foi adquirido com gradientes selecionados usando a abordagem eco-antieco, com um total de 8K $\times$ 512 pontos de dados; o desacoplamento ^{13}C foi

efetuado com a sequência GARP (figura 16). Procedimento análogo foi realizado para o espectro 2D ^1H - ^{13}C HSQC. Os dados foram processados a $2\text{K} \times 2\text{K}$, com filtros tipo cosseno-quadrado (SSB=2).

Figura 16. Espectro ^{15}N -HSQC da proteína HPr(H15A) de *S. aureus*, em água, obtido no espectrômetro Bruker de 800 MHz. T= 303K; pH=7,0.

HNCA e HNCO. Os experimentos 3D heteronucleares HNCA e HNCO foram realizados no espectrômetro de 500 MHz. Foram adquiridos $2\text{K} \times 64 \times 128$ pontos de dados no domínio

temporal, os quais foram processados a $2K \times 256 \times 256$, com predição linear em t_2 e t_1 e filtros cosseno-quadrado (SSB=2).

NOESY-HSQC. Os espectros 1H - 1H - ^{15}N NOESY-HSQC e 1H - 1H - ^{13}C NOESY-HSQC foram adquiridos com um total de $2K \times 64 \times 128$ pontos de dados; o tempo de mistura foi de 100ms. Os dados foram processados a $2K \times 256 \times 256$ com predição linear em t_2 e t_1 e filtros cosseno-quadrado (SSB=2).

3.1.3. Estratégia de Assinalamento

O assinalamento 1H dos sistemas de spin foi realizado conforme descrito em 2.2.1, através dos espectros 2D TOCSY e COSY e com uso do programa AURELIA [73] para visualização. Com o tempo de mistura de 60 ms para o experimento TOCSY, a maioria dos sistemas de spin se mostrou completa, facilitando sua identificação. O assinalamento sequencial foi inicialmente realizado através do espectro 2D NOESY (região “fingerprint”) e depois confirmado via 3D HNCA. Os deslocamentos químicos ^{13}C e ^{15}N foram obtidos dos espectros 2D heteronucleares ^{13}C -HSQC e ^{15}N -HSQC; as carbonilas ^{13}C foram assinaladas pelo espectro 3D HNCO.

3.1.4. Restrições Experimentais

Como restrições experimentais foram consideradas as constantes de acoplamento $^3J_{NH\alpha}$ e as distâncias interprotônicas.

Constantes de acoplamento. As constantes de acoplamento foram obtidas por simulação de doublets de linhas Lorentzianas no espectro COSY transformado (região NH-H α , “strip-transform”). O processamento com utilização de um filtro de multiplicação exponencial com baixo LB em t_2 garante a ausência de efeitos cosméticos na forma de linha, e portanto de artefatos no valor da constante de acoplamento. Aos valores encontrados foram incorporados limites de erro superior e inferior de 0,5Hz.

Distâncias interprotônicas. Para a quantificação dos picos NOEs dos espectros NOESY foi realizada uma busca automática de picos (“peak-picking”) dentro do programa AURELIA, resultando em ~3700 NOEs a serem analisados. A integração dos NOEs se deu em uma caixa de integração de tamanho $S_1=40Hz$ e $S_2=40 Hz$, com número de passos de segmentação $N_s=8$ e até um nível relativo de integração $a_1=0,1$. Os volumes foram transformados em distâncias pela equação (50), considerando-se um tempo de correlação rotacional constante; para calibração foram usados os 6 picos aromáticos dos

contatos $H\delta-H\epsilon$ das 3 Tirosinas existentes. Às distâncias foram incorporados intervalos de erros superior e inferior de 20%, exceto nos seguintes casos:

- para cada prótom metilênico degenerado (HB^* , HG^* , etc.) eram adicionados 10% de erro ao limite superior, visto que esses átomos são substituídos por um pseudo-átomo no protocolo CNS;
- para cada próton metílico envolvido o erro superior era aumentado em 20%, também devido à criação do pseudo-átomo;
- aos prótons em grupos sec-propil degenerados (Valinas, Leucinas) eram adicionados 40% no erro superior (pseudo-átomos);
- aos prótons metilênicos não degenerado e sem identificação estereoespecífica, os limites superior e inferior eram ajustado de tal forma que ambos os intervalos fossem idênticos, possibilitando uma troca entre os prótons;
- no caso de um próton metilênico não degenerado e não identificado estereoespecificamente, cujo adjacente não fornecesse um pico no espectro NOESY, ao limite superior eram adicionados 1,8Å (distância máxima entre prótons metilênicos adjacentes);
- para os prótons equivalentes em anéis, degenerados, o limite superior era aumentado em 30% (pseudo-átomo);
- para os prótons equivalentes em anéis, não-degenerados, ao limite superior eram adicionados 2,5Å (a fim de possibilitar a troca entre eles);

Os NOEs foram assinalados de forma interativa, alterando-se com cálculos de estruturas. Partiu-se de uma cadeia linear, assinalando-se primeiramente os NOEs envolvidos em hélices. Evidências de três hélices podiam ser observadas na análise dos deslocamentos químicos com relação aos valores-padrão, onde os valores sistematicamente positivos indicavam uma estrutura helicoidal. Estruturas foram calculadas utilizando-se o programa DYANA e analisadas com o programa MOLMOL, NOEs incompatíveis foram retirados. Baseado nas distâncias inter-protônicas esperadas para as hélices geradas, novos NOEs puderam ser encontrados. Uma parte significativa da ambigüidade dos NOEs pôde ser removida dessa forma. Uma vez bem moldadas as 3 hélices iniciou-se a geração da estrutura terciária pela introdução de restrições de distância de longo alcance. Foram primeiramente introduzidos somente NOEs univocamente assinalados nos espectros 2D NOESY ou, no caso de ambigüidades, quando essas pudessem se resolvidas por inspeção dos espectros 3D NOESY. Obtidas estruturas enoveladas, novas dubiedades puderam ser eliminadas e novos NOEs introduzidos.

3.1.5. Cálculo de Estruturas

O programa DYANA foi utilizado até que uma alta convergência de estruturas foi obtida. Partiu-se, então, para os protocolos de SA, utilizando-se o programa CNS. Foram geradas 1000 estruturas, contando com 2142 restrições de distância e 71 restrições de ângulos torcionais. O protocolo partiu de 50000K, em uma fase quente de dinâmica de ângulos torcionais, sendo resfriado até 0K em uma fase fria de dinâmica cartesiana. Das 1000 estruturas geradas foram selecionadas as melhores 125, todas apresentando os 7 elementos de estrutura secundária esperados. As estruturas foram analisadas e, em regiões em que ocorriam grandes divergências, os NOEs foram novamente confirmados e NONOEs introduzidos. NONOEs são restrições experimentais com distância entre 4.5 e 40Å, provenientes da não-existência de NOEs nos espectros. Eles impedem que prótons, entre os quais não existe um NOE no espectro, se aproximem no cálculo da estrutura. Nova dinâmica foi realizada sobre as 125 estruturas (protocolo REFINE), desta vez apenas em fase fria de dinâmica cartesiana, partindo de 300K até 0K. A tabela 7 traz os parâmetros utilizados nos protocolos.

As 40 estruturas de mais baixa energia foram submetidas aos programas de análise de qualidade, AQUA e PROCHECK-NMR. Através da análise das figuras de Ramachandram aminoácido por aminoácido, puderam ser encontrados os modelos estereoquimicamente incoerentes, possibilitando a busca de novos NONOEs que eliminassem a convergência de estruturas para tais situações de mínimo local.

Tabela 7. Parâmetros utilizados nos protocolos de SA e refinamento. As constantes k referem-se às constantes de força nas respectivas energias.

	Protocolo SA			Protocolo REFINE	
	fase quente	resfriamento	Minimização	resfriamento	minimização
T (K)	50000 → 1000	1000 → 0		300 → 0	
N	1000	1000	2000	1000	4000
k_{bond}	1.0	1.0	1.0	1.0	1.0
k_{ang}	0.4	0.4	1.0	0.4	1.0
k_{imp}	1.0	1.0	1.0	1.0	1.0
k_{vdw}	0.1	0.1 → 4.0	1.0	0.1 → 4.0	1.0
k_{noe}	150	150	150	150	75
k_{coup}	1.0	1.0	1.0	1.0	1.0
Repel	0.9 → 0.8	0.9 → 0.8		0.9 → 0.8	
Asymptote	0.1, 1.0	1.0		1.0	

3.2. RESULTADOS E DISCUSSÃO

As 40 estruturas finais obtidas apresentam excelente geometria. A estrutura da proteína HPr (H15A) de *Staphylococcus aureus* mantém o padrão fita-hélice-fita-fita-hélice-fita-hélice encontradas em proteínas HPr de outros organismos, com mesmo enovelamento.

3.2.1. Assinalamento Seqüencial e Estrutura Secundária

O apêndice A traz os deslocamentos químicos ^1H , ^{13}C e ^{15}N em HPr(H15A) de *Staphylococcus aureus*, relativos ao DSS. Dada a excelente qualidade dos espectros, na maioria dos grupos metílenos não degenerados pôde-se efetuar o assinalamento estereoespecífico (apêndice A). A análise de estrutura secundária, através da diferença entre os deslocamentos químicos $\text{H}\alpha$, $\text{C}\alpha$ e CO em relação aos valores “random-coil”, é mostrada nas figuras 17 a 19.

Nas figuras podem ser claramente identificadas as regiões de estrutura secundária regular. Os limites das estruturas não fica perfeitamente definido, porém, de uma forma geral, pode-se esperar os seguintes elementos: $\beta 1$ entre resíduos 1 e 9, $\alpha 1$ entre resíduos 16 e 31, $\beta 2$ entre resíduos 33 e 37, $\beta 3$ entre resíduos 39 e 42, $\alpha 2$ entre resíduos 47 e 55, $\beta 4$ entre resíduos 60 e 67, $\alpha 3$ entre resíduos 70 e 84.

Figura 17. Análise de estrutura secundária: diferenças dos deslocamentos químicos $\text{H}\alpha$ em relação aos padrões Gly-Gly-X-Ala-Gly-Gly e Gly-Gly-X-Pro-Gly-Gly [71].

Figura 18. Análise de estrutura secundária: diferenças dos deslocamentos químicos $C\alpha$ em relação aos padrões Gly-Gly-X-Ala-Gly-Gly e Gly-Gly-X-Pro-Gly-Gly [71].

Figura 19. Análise de estrutura secundária: diferenças dos deslocamentos químicos CO em relação aos padrões Gly-Gly-X-Ala-Gly-Gly e Gly-Gly-X-Pro-Gly-Gly [71].

Os elementos de estrutura secundária previstos pelos deslocamentos químicos podem ser confirmado no mapa de contatos de NOEs sequenciais e de constantes de acoplamento $^3J_{NH-H\alpha}$ (figura 20). O mapa mostra uma representação esquemática dos NOEs entre prótons de aminoácidos nas posições i e j na sequência, e das constantes de acoplamento J encontradas. A folha- β formada pelas 4

fitas- β identificadas por diferença de deslocamentos químicos pode ser melhor caracterizada através de um diagrama de contatos de longo alcance, conforme ilustrado na figura 21. São mostradas apenas as distâncias $d < 3\text{\AA}$, através de setas entre os prótons envolvidos.

Figura 20. Mapa de contatos sequenciais de HPr(H15A). Os NOEs são divididos em fortes ($d \leq 2,8\text{\AA}$), médios ($2,8\text{\AA} < d \leq 3,3\text{\AA}$) e fracos ($d > 3,3\text{\AA}$), refletindo na espessura da linha. As constantes de acoplamento J são divididas em fortes ($J < 6\text{Hz}$), representadas por círculos preenchidos, e fracas ($J > 8\text{Hz}$), representadas por círculos vazios; valores intermediários são indicados por um X.

Figura 21. Contatos entre as 4 fitas- β identificadas por análise dos deslocamentos químicos. As setas indicam distâncias inter-protônicas $d < 3\text{\AA}$, características para a folha- β formada. A ausência de setas resulta da ambiguidade de NOEs em decorrência de sobreposição de picos no espectro.

3.2.2. Estrutura Tridimensional

A figura 22 mostra a estrutura obtida para HPr(H15A) de *S. aureus*. Em (a) estão sobrepostas as 40 melhores estruturas, na forma de um cilindro cujo raio corresponde ao RMSD da cadeia principal; (b) traz a estrutura média na forma de “ribbons”. Os valores de RMSD dos átomos da cadeia principal variam desde um mínimo de 0.24Å até um máximo de 0.84Å, indicando a boa sobreposição dos modelos, sugerindo a possibilidade de que a mutante HPr(H15A) adote uma única conformação, em contraste com outras proteínas HPr já publicadas.

Figura 22. Estrutura de HPr(H15A). (a) Sobreposição das 40 melhores estruturas, na forma de um cilindro cujo raio corresponde ao RMSD da cadeia principal; (b) representa uma estrutura média.

As 40 melhores estruturas foram submetidas ao programa PROCHECK-NMR. A figura 23 traz o gráfico de Ramachandran [92] do conjunto, estando 91.4% dos resíduos em regiões mais favoráveis (vermelho - A, B, L), e 8.6% nas regiões permitidas (amarelo escuro - a, b, l, p). As resoluções equivalentes são mostradas na figura 24. Dada a grande ambigüidade no assinalamento de cadeias laterais, somada à grande sobreposição de picos nos espectros NOESY, a resolução equivalente em χ_1 e χ_2 é baixa.

Figura 23. Gráfico de Ramachandran das 40 melhores estruturas de HPr(H15A). Os números nos quadrados e triângulos indicam o modelo considerado. Triângulos se referem às glicinas. Os resíduos em regiões não-permitidas estão indicados em vermelho. As regiões do gráfico coloridas em vermelho, amarelo-escuro e amarelo-pálido indicam, respectivamente, regiões mais favoráveis, permitidas e menos favorecidas. Regiões brancas são proibidas.

Um resumo da geometria do conjunto é mostrado na figura 25. O primeiro gráfico (25a) traz o desvio do ângulo χ_1 em relação ao valor médio encontrado para os 40 modelos. A dispersão indica flexibilidade nas cadeias laterais, que pode ser proveniente de uma flexibilidade real da molécula ou de baixa quantidade de restrições experimentais em decorrência de sobreposições no espectro ou falta de assinalamento estereo-específico. Na figura 25b vemos a boa planaridade das ligações peptídicas, dada a baixa dispersão do ângulo torcional ω . Em 25c estão mostrados os RMSD para cada aminoácido em relação ao valor médio, que se mantém basicamente constantes ao longo da cadeia. Não há indícios de regiões mal definidas ou muito flexíveis. A figura 25d traz uma definição da estrutura secundária, mediadas ao longo dos 40 modelos. Estão definidos os 7 elementos de estrutura secundária: as 4 fitas- β (β_1 , β_2 , β_3 e β_4) compreendem os resíduos 2-7, 34-37, 40-43 e 60-65 respectivamente; as 3 hélices- α (α_1 , α_2 , α_3) compreendem os resíduos 16-27, 47-51 e 70-84 respectivamente. Esse padrão é comum ao longo das diversas estruturas de HPr já publicadas. A acessibilidade ao solvente está

esquematisada na figura 25e. Regiões mais acessíveis correspondem a “loops” na proteína, enquanto que elementos de estrutura secundário são mais enterrados. A variância circular (CV) e o fator-G são mostrados respectivamente em 25f e 25g. Quanto mais alto a CV, mais escuro o disco; quanto mais escuro o quadrado representando o fator-G, menos usual a conformação. Para estruturas bem definidas e pouco flexíveis espera-se baixos valores de CV ao longo da cadeia principal, portanto em ϕ e ψ , o que pode ser observado na figura. Em conjunto com a análise dos fatores-G pode-se avaliar a qualidade do modelo. Em alguns resíduos a CV é bastante baixa, porém o fator-G indica uma conformação não muito usual (por ex. nos resíduos 8 e 9), o que pode ser indicativo de algum erro nas restrições experimentais. Nas cadeias laterais (e portanto em χ_1 e χ_2) as variâncias circulares encontradas podem ser maiores, reflexo dos diferentes rotâmeros (g^- , t e g^+) possíveis. Novamente, a qualidade da conformação global pode ser analisada através do fator-G.

HECK-NMR

Equivalent resolution

ref_HPr_best40 (40 models)

Figura 24. Resoluções equivalentes dos modelos obtidos: **(a)** em relação à qualidade do gráfico de Ramachandran; **(b)** qualidade das pontes de hidrogênio; **(c)** ângulo torcional χ_1 ; **(d)** ângulo torcional χ_2 ; **(e)** fator-G.

Residue properties ref_HPr_best40 (40 models)

Figura 25. Geometria do conjunto: (a) desvio do ângulo χ_1 em relação ao valor médio encontrado (linha horizontal); (b) dispersão do ângulo torcional α (c) RMSD para cada aminoácido; (d) previsão da estrutura secundária; (e) acessibilidade ao solvente; (f) variância circular; (g) fator-G;

Descrição da estrutura

A proteína HPr(H15A) de *Staphylococcus aureus* manteve o padrão estrutural observado nas outras espécies: um sanduíche-aberto, formado de três hélices- α empacotadas contra uma folha composta por quatro fitas- β anti-paralelas (figura 26). As 3 hélices estão alinhadas paralela- ou antiparalelamente com as fitas, conforme esperado [93]. A folha- β formada pelas 4 fitas possui

a topologia 2x, 1, -2x, com um “twist” de mão esquerda (quando visto perpendicularmente às fitas).

O centro (“core”) hidrofóbico mostra uma total ausência de resíduos aromáticos, sendo composto exclusivamente de uma quantidade marcante de isoleucinas e leucinas, com algumas alaninas e valinas. Esse centro forma um funil, mais largo na região inferior onde se encontra o sítio ativo; na superior, mais estreita, é encontrada uma maior concentração de alaninas. As 4 metioninas da proteína apontam para fora, o que é pouco comum dada a característica hidrofóbica desse resíduo. Espera-se, baseado na estrutura do complexo entre HPr de *E. coli* e a Enzima I [19], que alguns destes resíduos (particularmente Met48 e Met51) estejam envolvidos na interface de interação entre as duas moléculas. A única fenilalanina, aromático hidrofóbico, não aponta para o centro (figura 26).

Figura 26. Representação em forma de “ribbon” da estrutura média encontrada para a proteína HPr(H15A) de *Staphylococcus aureus*. São indicadas as 4 metioninas e a fenilalanina.

Figura 27. Sítio ativo de HPr(H15A). Em (a) estão sobrepostas as 40 melhores estruturas; (b) representa uma estrutura média.

A região Asp10-Pro18, que na proteína nativa constitui o sítio ativo para a transferência do grupo fosfato entre as enzimas EI e EIIA, forma uma espécie de “tampa” na parte mais larga do funil hidrofóbico central; a Ile14 aponta para dentro do funil, agindo como uma âncora hidrofóbica vinculando a alça ao centro hidrofóbico. Os resíduos Asp10-Thr12 formam uma volta Asx, estabilizada por uma ligação de hidrogênio entre o oxigênio O_δ da Asp10 e a amida da Thr12. Entre a amida da Asp10 e a carbonila da Leu86 forma-se uma fraca ligação de hidrogênio. Em várias das estruturas calculadas a metila da Thr12 interage com a metila da Ala15 através de um contato van der Waals; essa interação não aparece nas outras espécies ou na proteína nativa, e talvez resulte na alteração dos ângulos ϕ, ψ da treonina, observada na estrutura da mutante. A Gly13 apresenta uma conformação não permitida para os demais aminoácidos, caindo na região ϵ do diagrama de Ramachandran nas melhores estruturas. Sua conservação em HPr de diversas espécies é consequência de seu papel estrutural: dobrar a cadeia, iniciando o “loop” do sítio ativo. A ausência de quaisquer contatos NOE entre a cadeia lateral da Arg17 e outros resíduos reflete em uma pobre sobreposição nas estruturas obtidas: o resíduo aponta para fora da proteína, estando, possivelmente, desordenado, como visto em outras estruturas de HPr determinadas por NMR. A Ala16 inicia a primeira hélice da proteína. Na estrutura de HPr de *Streptococcus faecalis* determinada por difração de raios-X, este resíduo demonstrou “strain” conformacional, com um ângulo $\phi = 59^\circ$ [30]. Os autores concluíram que a estrutura não-fosforilada estaria numa conformação tensionada. A estrutura aqui apresentada não demonstra nenhuma evidência de

tensão na Ala16, sendo que nas melhores estruturas o resíduo adota uma conformação dentro da região de hélice- α do diagrama de Ramachandran. Este resultado é compatível com estruturas mais recentes da HPrs determinadas tanto por NMR quanto difração de Raios-X, sugerindo que não existe uma conformação de alta energia para o “loop” do sítio ativo, relevante ao mecanismo de transferência de fosfato.

Estrutura secundária

A seguir são descritos os elementos de estrutura secundária encontrados.

Hélice $\alpha 1$ (Ala16-Ala26). Essa primeira hélice se inicia após o resíduo Ala15, que na estrutura nativa é a histidina fosforilável do sítio ativo. Nas diversas estruturas publicadas para HPr de outros organismos, a His15 é referida como sendo o “N-cap” da primeira hélice (definido como posição i). Sua estrutura é, porém, completamente incomum para formar um “N-cap”: sua cadeia lateral empacota perpendicularmente ao eixo da hélice, não formando ligação de hidrogênio com as amidas dos resíduos (i+2) e (i+3). Uma prolina (Pro18) na posição (i+2) é incomum em hélices- α de uma forma geral, mas altamente conservada em HPr. O anel da sua cadeia lateral empacota perpendicularmente ao grupo imidazol da His15 e sua conservação enfatiza a importância desta interação. A fosforilação da histidina, porém, disponibiliza oxigênios do grupo fosfato para formar a ligação de hidrogênio com as amidas da Ala16 e da Arg17. Na proteína mutante a carbonila da Arg17 forma uma ligação de hidrogênio fraca com a amida da Met21, em contraste com as estruturas de outros organismos. O aumento no comprimento desta ligação de hidrogênio, que é característica de uma hélice- α , indica que a primeira volta da hélice ligeiramente se desfaz. Esta observação foi confirmada pelos estudos comparativos entre a proteína mutante e a nativa, descritos mais adiante.

Hélice $\alpha 2$ (Ile47-Gly54). Uma serina (Ser46) na posição “N-cap” é comum, mas sua cadeia lateral é curta para realizar ligações de hidrogênio fortes com Met48 (i+2) e Gly49 (i+3). Sua fosforilação, porém, que acontece durante a regulação em organismos Gram positivos, deverá disponibilizar oxigênios para completar tais ligações. A distância entre a carbonila da Gly49 e a amida da Leu53 é muito longa para a existência de uma ligação de hidrogênio. Tal fato decorre da presença da Ser52, cuja hidroxila forma a ligação com a Gly49, assim enfraquecendo a ligação de hidrogênio normal da hélice. Em continuidade são formadas 2 ligações de hidrogênio do tipo hélice 3_{10} (entre Val50 e Leu53 e entre Met51 e Gly54) e a hélice termina com a Gly54 em uma conformação α_L (hélice de mão esquerda).

Hélice $\alpha 3$ (Glu70-Gly85). Possui um glutamato na posição “N-cap”, mas sem formar ligações de hidrogênio com as amidas da primeira volta. Entre Gly67 e Glu70 forma-se uma volta- β tipo I, com uma ligação de hidrogênio entre esses resíduos, satisfazendo a amida do resíduo Glu70. No final da hélice observa-se uma conformação interessante, conhecida por “Shellman box” [94]. Nessas estruturas as duas últimas carbonilas da hélice trocam de parceiros na formação de ligações de hidrogênio: Leu81 com Leu86, e Ser82 com Gly85. O motivo é estabilizado por contato van der Waals entre as cadeias laterais hifrofóbicas de Leu81 e Leu86 (figura 28). Além dessas interações, que definem a “Shellman box”, o final da hélice é estabilizado por uma ligação de hidrogênio (bifurcada) entre a carbonila da Leu81 e a amida de Thr87, e através dos oxigênios $O\gamma$ das cadeias laterais de Thr87 e Ser82. Essa região da molécula (Leu81-Leu86) fica espacialmente próxima ao “loop” do sítio ativo: entre a carbonila da Leu86 e a amida da Asp10 existe uma ligação de hidrogênio, e a cadeia lateral da Leu86 forma contatos hidrofóbicos com o “loop”.

Figura 28. “Shellman box” (Leu81-Thr87) e sua interação com o início do “loop” do sítio ativo Ile8-Asp10). São mostradas as amidas e carbonilas; para os resíduos Ile8, Leu81, Ser82, Leu86 e Thr87 são mostrados todos os átomos. Ligações de hidrogênio características do “Shellman box” estão indicadas em azul-claro. Em verde e laranja são mostradas, respectivamente, ligações de hidrogênio adicionais e vinculando o “Shellman box” com o “loop”.

Fitas- β . A fita $\beta 1$ (Glu2-Val7) está ligada à fita $\beta 4$ (Glu60-Ala65) através de um padrão convencional de ligações de hidrogênio, exceto na região da Ile8 onde, junto com o Asp58 e a Ala59 da quarta fita, forma uma saliência- β tipo G1 [95]. A fita $\beta 1$ termina em uma volta- γ inversa [96], formada pelos resíduos Val7-Ile9, imediatamente antes da entrada do “loop” do sítio ativo. A cadeia lateral da Ile8 da saliência faz contatos hidrofóbicos com a Leu86 do “Shellman box”. Esse conjunto de motivos parece importante para manter a interação entre o C-terminal da HPr e o “loop” do sítio ativo. A única outra volta- γ da estrutura ocorre antes do início da fita $\beta 2$ (Gln34-Glu36), entre resíduos Ser31-Ile33. Essa fita está conectada com $\beta 3$ (Lys41-Asn43) por uma volta do tipo I', como na maioria das HPrs de outras espécies, formando um grampo de cabelo tipo 2:4 [97]. O restante da folha- β não apresenta nenhuma conformação que mereça destaque.

Voltas- β . Na proteína são formadas 5 voltas- β , envolvendo os seguintes resíduos: Asp10-Gly13 (tipo I), Tyr37-Lys40 (tipo III'), Asn43-Ser46 (tipo I), Gly56-Ala59 (tipo II) e Gly67-Glu70 (tipo I). As voltas são semelhantes às encontradas nas estruturas de HPr de outros organismos [33].

Resíduos conservados

Uma comparação da sequência de HPr de *Staphylococcus aureus* e outras de estrutura conhecida permite uma racionalização da maioria dos 13 resíduos totalmente conservados. A Gly13 faz parte do “loop” do sítio ativo, e seu ângulo ϕ positivo sugere um papel estrutural importante na manutenção do “loop”. A His15 (na mutante Ala15) é a histidina fosforilada durante a transferência do grupo fosfato de EI para EII. A Arg17 também faz parte do “loop” do sítio ativo; sua cadeia lateral parece estar desordenada na proteína HPr(H15A), coerente com o consenso atual de que esta não assume uma conformação fixa tanto em HPr quanto em P-HPr (fosforilada). Possivelmente seu papel principal resida na interação com resíduos ácidos nos sítios ativos tanto de EI quanto de EII durante a formação dos complexos [19]. A Pro18 interage com a His15, formando um conjunto onde os dois anéis empacotam perpendicularmente entre si. Ala19, Ala26 e Ile63 fazem parte do funil hidrofóbico da proteína, exercendo, portanto, um papel puramente estrutural. A Ser46 é o sítio regulatório de fosforilação em bactérias Gram-positivas. A Gly67 possui um ângulo ϕ positivo e satisfaz o potencial de formação de ligação de hidrogênio do Glu70 na ausência da cadeia lateral adequada do “N-cap” (terceira hélice). Finalmente, o Asp69 forma uma ligação

de hidrogênio, através de sua cadeia lateral, com a amida do Asp30, assim travando o “loop” entre $\beta 4$ e $\alpha 3$ com o “loop” entre $\alpha 1$ e $\beta 2$ na região da Ser31. Com base na estrutura do complexo em *E. coli* espera-se que Lys45 faça parte do sítio de interação com a enzima I. Não há nenhuma explicação evidente para a conservação de Lys40 e Thr62.

3.2.3. Comparação com a proteína nativa HPr(wt) de *S. aureus*

A re-determinação da estrutura da proteína nativa HPr(wt) de *S. aureus* foi realizada em paralelo ao presente trabalho por Sebastian Meier, integrante do grupo do Prof. Dr. Dr. Hans-Robert Kalbitzer, Regensburg, Alemanha. As duas estruturas foram comparadas, mostrando-se bastante semelhantes, com algumas diferenças que poderiam ser decorrentes de condições experimentais (pH, temperatura, concentração da amostra, “shim” do equipamento, seqüências de pulsos), e forma de tratamento de dados distintas (parâmetros dos protocolos CNS e/ou DYANA). Além disso as diferenças reais, se pequenas, podem ser mascaradas pelos intervalos de erro associados aos NOEs.

A fim de eliminar diferenças provenientes de metodologias distintas foram preparadas novas amostras de HPr(H15A) e HPr(wt) simultaneamente, com mesmo solvente, mesma concentração (medida pelo método de Bradford [98]), o pH foi ajustado de forma idêntica; as medidas foram realizadas em seqüência no equipamento de 600 MHz, com mesmo “shim”, mesmo ganho e mesma seqüência de pulsos. A temperatura foi ajustada em 303K. Os espectros NOESY foram processados de maneiras idênticas. Na figura 29 é mostrada uma sobreposição da região entre prótons amídicos/aromáticos e de cadeias laterais dos espectros NOESY das duas proteínas. A maior parte das diferenças na sobreposição decorre das alterações nos deslocamentos químicos, principalmente para os resíduos no sítio ativo, ou devido aos contatos com o anel da His-15 na proteína nativa. As figuras 30 e 31 trazem a diferença dos deslocamentos químicos dos prótons amídicos e $H\alpha$ entre as duas proteínas. Diferenças significativas podem ser observadas na região do sítio ativo — Thr12, Gly13, Ile14, Ala/His15, Ala16, Arg17, Ala19 — e em aminoácidos próximos — Met51, Leu81, Ser82, Leu86. Prótons das cadeias laterais dos resíduos Thr12 ($H\beta$ e $H\gamma 2^*$), Pro18 (todos), Ala19, Met21 ($H\beta 2$ e $H\beta 3$), Leu22 ($H\beta 2/H\beta 3$) e Leu86 ($H\beta 1/H\beta 2$, $H\gamma$, $H\delta 2^*$ e $H\delta 3^*$) também tiveram seus deslocamentos químicos afetados, conforme pode ser visto nos apêndices B e C. Bastante interessante é o comportamento do resíduo Pro18: na proteína mutante os deslocamentos químicos dos prótons $H\gamma 2$ e $H\gamma 3$ são distintos, o mesmo ocorrendo para os prótons $H\delta 2$ e $H\delta 3$. Já na proteína nativa ocorre a degenerescência $H\gamma 2/H\gamma 3$ e $H\delta 2/H\delta 3$.

Figura 29. Comparação da região de contatos entre prótons amídicos/aromáticos e de cadeias laterais dos espectros NOESY de HPr(H15A) (azul) e HPr nativo (vermelho). Ambas as proteínas foram medidas em água, pH 7,0 e temperatura de 303K, no espectrômetro Bruker de 600 MHz.

Figura 30. Diferenças nos deslocamentos químicos dos prótons amídicos entre a proteína mutante HPr(H15A) e a proteína nativa HPr(wt), medidas em condições idênticas ($T=303K$, $pH7,0$). O deslocamento da His15 excede a escala ($\Delta H_N=0,79\text{ppm}$).

Figura 31. Diferenças nos deslocamentos químicos dos prótons H_α entre a proteína mutante HPr(H15A) e a proteína nativa HPr(wt), medidas em condições idênticas ($T=303K$, $pH7,0$). Os prótons $H_{\alpha 1}$ e $H_{\alpha 2}$ das glicinas são mostrados em azul e amarelo.

Para análise e quantificação dos NOEs foram abertas duas janelas com o programa AURELIA e os espectros simultaneamente analisados; apenas foram considerados os seguintes picos:

- NOEs não ambíguos comuns às duas amostras (~850 picos cada);
- NOEs não ambíguos presentes na proteína mutante e claramente ausentes na proteína nativa (NONOE);
- NOEs distintos resultantes da troca Ala15→His ou da degenerescência dos prótons H γ 2/H γ 3 e H δ 2/H δ 3 na Pro18 da proteína nativa.

Os NOEs foram integrados, convertidos em distâncias e organizados em dois arquivos de distâncias interprotônicas: *HPr(H15A).dist* (proteína mutante) e *HPr(wt).dist* (proteína nativa).

Conforme esperado, a maioria dos contatos se mostrou semelhante. Algumas diferenças nas distâncias foram encontradas, sendo, porém, na maioria menores do que o erro experimental a ser associado. Tal fato, associado à baixa quantidade de NOEs considerados, invalida os protocolos de “simulated annealing” aplicados a uma cadeia desenovelada, como gerador de estruturas passíveis de serem comparadas.

Partindo-se da hipótese de que as proteínas são bastante semelhantes, pode-se supor que os fatores causadores de erros experimentais (difusão de spin, diferentes tempos de correlação locais, etc.) são comuns em prótons equivalentes nas duas moléculas. Com base nisso, o seguinte procedimento alternativo foi utilizado:

- dos 40 modelos obtidos para a proteína HPr(H15A) foi escolhida a estrutura de mais baixa energia, *HPr_H15A.pdb*, e nela foram calculadas todas as distâncias interprotônicas $d < 5 \text{ \AA}$ (denominadas aqui de distâncias teóricas);
- realizou-se uma subtração entre as distâncias teóricas e as distâncias interprotônicas de *HPr(H15A).dist*;
- as diferenças foram somadas às distâncias interprotônicas de *HPr(wt).dist*;
- foi criada uma tabela de restrições de distâncias para a proteína nativa, considerando-se as novas distâncias obtidas, com um intervalo NULO de erro experimental (idem para a proteína mutante);
- para os NOEs envolvendo a cadeia lateral de His15 e os prótons H γ 2/H γ 3 e H δ 2/H δ 3 de Pro18 foi criada uma tabela de restrições de distâncias com os intervalos próprios de erro experimental (idem para a proteína mutante);
- foi gerado um arquivo de coordenadas *HPr_wt.pdb* para a proteína nativa, por substituição da Ala15 por His15 na estrutura *HPr_H15A.pdb*;

- o protocolo REFINE (tabela 5) foi aplicado sobre *HPr_wt.pdb*, utilizando-se como restrições experimentais apenas as duas tabelas criadas. Como controle, o mesmo procedimento foi realizado sobre *HPr_H15A.pdb*.

Foram calculadas 5 estruturas para a proteína nativa e 5 para a mutante, como controle. Em decorrência da diminuição significativa de NOEs, reflexo tanto da mais baixa resolução do espectrômetro utilizado (600MHz) quanto da perda pelos critérios utilizados para comparação, nas 5 estruturas obtidas para a HPr(H15A) a sobreposição não é tão boa quanto a encontrada para os 40 modelos anteriormente descritos (figura 22, capítulo III). Na figura 32 são mostradas as estruturas de HPr(H15A) e HPr(wt), na forma de cilindros cujos raios correspondem ao RMSDs das cadeias principais de cada uma.

Figura 32. Sobreposição das 5 melhores estruturas de HPr, na forma de um cilindro cujo raio corresponde ao RMSD das cadeias principais. (a) Mutante HPr(H15A); (b) Nativa HPr(wt).

Verifica-se que, para a proteína mutante, a segunda hélice perdeu resolução (RMSD aumentou); já na proteína nativa essa hélice manteve sua resolução, ocorrendo uma piora da sobreposição no sítio ativo, no início da primeira hélice e na região C-terminal. As diferenças entre as estruturas obtidas podem ser melhor observadas a partir de sua sobreposição, mostrada na figura 33. Diferenças significativas são encontradas na região do sítio ativo. A “tampa” do funil hidrofóbico central, formada pelos resíduos Asp10-Ala16, sofre um deslocamento para baixo na proteína mutante.

Figura 33. Sobreposição “stereo” das estruturas de HPr(H15A), em azul, e HPr(wt), em vermelho.

As duas estruturas foram analisadas cuidadosamente, revelando diferenças interessantes. Na proteína nativa, a cadeia lateral da His15 se alinha perpendicularmente à primeira hélice, fato comum a todas as estruturas já encontradas. Na substituição da histidina por uma alanina é eliminada essa interação, aparentemente permitindo que o “loop” se rearranje de forma a fornecer mais espaço para a Ile14. O rearranjo do “loop” reflete na estabilidade do início da primeira hélice: na proteína mutante as ligações de hidrogênio entre carbonila da Ala16 e amida da Thr20, e entre carbonila da Arg17 e amida da Met21 são significativamente alongadas ($\sim 0.5\text{\AA}$ de diferença) com relação às encontradas na proteína nativa (figura 34). A alongação da segunda já foi comentada na descrição do refinamento completo do mutante. Essa alteração conformacional justifica as diferenças de deslocamentos químicos entre os prótons da cadeia principal de H15A e wt, nessa região. A cadeia lateral da Thr12, voltada em direção à Ala/His15, percebe claramente a variação do ambiente químico: a diferença dos deslocamentos químicos de seus prótons $H\beta$ e $H\gamma$ são significativas (respectivamente $0,37$ e $0,14\text{\AA}$). O mesmo ocorre com os prótons $H\gamma$ e $H\delta$ da Pro18, o que, porém, não explica a degenerescência ocorrida na proteína nativa. Observa-se porém, que nas 5 estruturas obtidas para a proteína mutante o anel da prolina possui configuração em envelope, com o carbono $C\gamma$ fora do plano e apontando para a

Ala15. Nas estruturas da proteína nativa, porém, apesar de se formar o envelope, o carbono C γ se encontra em ambos os lados do plano do anel. Uma possível dinâmica no anel da prolina na proteína nativa justificaria a degenerescência encontrada para os prótons H γ e H δ .

Figura 34. “Loop” do sítio ativo e início da primeira hélice em HPr. **(a)** Proteína mutante (H15A); **(b)** proteína nativa (wt). Para os resíduos Ile14 e Ala/His15 são mostrados todos os átomos, para os demais apenas os átomos da cadeia principal (excluindo H α). As ligações de hidrogênio que estabilizam a hélice estão indicadas.

O rearranjo do “loop” do sítio ativo reflete, também, nos resíduos espacialmente próximos: Val50, Met51, Leu81, Ser82 e Leu86 (figura 35). A ligação de hidrogênio entre a amida de Asp10 e a carbonila de Leu86 é significativamente enfraquecida na mutante ($\sim 2,6\text{\AA}$ na nativa, $\sim 3,2\text{\AA}$ na mutante, com ângulos semelhantes) devido ao afastamento do “loop” em relação ao centro hidrofóbico. A cadeia lateral da Leu86 aponta na direção do sítio ativo, justificando as diferenças de deslocamento químico dos prótons da cadeia lateral encontradas entre proteína nativa e mutante. Em particular Leu 86 perde contato van der Waals com a Gly13 na mutante devido a distorção do “loop”, levando a um aumento de quase $2\text{\AA}$ na menor distância entre os resíduos.

Em estruturas cristalográficas de HPr de outras espécies é encontrada uma água formando ligações de hidrogênio com as carbonilas de Ile14 e Val50. Em *S. aureus* as distâncias encontradas entre essas carbonilas ($\sim 5,0$ e $\sim 5,5\text{\AA}$, respectivamente para proteína nativa e mutante) é coerente com tal interação mediada por duas ligações de hidrogênio, através de uma molécula de água. A orientação relativa das carbonilas muda bastante na mutante, enfraquecendo a suposta interação, liberando a Ile14, que sai um pouco do centro hidrofóbico, e a hélice $\alpha 2$, que se torna mais móvel.

Figura 35. Regiões espacialmente próximas ao “loop” do sítio ativo. **(a)** Sobreposição das estruturas mutante (azul) e nativa (rosa); **(b)** proteína mutante (H15A); **(c)** proteína nativa (wt). As cadeias principais são representadas na forma de “ribbon”. Os átomos das cadeias laterais de resíduos do “loop” e próximos a ele são mostrados.

3.2.4. Discussão e Conclusões

Na última década estruturas de HPr de diversos organismos foram resolvidas, tanto por Cristalografia de raios-X quanto por espectroscopia de Ressonância Magnética Nuclear. Apesar de todas manterem o mesmo padrão estrutural, diferenças significativas são encontradas na região do sítio ativo. Além disso, HPrs não são completamente intercambiáveis em sistemas heterólogos, indicando que aspectos específicos de cada espécie são críticos para a interação proteína-proteína que acompanha a fosfotransferência [28].

A proteína HPr de *Staphylococcus aureus* foi uma das primeiras a ter sua estrutura resolvida, por NMR, em 1993 [23]. A estrutura é de baixa resolução, em decorrência tanto da falta de experimentos heteronucleares quanto do uso de um espectrômetro de 500 MHz, divergindo bastante das estruturas determinadas posteriormente para outros organismos. A

re-determinação da estrutura de HPr de *S. aureus*, em alta resolução (800 MHz, experimentos tridimensionais heteronucleares) mostrou-se, portanto, fundamental, a fim de esclarecer se as divergências estruturais eram reais ou decorrentes de erros experimentais. O fato de uma mutante inativa, cuja histidina do sítio ativo foi substituída por uma alanina, possuir uma afinidade semelhante que a nativa para com as enzimas EI e EII do sistema PTS, motivou a determinação da estrutura de HPr(H15A) de *S. aureus*. Em complemento foram realizados experimentos de NMR visando uma comparação estrutural detalhada entre as proteínas nativa e mutante.

A proteína HPr(H15A) de *S. aureus* mostrou o mesmo padrão estrutural encontrado nas outras espécies, divergindo da estrutura inicial obtida para a proteína nativa conforme pode ser verificado na figura 36. Tais diferenças não são compatíveis com os estudos comparativos realizados entre proteína nativa e mutante, devendo ser decorrentes de erros no assinalamento dos NOEs na estrutura antiga em função de sua baixa resolução.

Figura 36. Comparação entre as estruturas das proteínas HPr mutante, nativa nova e nativa antiga (de baixa resolução) de *S. aureus*. (a) Mutante H15A resolvida por NMR de alta resolução. (b) Nativa obtida por estudos de NMR comparativos. (c) Nativa resolvida por NMR em 1993 [23], código pdb 1ZER. Apenas a primeira das estruturas obtidas é mostrada.

As principais diferenças encontradas entre as estruturas nativa HPr(wt) e mutante HPr(H15A), ambas resolvidas utilizando a mesma sistemática experimental e de assinalamento, foram no “loop” do sítio ativo e resíduos espacialmente próximos a ele. A proteína HPr possui um centro hidrofóbico na forma de um funil, mais largo na região inferior (“loop” do sítio ativo). A interação entre o início da primeira hélice e o anel da His15

segura o loop, introduzindo a Ile14 em uma pequena saliência do funil. Com a substituição da His15 por Ala15 essa interação desaparece, resultando em um afrouxamento do “loop”. Como consequência o início da primeira hélice e a segunda hélice perdem parte de sua estabilidade devido a um enfraquecimento de ligações de hidrogênio com resíduos do “loop”.

Visando entender melhor a alteração conformacional resultante da mutação His15→Ala, as estruturas encontradas para as proteínas nativa e mutante foram comparadas com as já publicadas para os diversos organismos, em particular *Escherichia coli*. Para essa bactéria já foram resolvidas as estruturas de HPr não-fosforilada [14,15], fosforilada (P-HPr) [16] e complexada à enzima EI [19], motivando tais comparações, apesar da natureza Gram negativa da bactéria. As estruturas de HPr e P-HPr de *E. coli*, quando sobrepostas, revelam-se bastante semelhantes, não sendo observada a alteração conformacional observada entre HPr(H15A) e HPr(wt). Em analogia, parece pouco provável que a mutante H15A represente a conformação do estado fosforilado de HPr de *S. aureus*, apesar de sua igual capacidade de interação com as enzimas EI e EII.

Figura 37. Sobreposição do “loop” do sítio ativo entre HPr de *S. aureus* e de *E. coli*. **(a)** A estrutura de HPr(H15A) de *S. aureus* (átomos coloridos por tipo) está sobreposta à de *E. coli* no complexo HPr/EI (em vermelho). O RMSD da cadeia principal é de 0,66Å. **(b)** A sobreposição entre HPr(H15A) de *S. aureus* (átomos coloridos por tipo) e de P-HPr de *E. coli* em solução (em vermelho) resulta em um RMSD de 1,13Å.

As estruturas de HPr e P-HPr de *E. coli* e a estrutura para HPr(H15A) de *S. aureus* foram sobrepostas à HPr do complexo HPr/EI. Para tal finalidade, a sobreposição foi baseada apenas átomos da cadeia principal dos resíduos 10-21, uma vez que era de interesse estudar a conformação do “loop” em si. Surpreendentemente a estrutura que melhor se sobrepôs à HPr do complexo foi a HPr(H15A) de *S. aureus* (RMSD de 0,66Å) e não as estruturas de *E. coli* (RMSD de 1,14 e 1,27Å respectivamente para P-HPr e HPr), conforme pode ser visto na figura 37. Além

disso, entre todas as estruturas de HPr, resolvidas por difração de raios-X ou NMR até o momento, a conformação do “loop” que mais se aproxima à da mutante é a forma complexada de *E. coli*. Este resultado é particularmente interessante, uma vez que *E. coli* é uma bactéria Gram negativa, evolucionariamente mais distante de *S. aureus*.

Tal observação leva à hipótese de que, durante a interação de HPr com EI, a proteína sofra uma mudança conformacional bastante semelhante à encontrada em decorrência da mutação His15→Ala. Caso os mecanismos que levam a essa mudança conformacional sejam análogos aos apresentados para a mutação, durante a formação do complexo HPr/EI alguma interação deva enfraquecer a interação entre a His15 e a primeira hélice. Pode-se especular que tal distorção do “loop” deva favorecer a transferência do fosfato da enzima I para His15 de HPr. Uma vez fosforilada, o dipolo da hélice $\alpha 1$ de HPr agiria para atrair a cadeia lateral da histidina, levando-a para sua conformação original e assim enfraquecendo o complexo e liberando o P-HPr e EI livres.

4. Peptídeo C da Proinsulina Humana

4.1. MATERIAIS E MÉTODOS

No processo de fabricação da insulina comercial, o peptídeo-C é, em geral, descartado juntamente com outros resíduos e solventes. Dadas as recentes evidências de uma relevância biológica desse peptídeo a Empresa BioBras, situada em Montes Claros, MG, se mostrou interessada em um trabalho de colaboração com o Centro de Biotecnologia Molecular Estrutural (CBME), com sede no Instituto de Física de São Carlos, visando estudar a conformação do peptídeo C da insulina humana em solução usando a técnica de espectroscopia por Ressonância Magnética Nuclear. Peptídeo-C foi fornecido pelo Sr. Luciano Villela, diretor da Unidade de Biotecnologia da BioBras. Os experimentos foram realizados na Alemanha, em trabalho de colaboração com o Prof. Dr. Dr. Hans-Robert Kalbitzer, Instituto de Biofísica Bioquímica Física, Universidade de Regensburg.

4.1.1. Preparação das Amostras

Peptídeo-C da proinsulina humana foi purificado e fornecido pela BioBras em solução aquosa contendo 1mg/ml de peptídeo, 0,1% de ácido trifluoracético (TFA) e 10% de acetonitrila (ACN). A fim de minimizar os efeitos do TFA e da ACN, tanto sobre as possíveis propriedades fisiológicas do peptídeo-C quanto pelo efeito perturbativo desses solventes nos espectros de NMR, procedeu-se com sua retirada da seguinte forma:

- 20 ml da solução foram diluídas a 60ml por adição de água bidestilada e o pH inicial de 1,8 foi elevado a 5,8 pela adição de NaOH;
- a solução foi reconcentrada por centrifugação a vácuo até 10ml;
- nova diluição a 60ml por adição de água foi realizada, seguida de centrifugação a vácuo a um volume final de 10ml; tal procedimento foi repetido mais duas vezes; o pH foi controlado e mantido próximo ao neutro a fim de evitar uma precipitação do peptídeo, a qual ocorre entre pH 3 e 4.
- em uma última etapa de centrifugação a vácuo, a solução foi concentrada a 1,0ml (~6mM peptídeo-C) e seu pH ajustado em 7,0 com NaOH.

Optou-se pela diluição/reconcentração, ao invés de uma simples liofilização, por não se ter conhecimentos dos efeitos da liofilização sobre a conformação do peptídeo.

A partir da nova solução foram preparadas três amostras para os experimentos de NMR:

- *Peptídeo-C em água* (~6mM) — A 500µl da solução foram adicionados 2,5µl de DSS (padrão) e 25µl de D₂O (lock);
- *Peptídeo-C em 50% água e 50% TFE-d₂* (~3mM) — A 250µl da solução foram adicionados 250µl de trifluoretanol bideuterado (TFE-d₂ = CF₃CD₂OH) e 2,5µl de DSS;
- *Peptídeo-C em 20% água e 80% TFE-d₂* (~1,2mM) — A 100µl da solução foram adicionados 400µl de trifluoretanol bideuterado e 2,5µl de DSS.

O pH das 3 amostras foi reajustado em 7,0. Os tubos de NMR foram selados com parafilme a fim de evitar evaporação de TFE. As medidas foram realizadas a uma temperatura de 283K.

4.1.2. Espectros de NMR

Os espectros de NMR foram obtidos no espectrômetro Bruker DRX-600, operando em frequência de próton de 600MHz. O sinal da água foi suprimido através de pré-saturação seletiva; todos os espectros foram coletados no modo sensível à fase usando TPPI. Para calibração dos deslocamentos químicos ¹H utilizou-se DSS como referência interna. O software XWINNMR foi utilizado para aquisição e processamento dos dados, em estações Silicon Graphics. A seguir são descritos os experimentos realizados em cada uma das 3 amostras.

COSY. Espectros DQF-COSY foram adquiridos com 16K×1K pontos de dados no domínio temporal. O processamento de cada espectro foi realizado com 4K×2K, utilizando-se filtros de multiplicação Gaussiana em ambas as dimensões (LB=-6Hz e GB=0.12 em t_2 ; LB=-8Hz e GB=0.1 em t_1). Para a análise do acoplamento J foi realizada uma “strip transform” da região NH-H α , com 16K×1K e um filtro de multiplicação exponencial em t_2 (LB=0.1) e senoidal (SSB=0) em t_1 .

TOCSY e NOESY. Espectros TOCSY e NOESY bidimensionais foram adquiridos com 4K×1K pontos de dados no domínio temporal. Nos experimentos TOCSY o tempo de trava de spin foi de 80ms, utilizando a sequência MLEV-17. Nos experimentos NOESY o tempo de mistura foi de 200ms. O processamento de cada espectro foi realizado com 2K×2K pontos de dados no domínio da frequência, utilizando-se filtros idênticos aos do COSY. Nas figuras 38 e 39 são mostrados espectros TOCSY e NOESY sobrepostos para as amostras em água e água/TFE 50%/50%. A figura 40 mostra as regiões de contatos NH-NH para as 3 amostras.

Figura 38. Sobreposição dos espectros TOCSY (vermelho) e NOESY (azul) para o peptídeo-C em água (600MHz, 283K, pH=7,0). (a) Região dos contatos entre prótons HN e aromáticos com prótons alifáticos. (b) Região de contatos entre prótons alifáticos.

Figura 39. Sobreposição dos espectros TOCSY (vermelho) e NOESY (azul) para o peptídeo-C em 50% de água e 50% de TFE (600MHz, 283K, pH=7,0). (a) Região dos contatos entre prótons HN e aromáticos com prótons alifáticos. (b) Região de contatos entre prótons alifáticos.

Figura 40. Região HN-HN dos espectros NOESY (600MHz, 283K, pH 7,0) para o peptídeo-C da insulina humana: **(a)** em água, **(b)** em 50% água, 50% TFE-d₂ e **(c)** em 20% água, 80% TFE-d₂. Em água estão presentes poucos e fracos picos sequenciais, indicando uma fraca tendência helicoidal. Com a adição de 50% de TFE os contatos aumentam em número e intensidade. Observa-se, também, um deslocamento da maioria dos picos para mais baixas frequências, decorrente da alteração do solvente. Em 80% de TFE um novo deslocamento provoca a sobreposição de diversos picos.

Os espectros da amostra em água/TFE 20%/80% se mostraram de qualidade inferior do que para as outras 2 amostras, devido aos efeitos de troca entre a hidroxila protonada do TFE-d₂ e a água. A troca leva a dificuldades em obter um bom “shim”, e portanto a um alargamento dos picos, refletindo em uma baixa relação sinal-ruído.

4.1.3. Estratégia de Assinalamento

O assinalamento ¹H dos sistemas de spin para cada amostra foi realizado conforme descrito em 2.2.1, através dos espectros TOCSY e COSY e com uso do programa AURELIA para visualização. O assinalamento seqüencial foi realizado através do espectro NOESY (região “fingerprint”). Na figura 41 é mostrado o “NOE-pass” para a amostra de Peptídeo-C em água; o padrão é mantido para as outras 2 amostras.

Figura 41. “NOE-pass” do Peptídeo-C em água (600MHz, 283K, pH7,0).

4.1.4. Restrições Experimentais

Como restrições experimentais foram consideradas as constantes de acoplamento $^3J_{\text{NH}\alpha}$ e as distâncias interprotônicas.

Constantes de acoplamento. As constantes de acoplamento para as amostras em água e água/TFE 50%/50% foram obtidas por simulação de dubletos de linhas Lorentzianas nos picos dos respectivos espectros COSY transformados (região NH-H α , “strip-transform”). Aos valores encontrados foram incorporados limites de erro de 0,5Hz. A baixa qualidade do espectro COSY para a amostra em água/TFE 20%/80% impediu o cálculo das constantes de acoplamento.

Distâncias interprotônicas. Após a busca automática dos picos nos espectros NOESY de cada amostra com o programa AURELIA, os NOEs foram integrados em uma caixa de tamanho $S_1=30\text{Hz}$ e $S_2=30\text{ Hz}$, com $N_s=10$ passos de segmentação e até um nível relativo $a_T=0,1$. Os volumes foram transformados em distâncias pela equação (50), considerando-se um tempo de correlação rotacional constante; para calibração foram usados os picos metilênicos. Os intervalos de erro foram calculados conforme descrito previamente (3.1.4.).

4.1.5. Cálculo de Estruturas

A sobreposição dos espectros TOCSY e NOESY para a amostra em água (figura 38) mostra uma quantidade mínima de picos NOEs diferentes dos intra-residuais ou sequenciais HA(i)→HN(i+1), forte indício de que o peptídeo se encontra desestruturado nessas condições. Os fracos e poucos contatos NH-NH (figura 40a) reforçam tal hipótese. Na amostra em água/TFE 20%/80% a baixa qualidade dos espectros leva a uma grande sobreposição de picos, resultando em poucas restrições experimentais de distância. Em ambas as amostras, a pequena quantidade de restrições experimentais seria insuficiente para uma convergência dos protocolos CNS a alguma estrutura definida. Por esse motivo apenas foram calculadas estruturas para o peptídeo-C em água/TFE 50%/50%.

A partir de uma estrutura inicial totalmente linear foram geradas 800 estruturas com o protocolo de SA, contando com 253 restrições de distância e 14 restrições de ângulos torcionais. O protocolo partiu de 3000K, em uma fase quente de dinâmica de ângulos torcionais, sendo resfriado até 0K em uma fase fria de dinâmica cartesiana. Das 800 estruturas geradas foram selecionadas as 40 de mais baixa energia; essas estruturas foram analisadas e, em regiões em que ocorriam divergências, os NOEs foram novamente confirmados e NONOEs introduzidos. As 40 estruturas foram submetidas ao protocolo REFINE, de 75K a 0K. Os parâmetros utilizados nos dois protocolos foram idênticos aos descritos anteriormente (3.1.5., tabela 6), exceto pelas temperaturas iniciais.

4.2. RESULTADOS E DISCUSSÃO

O peptídeo-C em 50% de água e 50% de trifluoretanol apresentou alguns elementos de estrutura secundária bem definidos ao longo das estruturas calculadas, principalmente nas regiões N- e C-terminal.

4.2.1. Assinalamento Seqüencial e Estrutura Secundária

Os deslocamentos químicos ^1H do Peptídeo-C em água, água/TFE 50%/50% e água/TFE 20%/80%, relativos ao DSS, podem ser encontrados no apêndice B. A análise de estrutura secundária para as 3 amostras, através da diferença dos deslocamentos químicos $\text{H}\alpha$ em relação aos valores “random-coil”, é mostrada na figura 42. São identificadas 2 regiões com deslocamentos negativos sistemáticos, potencialmente helicoidais: entre Glu1 e Glu11 e entre Gln22 e Glu27. A região central, entre Leu12 e Leu21, e a C-terminal, a partir de Gly28, não é compatível com uma estrutura helicoidal, podendo consistir de voltas- β . Essas tendências estruturais são potencializadas com a adição de TFE. Vale a pena ressaltar que as diferenças de deslocamento químico foram calculadas com relação ao padrão *em água*; o efeito do TFE sobre os valores de Wishart [71] não são conhecidos, podendo resultar em uma interpretação incorreta do gráfico.

Figura 42. Análise de estrutura secundária: diferenças dos deslocamentos químicos $\text{H}\alpha$ em relação aos padrões Gly-Gly-X-Ala-Gly-Gly e Gly-Gly-X-Pro-Gly-Gly [71], para o peptídeo-C em água (azul), em 50% de água e 50% de TFE- d_2 (vinho) e em 20% de água e 80% de TFE- d_2 (amarelo).

Henriksson *et al.* [57] realizaram estudos por NMR no peptídeo-C da proinsulina humana em 95% de TFE, a uma temperatura de 313K e em pH 3,6. O gráfico por eles apresentado é bastante semelhante ao da figura 42, exceto entre Gln22 a Glu27: nessa região não é encontrada a mesma tendência helicoidal. Tal fato pode ser certamente atribuído às condições experimentais por eles escolhida: pH bastante ácido e alta temperatura. A tendência helicoidal encontrada para os 10 primeiros aminoácidos, assim como a ausência da mesma na região central, é comum em todas as previsões, tanto empíricas (por previsão de estrutura secundária, etc.) [59,63] quanto experimentais (CD, NMR, FTIR), até o momento publicadas [57,61,62].

Figura 43. Mapa de contatos do peptídeo-C em H₂O/TFE (a) 50%/50% e (b) 20%/80%. Os NOEs são divididos em fortes ($d \leq 2,8 \text{ \AA}$), médios ($2,8 \text{ \AA} < d \leq 3,3 \text{ \AA}$) e fracos ($d > 3,3 \text{ \AA}$), refletindo na espessura da linha. As constantes de acoplamento J são divididas em fortes ($J < 6 \text{ Hz}$), representadas por círculos preenchidos, e fracas ($J > 8 \text{ Hz}$), representadas por círculos vazios; valores intermediários são indicados por um X. Espaços vazios não simbolizam a ausência de contatos, sendo, na maioria dos casos, decorrente de picos ambíguos. Linhas cinzas indicam NONOEs confirmados.

A tendência helicoidal das duas regiões, porém, se mostrou incompatível com o mapa de contatos de NOEs sequenciais e de constantes de acoplamento $^3J_{NH-H\alpha}$, mostrado na figura 43 para o peptídeo-C nas duas concentrações de TFE. Para que uma estrutura seja helicoidal as distâncias interprotônicas do tipo $H\alpha(i) \rightarrow HN(i+1)$ devem ser maiores do que $3,3 \text{ \AA}$, resultantes de NOEs fracos, o que claramente não ocorre nessas regiões. Visando eliminar a possibilidade de

que essa discrepância entre as duas informações estruturais pudesse ser decorrente de um erro nas distâncias calculadas, ocorrido por uma escolha muito alta para o tempo de mistura no experimento NOESY, novos experimentos a tempos mais baixos foram realizados. A figura 44 mostra a diferença relativa entre as distâncias interprotônicas, obtidas a 150, 100 e 80ms, com relação às distâncias a 200ms. Utilizou-se a amostra em 50% de água e 50% de TFE; apenas as distâncias entre prótons HN e alifáticos são mostradas. As diferenças encontradas se mostraram, na maioria, inferiores a 10% e portanto dentro dos limites dos erros atribuídos às distâncias. As diferenças maiores encontradas, entre 10 e 25%, se referem a contatos entre prótons amídicos e prótons metílicos nas valinas, leucinas e alaninas, cujos limites de erro superior e inferior atribuídos foram de, no mínimo, 40%. Dessa forma, fica eliminada a possibilidade de um erro nas restrições experimentais de distâncias decorrente de um tempo de mistura inadequado.

Figura 44. Diferenças relativas entre as distâncias entre prótons HN-alifáticos para o Peptídeo-C em água/TFE 50%/50%, provenientes de espectros NOESY a diferentes tempos de mistura τ_{mix} com relação a $\tau_{mix}=200ms$. (a) $\tau_{mix}=150ms$; (b) $\tau_{mix}=100ms$; (c) $\tau_{mix}=80ms$.

4.2.2. Elementos Estruturais

As 40 estruturas de mais baixa energia obtidas para o peptídeo-C em água e TFE 50%/50% foram analisadas com auxílio do programa MOLMOL. Uma sobreposição das estruturas, em função da cadeia principal, não revelou uma estrutura terciária definida. Uma busca de regiões estruturadas foi realizada através de uma sobreposição em setores da cadeia principal. Foram calculados os valores RMSD para seqüências dentro de uma janela de 4, 5 e 6 aminoácidos, que são mostrados na figura 45.

Figura 45. Valores RMSD calculados para a sobreposição dos átomos da cadeia principal nas 40 melhores estruturas, em seqüências de 4 (vermelho), 5 (azul claro) e 6 (azul escuro) aminoácidos.

Figura 46. Um dos 40 modelos para o peptídeo-C. Em azul são indicadas as restrições experimentais de distância utilizadas para gerar as estruturas; 5 setores independentes parecem se formar. Em átomos degenerados são gerados pseudo-átomos, excluídos da figura para melhor visualização. Os contatos envolvendo pseudo-átomos terminam em pontos vazios na figura.

Nas regiões A2-L5, Q9-L12, G15-A18 e Q22-A25 é encontrada uma sobreposição razoável, a qual pode ser vista pela presença de pequenos mínimos na curva vermelha ($\text{RMSD} \leq 0.6\text{\AA}$), indicando a presença de elementos mais estruturados no peptídeo. A melhor sobreposição ocorre, porém, na região C-terminal, para os últimos 6 resíduos L26-Q31. Essas regiões apresentam uma maior concentração de restrições experimentais, como pode ser visto na figura 46: o peptídeo parece estar subdividido em 5 setores, conectados entre si por uma quantidade muito pequena de NOEs. A falta de contatos entre esses setores pode ser reflexo tanto de uma real separação no espaço entre os prótons ($>4\text{\AA}$ para a resolução do espectro NOESY) quanto de picos ambíguos, impossíveis de serem assinalados. As 5 regiões foram analisadas separadamente.

Região Ala2-Leu5. A sobreposição dos átomos da cadeia principal revela a existência de duas famílias de estruturas com conformações distintas, populadas a 85% e 15% (figura 47).

Figura 47. Sobreposição da região Ala2-Leu5 para as 40 melhores estruturas obtidas, com relação aos átomos da cadeia principal. **(a)** Duas famílias de conformações distintas podem ser observadas, populadas a 85% e 15%. **(b)** Uma representante de cada família é mostrada em vermelho (mais populada) e azul (menos populada). As orientações das estruturas foram mantidas.

Em todas as estruturas as distâncias entre O(2) (oxigênio da carbonila do resíduo 2) e $\text{H}_\text{N}(5)$ (hidrogênio amídico do resíduo 5), assim como os ângulos definidos entre O(2)- $\text{H}_\text{N}(5)$ - $\text{N}_\text{H}(5)$ são compatíveis com a presença de ligações de hidrogênio. Na família menos populada essas ligações são menores, variando entre 1,5 e 1,7 $\text{\AA}$, enquanto que os ângulos variam entre

25 e 29°. Também as distâncias entre O(3) e H_N(6) ficam entre 1,9 e 2,2Å, com ângulos entre 29 e 37°, indicando a presença de uma segunda ligação de hidrogênio. Já na família mais populada o resíduo 6 não fica bem definido: uma ligação de hidrogênio entre O(3) e H_N(6) não parece existir. As ligações de hidrogênio entre O(2) e H_N(5) possuem comprimento entre 2,0 e 3,1Å e ângulos entre 22 e 37°. A figura 48 mostra uma representante de cada família; as ligações de hidrogênio são indicadas.

Figura 48. Representante de cada família conformacional da região Ala2-Leu5. **(a)** Mais populada; **(b)** menos populada. As prováveis ligações de hidrogênio, responsáveis pela estabilidade das estruturas, estão indicadas.

O gráfico de Ramachandran para a região é mostrado na figura 49. Nas 6 estruturas da família menos populada os ângulos ϕ, ψ dos resíduos 3, 4 e 5 são compatíveis com duas volta- β do tipo III', início típico de uma estrutura em hélice 3_{10} de mão esquerda, em coerência com as ligações de hidrogênio encontradas. Visto que o resíduo 6 se distribui aleatoriamente no gráfico de Ramachandran, a hélice é interrompida. Já os resíduos 3 e 4 da família mais populada possuem ângulos típicos de uma volta- β tipo I.

Figura 49. Gráfico de Ramachandran da região Ala2-Leu5 para as 40 melhores estruturas obtidas.

Regiões Gln9-Leu12, Gly15-Ala18 e Glu22-Ala25. Nessas 3 regiões a sobreposição dos átomos da cadeia principal não leva a elementos estruturais bem definida, o que pode ser verificado na figura 50, por exemplo, para a região Gln9-Leu12. O gráfico de Ramachandran apresenta uma grande dispersão para os quatro resíduos, nas 3 regiões. A quantidade de restrições experimentais encontrada é pequena, insuficiente para fixar a estrutura; a introdução de NONOEs não foi possível, dada a grande sobreposição de picos no espectro NOESY. Cálculos retroativos através de RELAX não mostram nenhuma violação de restrições de distâncias. Na maioria dos modelos não são observadas distâncias coerentes com ligações de hidrogênio. A leve tendência a uma volta em cada uma das regiões pode ser classificada como dobra (“bend”), cujas distâncias entre resíduos i e $i+3$ são menores do que 7Å.

Figura 50. Sobreposição da região Gln9-Leu12 para as 40 melhores estruturas obtidas, com relação aos átomos da cadeia principal. (a) Modelos sobrepostos; (b) um dos modelos é destacado, com os átomos coloridos por tipo.

Figura 51. Sobreposição da região Leu26-Gln31 para as 40 melhores estruturas obtidas, com relação aos átomos da cadeia principal. (a) Modelos sobrepostos; (b) um dos modelos é destacado, com os átomos coloridos por tipo.

Região Leu26-Gln31. Como pode ser verificado na figura 41, essa região se destaca pela excelente sobreposição entre os átomos da cadeia principal dos 40 modelos. O RMSD para esse hexapeptídeo ($<0.3\text{\AA}$) é menor do que nas demais regiões consideradas, mesmo para apenas os tetrapeptídeos, demonstrando que é de longe a região mais estruturada. As cadeias laterais também parecem bem estruturadas, em especial o resíduo Leu30. A distância entre $O(27)$ e $H_N(30)$ é compatível com a formação de uma ligação de hidrogênio, variando entre 1,7 e 2,6 $\text{\AA}$; o ângulo para a ligação está no intervalo entre 21 e 28°. Essa ligação parece estar sendo bifurcada, envolvendo o átomo $H_N(31)$: as distâncias e ângulos se encontram, respectivamente, nos intervalos 1,8-2,7 $\text{\AA}$ e 9-26°. As duas ligações de hidrogênio estão indicadas na figura 51. Os ângulos ϕ, ψ dos resíduos Gly28 e Ser29 caracterizam uma volta- β tipo III', conforme pode ser verificado no gráfico de Ramachandran (figura 52). O resíduo Leu30 demonstra uma conformação de cadeia principal pouco favorável, porém não proibida, para uma leucina.

Figura 52. Gráfico de Ramachandran da região Leu26-Gln31 para as 40 melhores estruturas obtidas. Os resíduos 28, 29 e 30 são destacados, respectivamente em verde, vermelho e azul.

A volta é extremamente bem definida: nos 800 modelos inicialmente gerados, todos os valores ϕ, ψ encontrados para os resíduos Gly28-Leu30 se concentravam na mesma região do gráfico de Ramachandran. Visando eliminar a possibilidade de que a volta pudesse ter se formado devido a um artefato nas energias do protocolo SA ou REFINE, as estruturas obtidas foram submetidas a novo refinamento, a baixa temperatura (75K), sem a presença de restrições experimentais. Em nenhum dos modelos a volta se manteve. A característica incomum de uma volta- β do tipo III' sugeriu a possibilidade de que as restrições experimentais pudessem ser insuficientes para diferenciá-la de uma volta- β do tipo III. Como teste, em uma das estruturas foram alterados os ângulos ϕ, ψ dos resíduos 29 e 30, de forma a transformar a volta- β em tipo III. As distâncias inter-atômicas encontradas para a nova estrutura eram incoerentes com o espectro NOESY.

4.2.3. Discussão e Conclusões

Com o ressurgimento do interesse nas atividades fisiológicas e farmacológicas do peptídeo-C, duas novas hipóteses surgiram a respeito de sua interação com um eventual receptor. Ido *et al.* propuseram, em 1997 [54], a existência de uma região central do peptídeo-C, com interações aquirais, importantes para sua atividade biológica. Em oposição, Ohtomo *et al.* [55] demonstraram, em 1998, a existência de interações estereo-específicas para o peptídeo-C, típicas de receptores protéicos em membranas. Tais interações quirais sugerem a existência de uma estrutura tridimensional bem definida para pelo menos uma parte do peptídeo-C. Buscando tal estrutura foram realizados os estudos de NMR aqui apresentados.

O peptídeo-C da proinsulina humana, dissolvido em 50% de água e 50% de TFE, não apresentou uma estrutura terciária definida, quando resolvido por espectroscopia de NMR bidimensional. No espectro NOESY não foi encontrado nenhum contato intra-residual de longo alcance, necessário para a convergência dos modelos a alguma estrutura terciária. Uma análise detalhada das estruturas obtidas revelou a existência de 5 regiões, com seqüências mínimas de 4 resíduos, onde ocorre uma melhor sobreposição dos átomos da cadeia principal.

A primeira região, N-terminal (Ala2-Leu5), parece estar constituída de duas famílias conformacionais, possuindo características de voltas- β tipo I (mais populada) e I' (menos populada). A possível estrutura em hélice iniciada por essas voltas é compatível com os deslocamentos químicos para mais baixas frequências encontradas (figura 42), e também coerente com os resultados experimentais prévios [57,61,62]. A hélice é, porém, interrompida após a Leu5: o deslocamento químico do resíduo Gln6 se torna próximo ao encontrado em estruturas

randômicas e nos diferentes modelos os valores ϕ, ψ para a Gln6 se distribuem quase aleatoriamente no gráfico de Ramachandran. A existência de um NONOE entre o próton $H\alpha$ do Asp4 e o próton HN da Val7 confirma a quebra da hélice. Apesar do resíduo Gln6 ser altamente conservado em peptídeos-C, a Leu5 é, em todas as outras espécies diferentes da humana, substituída por uma prolina, resíduo que interrompe hélices.

Experimentos realizados com proinsulina, contendo deleções no peptídeo-C mas mantendo os pares de aminoácidos dibásicos (ver figura 3, capítulo I), mostram que a mesma não é corretamente processada [99,100]. A presença de um tetrapeptídeo N-terminal bastante conservado (ver tabela 3, capítulo I) sugere a presença de um “sítio” na junção BC, importante para o reconhecimento pela endopeptidase tipo I e portanto o seu processamento [100]. A estrutura em volta- β encontrada para Ala2-Leu5 é coerente com tal sugestão.

Nas 3 regiões centrais (Gln9-Leu12, Gly15-Ala18 e Glu22-Ala25) a sobreposição não é tão boa quanto na primeira, indicando algum tipo de volta, porém sem corresponder a um tipo de estrutura secundária própria. A quantidade de NOEs é pequena, em decorrência de uma grande sobreposição de picos no espectro NOESY a qual impossibilita o assinalamento. Não foi possível concluir, portanto, a partir dos experimentos bidimensionais realizados, se essas 3 regiões centrais são realmente desordenadas. Experimentos de FTIR [62], NMR e CD [57,61] confirmam a tendência desordenada dessa região. Estudos de atividade Na^+K^+ -ATPase em segmentos tubulares renais em ratos (cujo estímulo pelo peptídeo-C foi previamente demonstrado [51]), realizados com fragmentos de peptídeo-C [55], mostraram a existência de um fragmento na região central (resíduos 11-19) o qual apresentou alguma atividade biológica. A atividade porém é incompleta, não-específica aos resíduos e ocorre também para um peptídeo feito de aminoácido D. É sugerida a presença de um “sítio ativo” com interações não-específicas na forma de um receptor atípico [58]. Os resíduos 13-19 (GGPGAG) possuem uma característica destacada: são todos apolares, com uma sequência que é quase simétrica em relação à prolina central, com um alto conteúdo do aminoácido aquiral glicina. Já em estudos anteriores [54] foi proposto que esses resíduos formassem uma estrutura em volta- β para realizar interações não quirais com membranas. A leve tendência a uma volta (“bend”) encontrada nas 3 regiões centrais é compatível com tais resultados, porém, por falta de NOEs não ambíguos, uma conclusão final não pode ser tirada.

A região C-terminal (Leu26-Gln31) se destaca pela excelente sobreposição nos 40 modelos. O tetrapeptídeo definido por Glu27-Leu30 forma uma volta- β do tipo III', estabilizada pela ligação de hidrogênio entre a carbonila do Glu27 e a amida da Leu30. Essa ligação é bifurcada, envolvendo

também a amida do Gln31. Tais resultados são coerentes com a formação de uma estrutura bem definida para o pentapeptídeo **EGSLQ** C-terminal. A presença de uma volta tipo III' não é muito comum, sendo, porém, facilitada pela presença de uma glicina na posição (i+1) e de uma serina na posição (i+2), as quais podem adotar conformações de hélice de mão esquerda. O resíduo Leu30 possui uma destacada sobreposição dos átomos da cadeia lateral, conforme pode ser visto na figura 53. A boa definição de uma cadeia lateral, envolvendo metilenos e metilas não estereoespecificamente determinadas, não é comum em estruturas de NMR em decorrência aos grandes intervalos de erro impostos às distâncias. Um exemplo é a Ser29, também evidenciada na figura 53. A conformação da Leu30 parece estar sendo estabilizada por interações de van der Waals com a cadeia principal.

Figura 53. Sobreposição da região Leu26-Gln31 para as 40 estruturas obtidas, com relação aos átomos da cadeia principal. Apenas as cadeias laterais da Ser29 e da Leu30 são mostrados. O grupo sec-metil da leucina é extremamente bem definido, em contraste com o metileno da serina.

Em experimentos de 2D-NMR comparativos, realizados entre proinsulina e insulina [61], foi observada uma mudança significativa entre os deslocamentos químicos das duas moléculas, em alguns resíduos hidrofóbicos localizados no centro da insulina. A clivagem da proinsulina na junção CA (C-terminal da cadeia C) eliminou essas diferenças, o mesmo não ocorrendo após a clivagem na junção BC. Os autores sugerem a existência de uma estrutura estável local na junção CA, denominando-a “CA knuckle”; tal estrutura é acessível à endopeptidase II, envolve uma estrutura secundária não-padrão (de acordo com estudo por CD) e é local (ou seja, não envolve regiões distantes do peptídeo-C). Tais resultados são compatíveis com a estrutura aqui apresentada.

Nos mesmos experimentos de atividade realizados com fragmentos de peptídeo-C [55] foi encontrado um segundo fragmento na região C-terminal, compreendendo os resíduos 27-31, o qual apresentou atividades biológicas característica de interação receptor-ligante, substituindo inteiramente a atividade do peptídeo-C em termos de estimular $\text{Na}^+\text{-K}^+\text{-ATPase}$ e ligação a membranas de células de diversos tecidos, incluindo fibroblastos, células endoteliais e células do tecido renal [55,56]. Os efeitos são específicos aos resíduos, indicando a presença de um receptor específico ao peptídeo-C, o qual reconhece o pentapeptídeo C-terminal [55,58]. Tal fato sugere uma estrutura secundária bem definida para o pentapeptídeo EGSLQ, possivelmente a estrutura em “CA knuckle” anteriormente prevista [61]. De acordo com os resultados encontrados no presente trabalho, a região Leu26-Gln31 é, de longe, a mais estruturada na molécula, dando apoio essas idéias.

O processamento da proinsulina, primeiramente por duas distintas endopeptidases e em seguida por perda dos pares de aminoácidos dibásicos tanto no peptídeo-B quanto no peptídeo-C, poderia ser um indício da função biológica da região C-terminal do peptídeo-C. Sabe-se que, no caso da cadeia B da insulina, a remoção das duas argininas C-terminais pela carboxipeptidase H é essencial para a sua interação com o receptor [101]. O processamento da região C-terminal do peptídeo-C expõe o resíduo Gln-31, totalmente conservado nas seqüências conhecidas, sugerindo um papel funcional importante para o mesmo. A seqüência de resíduos Gly28-Ser29 no pentapeptídeo C-terminal não é absolutamente conservada (e particularmente não está presente em rato). Tal fato pode estar relacionado com receptores espécie-específicos, ou com a possibilidade de que a atividade do peptídeo-C não seja conservada em todas as espécies. A ligação do peptídeo-C de rato a células humanas não foi observada [58], sugerindo que a diferença seqüencial não permita a formação da estrutura encontrada no presente trabalho. No rato a seqüência GS do pentapeptídeo C-terminal é substituída por VA. A valina, devido à bifurcação do carbono- β , muito dificilmente participará de uma volta- β do tipo III' [90].

Vale a pena ressaltar que a estrutura obtida no presente trabalho se refere à molécula em 50% TFE e 50% de água. Em água tal estrutura parece não existir, refletindo o estado do peptídeo em um preparado isolado. A estrutura adquirida pelo peptídeo-C na presença de TFE, em especial a extremamente bem definida encontrada para a região C-terminal, pode ser indício de sua estruturação em um meio menos aquoso, por exemplo, quando interagindo de uma forma estereoespecífica com um provável receptor ou com uma membrana. Em geral a água é excluída de tal interação, possibilitando a formação de tal estrutura. Dessa forma, apesar das condições não-fisiológicas utilizadas (presença de TFE), a estrutura encontrada para o pentapeptídeo C-terminal pode ser relevante fisiologicamente.

Índice de Figuras

Figura 1. Sistema fosfotransferase (PTS), interações e regulação por HPr.	3
Figura 2. Padrão estrutural em proteínas HPr.	5
Figura 3. Processamento da proinsulina para insulina.	7
Figura 4. Estrutura da insulina humana.	9
Figura 5. Orientações dos campos magnéticos e comportamento da magnetização no referencial girante “rotating frame”.	17
Figura 6. Níveis de energia e espectros de NMR esperados para um sistema de spin AX.	24
Figura 7. Relaxação em um sistema de spin- $\frac{1}{2}$ homonucleares, composto de $2N$ spins I e $2N$ spins S, não acoplados escalarmente.	26
Figura 8. Esquema geral para espectroscopia de NMR bidimensional.	28
Figura 9. Dependência do sinal de NMR com o período de evolução, t_1 .	29
Figura 10a. Espectros da Glicina, Alanina, Treonina e Valina.	34
Figura 10b. Espectros da Leucina e Isoleucina.	35
Figura 10c. Espectros da Serina, Cisteína, Aspartato e Asparagina.	36
Figura 10d. Espectros da Histidina, Tirosina, Fenilalanina, Triptofano.	37
Figura 10e. Espectros da Metionina, Glutamato, Glutamina, Prolina, Arginina e Lisina.	38
Figura 11. Espectros fictícios do tetrapeptídeo Met-Gly-Ala-Glu.	39
Figura 12. Características da molécula de TFE.	48
Figura 13. Espectro TOCSY da proteína HPr(H15A) de <i>Staphylococcus aureus</i> .	58
Figura 14. Espectro NOESY da proteína HPr(H15A) de <i>Staphylococcus aureus</i> .	59
Figura 15. Espectro DQF-COSY da proteína HPr(H15A) de <i>Staphylococcus aureus</i> .	60
Figura 16. Espectro ^{15}N -HSQC da proteína HPr(H15A) de <i>S. aureus</i> .	61
Figura 17. Análise de estrutura secundária.	65
Figura 18. Análise de estrutura secundária.	66
Figura 19. Análise de estrutura secundária.	66
Figura 20. Mapa de contatos sequenciais de HPr(H15A).	67
Figura 21. Contatos entre as 4 fitas- β identificadas por análise dos deslocamentos químicos.	67
Figura 22. Estrutura de HPr(H15A).	68

Figura 23. Gráfico de Ramachandran das 40 melhores estruturas de HPr(H15A).	69
Figura 24. Resoluções equivalentes dos modelos obtidos.	71
Figura 25. Geometria do conjunto.	71
Figura 26. Representação em forma de “ribbon” da estrutura média encontrada para a proteína HPr(H15A) de <i>Staphylococcus aureus</i> .	72
Figura 27. Sítio ativo de HPr(H15A)..	73
Figura 28. “Shellman box” (Leu81-Thr87) e sua interação com o início do “loop” do sítio ativo Ile8-Asp10).	75
Figura 29. Comparação da região de contatos entre prótons amídicos/aromáticos e de cadeias laterais dos espectros NOESY de HPr(H15A) (azul) e HPr nativo (vermelho).	78
Figura 30. Diferenças nos deslocamentos químicos dos prótons amídicos entre a proteína mutante HPr(H15A) e a proteína nativa HPr(wt).	79
Figura 31. Diferenças nos deslocamentos químicos dos prótons H α entre a proteína mutante HPr(H15A) e a proteína nativa HPr(wt).	79
Figura 32. Sobreposição das 5 melhores estruturas de HPr, na forma de um cilindro cujo raio corresponde ao RMSD das cadeias principais..	81
Figura 33. Sobreposição das estruturas de HPr(H15A), em azul, e HPr(wt), em vermelho	82
Figura 34. “Loop” do sítio ativo e início da primeira hélice em HPr.	83
Figura 35. Regiões espacialmente próximas ao “loop” do sítio ativo.	84
Figura 36. Comparação entre as estruturas das proteínas HPr mutante, nativa nova e nativa antiga (de baixa resolução) de <i>S. aureus</i> .	85
Figura 37. Sobreposição do “loop” do sítio ativo entre HPr de <i>S. aureus</i> e de <i>E. coli</i> .	86
Figura 38. Sobreposição dos espectros TOCSY (vermelho) e NOESY (azul) para o peptídeo-C em água (600MHz, 283K, pH=7,0).	90
Figura 39. Sobreposição dos espectros TOCSY (vermelho) e NOESY (azul) para o peptídeo-C em 50% de água e 50% de TFE (600MHz, 283K, pH=7,0).	90
Figura 40. Região HN-HN dos espectros NOESY (600MHz, 283K, pH 7,0) para o peptídeo-C da insulina humana.	91
Figura 41. “NOE-pass” do Peptídeo-C em água (600MHz, 283K, pH7,0).	92
Figura 42. Análise de estrutura secundária.	94
Figura 43. Mapa de contatos do peptídeo-C em H ₂ O/TFE.	95

Figura 44. Diferenças relativas entre as distâncias entre prótons HN-alifáticos para o Peptídeo-C em água/TFE 50%/50%, provenientes de espectros NOESY a diferentes tempos de mistura τ_{mix} com relação a $\tau_{mix}=200ms$.	96
Figura 45. Valores RMSD calculados para a sobreposição dos átomos da cadeia principal nas 40 melhores estruturas.	97
Figura 46. Um dos 40 modelos para o peptídeo-C.	97
Figura 47. Sobreposição da região Ala2-Leu5 para as 40 melhores estruturas obtidas, com relação aos átomos da cadeia principal.	98
Figura 48. Representante de cada família conformacional da região Ala2-Leu5.	99
Figura 49. Gráfico de Ramachandran da região Ala2-Leu5 para as 40 melhores estruturas obtidas.	100
Figura 50. Sobreposição da região Gln9-Leu12 para as 40 melhores estruturas obtidas.	101
Figura 51. Sobreposição da região Leu26-Gln31 para as 40 melhores estruturas obtidas.	101
Figura 52. Gráfico de Ramachandran da região Leu26-Gln31 para as 40 melhores estruturas.	102
Figura 53. Sobreposição da região Leu26-Gln31 para as 40 estruturas obtidas.	105

Índice de Tabelas

Tabela 1. Estruturas de HPr resolvidas até o momento (março de 2001).	4
Tabela 2. Alinhamento de proteínas HPr de algumas espécies.	5
Tabela 3. Alinhamento entre peptídeos-C de algumas espécies.	8
Tabela 4. Distâncias inter-protônicas sequenciais típicas (em Å) dos elementos de estrutura secundária.	43
Tabela 5. Caracterização das três conformações possíveis em torno da ligação $C\alpha-C\beta$.	45
Tabela 6. Caracterização dos três rotâmeros possíveis em torno de $C\alpha-C\beta$ ligados em valinas.	46
Tabela 7. Parâmetros utilizados nos protocolos de SA e refinamento.	64

Abreviações

1D, 2D, 3D	Uni-, bi-, tri-dimensional
AQUA	Analysing the quality
AURELIA	Automatic resonance line assignment
CD	Circular dichroism spectroscopy
CNS	Crystallography and NMR system
COSY	Correlated spectroscopy
CcpA	Catabolite control protein
DQF	Double-quantum-filter
DSS	Sodium 2,2-dimethyl-2-silapentane-5-sulfonate
DTE	1,4-Dithioerythritol
DYANA	Dynamics algorithm for NMR applications
EDTA	Ethylenediaminetetraacetic acid
FID	Free induction decay
FTIR	Fourier transform infrared spectroscopy
FT-NMR	Fourier transform nuclear magnetic resonance spectroscopy
HOHAHA	Homonuclear Hartmann-Hahn
HPr	Histidine-containing phosphocarrier protein
HSQC	Heteronuclear single quantum coherence
IUPAC	International Union of Pure and Applied Chemistry
MD	Molecular dynamics
NMR	Nuclear magnetic resonance spectroscopy
NOE	Nuclear Overhauser enhancement
NOESY	Nuclear Overhauser-effect spectroscopy
PDB	Protein data bank
PEP	Phosphoenolpyruvate
ppm	Parts per million
PROCHECK	Protein check

PTS	Phospho <i>enol</i> pyruvate:sugar phosphotransferase system
RMSD	Root mean squared deviation
TFE	2,2,2-trifluoroethanol
TOCSY	Total correlated spectroscopy
TPPI	Time-propotional phase-incrementation

Símbolos

B_0	Amplitude do campo magnético estático
B_1	Amplitude do campo magnético de rádio-freqüência
$B_{rf}(t)$	Campo magnético de rádio-freqüência
$B(t)$	Vetor de campo magnético
Da	Dalton (1 Da = massa de um átomo de 1H)
δ	Deslocamento químico (ppm)
ΔB_0	Amplitude do campo magnético reduzido
$\Delta\nu$	Largura de linha em Hertz
Δt	Intervalo de tempo
E	Energia
γ	Fator giromagnético
$h=2\pi\hbar$	Constante de Planck
I	Número quântico de momento angular de spin
$\mathbf{I}$	Vetor de momento angular
I_z	Componente z do vetor de momento angular
$J(\omega)$	Função de densidade espectral
$^nJ_{ij}$	acoplamento escalar de n -ligações, entre o spin I e o spin j
k_B	Constante de Boltzmann
m	Número quântico magnético
M_0	Magnetização no equilíbrio

$\mathbf{M}(t)$	Vetor de magnetização
M_r	Massa molecular relativa
μ	Vetor de momento magnético
μ_0	Permissividade do vácuo
μ_z	Componente z do vetor de momento magnético
ν_0	Frequência de ressonância em Hertz
ω	Velocidade angular
ω_0	Frequência de ressonância de Larmor
ω_l	Intensidade do campo de rf
ω_f	Frequência do campo de rf
Ω	“offset” do campo estático reduzido
pH	$-\log a(\text{H}^+)$ (a = Atividade)
ϕ	Fase do campo de rf
R_1	Taxa de relaxação longitudinal ou spin-rede
R_2	Taxa de relaxação transversal ou spin-spin
R_2^*	Taxa de relaxação transversal não-homogênea
R_{inom}	Constante de alargamento de linha não-homogêneo
$S(\omega)$	Sinal no domínio da frequência
$s(t)$	Sinal no domínio do tempo
σ	Blindagem nuclear isotrópica
σ_{ij}	Relaxação cruzada entre núcleos i e j
T_1	Tempo de relaxação longitudinal ou spin-rede
T_2	Tempo de relaxação transversal ou spin-spin
$t_1, t_2, \dots$	Tempo direto de aquisição, ou indireto de evolução
τ_c	Tempo de correlação
τ_{mix}	Tempo de mistura
$u(\omega)$	parte imaginária do espectro complexo no domínio da frequência
$v(\omega)$	parte real do espectro complexo no domínio da frequência

Bibliografia

PUBLICAÇÕES EM PERIÓDICOS

- [1] Wider, G. (2000). Structure determination of biological macromolecules in solution using NMR spectroscopy. *Biotechniques*, **29**, 1278-1294.
- [2] Bernstein F.C., Koetzle T.F., Williams G.J.B, Meyer E.F.Jr., Brice M.D., Rogers J.R., Kennard O., Shimanouchi T. & Tasumi M. (1977). The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.*, **112**, 535-542.
- [3] Kundig W.S., Ghosh S. & Roseman S. (1964). Phosphate bound to a histidine in a protein as an intermediate in a novel phosphotransferase system. *Proc. Natl. Acad. Sci. USA*, **52**, 1067-1074.
- [4] Meadow N.D., Fox D.K. & Roseman S. (1990). The bacterial phosphoenolpyruvate: glycose phosphotransferase system. [Review]. *Annu. Rev. Biochem.*, **59**, 497-542.
- [5] Postma P.W., Lengler L.W. & Jacobson G.R. (1993). Phosphoenolpyruvate: carbohydrate phosphotransferase system in bacteria. *Microbiol. Rev.*, **57**, 543-594.
- [6] Saier M.H.Jr. & Reizer J. (1990). Domain shuffling during evolution of the proteins of the bacterial phosphotransferase system. *Res. Microbiol.*, **141**, 1033-1038.
- [7] Postma P.W. & Lengler L.W. (1985). Phosphoenolpyruvate:carbohydrate phosphotransferase system of bacteria. [Review]. *Microbiol. Rev.*, **49**, 232-269.
- [8] Jones B.E., Rajagopal P. & Klevit R.E. (1997). Phosphorylation on histidine is accompanied by localized structural changes in the phosphocarrier protein, HPr from *Bacillus subtilis*. *Protein Sci.*, **6**, 2107-2119.

- [9] Galinier A., Kravanja M., Engelmann R., Hengstenberg W., Kilhoffer M.C., Deutscher J. & Haiech J. (1998). New protein kinase and protein phosphatase families mediate signal transduction in bacterial catabolite repression. *Proc. Natl. Acad. Sci. USA*, **95**, 1823-1828.
- [10] Galinier A., Deutscher J. & Martin-Verstraete I. (1999). Phosphorylation of either Crh or HPr mediates binding of CcpA to the *Bacillus subtilis xyn cre* and catabolite repression of the *xyn* operon. *J. Mol. Biol.*, **286**, 307-314.
- [11] Kravanja M., Engelmann R., Dossonnet V., Bluggel M., Meyer H.E., Frank R., Galinier A., Deutscher J., Schnell N. & Hengstenberg W. (1999). The *hprK* gene of *Enterococcus faecalis* encodes a novel bifunctional enzyme: the HPr kinase/phosphatase. *Mol. Microbiol.*, **31**, 59-66.
- [12] Deutscher J. & Saier M.H.Jr. (1983). ATP-dependent protein kinase-catalyzed phosphorylation of a seryl residue in HPr, a phosphate carrier protein of the phosphotransferase system in *S. pyogenes*. *Proc. Natl. Acad. Sci. USA*, **80**, 6790-6795.
- [13] Reizer J., Sutrina S.L., Saier M.H.Jr., Stewart G.C., Peterkofsky A. & Reddy P. (1989). Mechanistic and physiological consequences of HPr(Ser)phosphorylation on the activities of the phosphoenolpyruvate:sugar phosphotransferase system in Gram-positive bacteria: studies with sitespecific mutants of HPr. *EMBO J.*, **8**, 2111-2120.
- [14] Jia Z., Quail J.W., Waygood E.B. & Delbaere L.T. (1993). The 2.0-Å resolution structure of *Escherichia coli* histidine-containing phosphocarrier protein HPr. A redetermination. *J. Biol. Chem.*, **268**, 22490-22501.
- [15] van Nuland N.A., Hangyi I.W., van Schaik R.C., Berendsen H.J., van Gunsteren W.F., Scheek R.M. & Robillard G.T. (1994). The high-resolution structure of the histidine-containing phosphocarrier protein HPr from *Escherichia coli* determined by restrained molecular dynamics from nuclear magnetic resonance nuclear Overhauser effect data. *J. Mol. Biol.*, **237**, 544-559.

- [16] van Nuland N.A., Boelens R., Scheek R.M. & Robillard G.T. (1995). High-resolution structure of the phosphorylated form of the histidine-containing phosphocarrier protein HPr from *Escherichia coli* determined by restrained molecular dynamics from NMR-NOE data. *J. Mol. Biol.*, **246**, 180-193.
- [17] Napper S., Anderson J.W., Georges F., Quail J.W., Delbaere L.T.J. & Waygood E.B. (1996). Mutation of serine-46 to aspartate in the histidine-containing protein of *Escherichia coli* mimics the inactivation by the phosphorylation of serine-46 in HPrs from Gram-positive bacteria. *Biochemistry*, **35**, 11260.
- [18] Prasad L., Waygood E.B, Lee J.S. & Delbaere L.T.J. (1998). The 2.5 Å structure of the Jel42 Fab fragment/HPr complex. *J. Mol. Biol.*, **280**, 829.
- [19] Garrett D.S., Seok Y.J., Peterkofsky A., Gronenborn A.M.& Clore G.M. (1999). Solution structure of the 40,000 M_r phosphoryl transfer complex between the N-terminal domain of enzyme I and HPr. *Nat. Struct. Biol.*, **6**, 166-173.
- [20] Napper S., Delbaere L.T.J. & Waygood E.B. (1999). The aspartyl replacement of the active site histidine in histidine-containing protein, HPr, of the *Escherichia coli* phosphoenolpyruvate:sugar phosphotransferase system can accept and donate a phosphoryl group. *J. Biol. Chem.*, **274**, 21776-21782.
- [21] Herzberg O., Reddy P., Sutrina S., Saier M.H., Reizer J. & Kapadia G. (1992). Structure of the histidine-containing phosphocarrier protein HPr from *Bacillus subtilis* at 2.0-Å resolution. *Proc. Natl. Acad. Sci. USA*, **89**, 2499-2503.
- [22] Liao D.-I., Herzberg O. (1994). Refined structures of the active Ser83→Asp and impaired Ser46→Asp histidine-containing phosphocarrier proteins. *Structure*, **2**, 1203-1216.
- [23] Kalbitzer H.R. & Hengstenberg W. (1993). The solution structure of the histidine-containing protein (HPr) from *Staphylococcus aureus* as determined by two-dimensional ¹H-NMR spectroscopy. *Eur. J. Biochem.*, **216**, 205-214.

- [24] Görler A., Hengstenberg W., Kravanja M., Beneicke W., Maurer T. & Kalbitzer H.R. (1999). Solution structure of the histidine-containing phosphocarrier protein from *Staphylococcus carnosus*. *Appl. Magn. Reson.*, **17**, 465-480.
- [25] Jia Z., Vandonselaar M., Hengstenberg W., Quail J.W. & Delbaere L.T. (1994). The 1.6 Å structure of histidine-containing phosphotransfer protein HPr from *Streptococcus faecalis*. *J. Mol. Biol.*, **236**, 1341-1355.
- [26] Maurer T., Döker R., Hahmann M., Görler A., Hengstenberg W. & Kalbitzer H.R. (2001). Three dimensional structure of the histidine-containing phosphocarrier protein (HPr) from *Enterococcus faecalis* in solution. *Eur. J. Biochem.*, **268**, 635.
- [27] Pieper U., Kapadia G., Zhu P.-P., Peterkofsky A. & Herzberg O. (1995). Structural evidence for the evolutionary divergence of mycoplasma from Gram-positive bacteria: the histidine-containing phosphocarrier protein. *Structure*, **3**, 781-790.
- [28] Reizer J., Sutrina S.L., Wu L.F., Deutscher J., Reddy P. & Saier M. (1992). Functional interactions between proteins of the phosphoenolpyruvate:sugar phosphotransferase systems of *Bacillus subtilis* and *Escherichia coli*. *J. Biol. Chem.*, **267**, 9158-9169.
- [29] van Nuland N.A., Grötzinger J., Dijkstra K., Scheek R.M. & Robillard G.T. (1992). Determination of the three-dimensional solution structure of the histidine-containing phosphocarrier protein HPr from *Escherichia coli* using multidimensional NMR spectroscopy. *Eur. J. Biochem.*, **210**, 881-891.
- [30] Jia Z., Vandonselaar M., Quail J.W. & Delbaere L.T. (1993). Active-centre torsion-angle strain revealed in 1.6 Å-resolution structure of histidine-containing phosphocarrier protein. *Nature*, **361**, 94-97.
- [31] Kruse R., Hengstenberg W., Beneicke W. & Kalbitzer H.R. (1993). Involvement of the various amino- and carboxyl-terminal residues in the active site of the histidine-containing protein HPr of the phosphoenolpyruvate-dependent phosphotransferase system of

Staphylococcus carnosus: site-dependent mutagenesis with the *ptsH* gene, biochemical characterization and NMR studies of the mutant proteins. *Protein Engineering*, **6**, 417-423.

- [32] van Nuland N.A., Kroon G.J.A., Dijkstra K., Wolters G.K., Scheek R.M. & Robillard G.T. (1993). The NMR determination of the IIA^{mtl} binding site on the HPr of the *Escherichia coli* phosphoenol pyruvate-dependent phosphotransferase system. *FEBS Lett.*, **315**, 11-15.
- [33] Jia Z., Quail J.W., Delbaere L.T. & Waygood E.B. (1994). Structural comparison of the histidine-containing phosphocarrier protein HPr. *Biochem. Cell Biol.*, **72**, 202-217.
- [34] Rajagopal P., Waygood E.B. & Klevit R.E. (1994). Structural consequences of the histidine phosphorylation: NMR characterization of the phosphohistidine form of the histidine-containing protein from *Bacillus subtilis* and *Escherichia coli*. *Biochemistry*, **33**, 15271-15282.
- [35] Anderson J.W. (1995). HPr: a model protein. [Review]. *Biochem. Cell Biol.*, **73**, 219-222.
- [36] Bowers P.M. & Klevit R.E. (1996). Hydrogen bonding and equilibrium isotope enrichment in the histidine-containing proteins. *Nat. Struc. Biol.*, **3**, 522-531.
- [37] Liao D.-I., Silverton E., Seok Y.-J., Lee B.R., Peterkofsky A. & Davies D.R. (1996). The first step in sugar transport: crystal structure of the amino terminal domain of enzyme I of the *E. coli* PEP: sugar phosphotransferase system and a model of the phosphotransfer complex with HPr. *Structure*, **4**, 861-872.
- [38] Thapar R., Nicholson E.M., Rajagopal P., Waygood E.B., Scholtz J.M. & Klevit R.E. (1996). Influence of N-Cap mutations on the structure and stability of the *Escherichia coli* HPr. *Biochemistry*, **35**, 11268-11277.
- [39] Garrett D.S., Seok Y.J., Liao D.I., Peterkofsky A., Gronenborn A.M. & Clore G.M. (1997). Solution structure of the 30 kDa N-terminal of enzyme I of the *Escherichia coli* phosphotransferase system by multidimensional NMR. *Biochemistry*, **36**, 2517-2530.

- [40] Garrett D.S., Seok Y.J., Peterkofsky A., Clore G.M. & Gronenborn A.M. (1997). Identification by NMR of the binding surface for the histidine-containing phosphocarrier protein HPr on the N-terminal domain of enzyme I of the *Escherichia coli* phosphotransferase system. *Biochemistry*, **36**, 4393-4398.
- [41] Hahmann M., Maurer T., Lorenz M., Hengstenberg W., Glaser S. & Kalbitzer H.R. (1998). Structural studies of histidine-containing phosphocarrier protein from *Enterococcus faecalis*. *Eur. J. Biochem.*, **252**, 51-58.
- [42] Steiner D., Cunningham D. & Tager H. (1967). Insulin biosynthesis: evidence of a precursor. *Science*, **157**, 697-700.
- [43] Steiner D. & Oyer P. (1967). Biosynthesis of insulin and a probable precursor of insulin by a human islet cell adenoma. *Proc. Natl. Acad. Sci. USA*, **57**, 473-481.
- [44] Rubenstein A., Clark J., Melani F. & Steiner D. (1969). Secretion of proinsulin, C-peptide by pancreatic beta cells and its circulation in blood. *Nature*, **224**, 697-699.
- [45] Davidson H.W., Rhodes C.J. & Hutton J.C. (1988). Intra-organelle Ca^{2+} and pH control proinsulin cleavage in the pancreatic β -cell via two distinct site-specific endopeptidases. *Nature*, **333**, 93-96.
- [46] Clark, P.M. (1999). Assays for insulin, proinsulin(s) and C-peptide. [Review]. *Ann. Clin. Biochem.*, **36**, 541-564.
- [47] Johansson B.-L., Sjöberg S. & Wahren J. (1992). The influence of human C-peptide on renal function and glucose utilization in type 1 (insulin-dependent) diabetic patients. *Diabetologia*, **35**, 121-128.
- [48] Johansson B.-L., Linde B. & Wahren J. (1992). Effects of C-peptide on blood flow, capillary diffusion capacity and glucose utilization in the exercising forearm of type 1 (insulin-dependent) diabetic patients. *Diabetologia*, **35**, 1151-1158.

- [49] Sjoquist M., Huang W. & Johansson B.-L. (1998). Effects of C-peptide on renal function at the early stage of experimental diabetes. *Kidney Int.*, **54**, 758-764.
- [50] Wu W., Oshida Y., Yang W.-P., Ohsawa I., Sato J., Iwao S., Johansson B.-L., Wahren J. & Sato Y. (1996). Effect of C-peptide administration on whole body glucose utilization in STZ-induced diabetic rats. *Acta Physiol. Scand.*, **157**, 253-258.
- [51] Forst T., Kunt T., Pohlmann T., Goitom K., Engelbach M., Beyer J. & Pfützner A. (1998). Biological activity of C-peptide on the skin microcirculation in patients with insulin-dependent diabetes mellitus. *J. Clin. Invest.*, **101**, 2036-2041.
- [52] Johansson B.-L., Borg K., Fernqvist-Forbes E., Odergren T., Remahl S. & Wahren J. (1996). C-peptide improves autonomic nerve function in IDDM patients. *Diabetologia*, **39**, 687-695.
- [53] Zierath J.R., Handberg A., Tally M. & Wallberg-Henriksson H. (1996). C-peptide stimulates glucose transport in isolated human skeletal muscle independent of insulin receptor and tyrosine kinase activation. *Diabetologia*, **39**, 306-313.
- [54] Ido Y., Vindigni A., Chang K., Stramm L., Chance R., Heath W.F., DiMarchi R.D., Di Cera E. & Williamson (1997). Prevention of Vascular and Neural Dysfunctionb in Diabetic Rats by C-Peptide. *Science*, **277**, 563-566.
- [55] Ohtomo Y., Bergman T., Johansson B.-L., Jörnvall H. & Wahren J. (1998). Differential effects of proinsulin C-peptide fragments on Na^+, K^+ -ATPase activity of renal tubule segments. *Diabetologia*, **41**, 287-291.
- [56] Rigler R., Pramanik A., Jonasson P., Kratz G., Jansson O.T., Nygren P.-Å., Ståhl S., Ekberg K., Johansson B.-L., Uhlén S., Uhlén M., Jörnvall H. & Wahren J. (1999). Specific binding of proinsulin C-peptide to human cell membranes. *PNAS*, **96**, 13318-13323.
- [57] Henriksson M., Shafqat J., Liepinsh E., Tally M., Wahren J., Jörnvall H., Johansson J.. (2000). Unordered structure of proinsulin C-peptide in aqueous solution and in the presence of lipid vesicles. *Cell. Mol. Life Sci.*, **57**, 337-342.

- [58] Wahren J., Ekberg K., Johansson J., Henriksson M., Pramanik A., Johansson B.-L., Rigler R. & Jörnvall H. (2000). Role of C-peptide in human physiology. *Am. J. Physiol. Endocrinol. Metab.*, **278**, E759-E768.
- [59] Snell C.R. & Smyth D.G. (1975). Proinsulin: a proposed three-dimensional structure. *J. Biol. Chem.*, **250**, 6291-6295.
- [60] Hua, Q.X., Gozani, S.N., Chance, R.E., Hoffmann, J.A., Frank, B.H., Weiss, M.A. (1995). Structure of a protein in a kinetic trap. *Nat. Struct. Biol.*, **2**, 129-138.
- [61] Weiss M.A., Frank B.H., Khait I., Pekar A., Heiney R., Shoelson S. & Neuriger L.J. (1990). NMR and photo-CIDNP studies of human proinsulin and prohormone processing intermediates with application to endopeptidase recognition. *Biochemistry*, **29**, 8389-8401.
- [62] Xie L., Tsou C.-L. (1993). Comparison of secondary structures of insulin and proinsulin by FTIR. *J. Prot. Chem.*, **12**, 483-487.
- [63] Chou P.Y. & Fasman G.D. (1978). Prediction of the secondary structure of proteins from their amino acid sequence. [Review]. *Adv. Enzymol.*, **47**, 45-148.
- [64] Steiner D.F. (1978). On the role of proinsulin C-peptide. *Diabetes*, **27**, 145-148.
- [65] Bloch F., Hansen W.W. & Packard M. (1946). Nuclear induction. *Phys. Rev.*, **69**, 127.
- [66] Purcell E.M., Torrey H.C. & Pound R.V. (1946). Resonance absorption by nuclear magnetic moments in a solid. *Phys. Rev.*, **69**, 37-38.
- [67] Bloch F. (1946). Nuclear induction. *Phys. Rev.*, **70**, 460-474.
- [68] Bloch F. & Siegert A. (1940). Magnetic resonance for nonrotating fields. *Phys. Rev.*, **57**, 522-527.

- [69] Ramsey N.F. & Purcell E.M. (1952). *Phys. Rev.*, **85**, 143-144.
- [70] Wüthrich K., Billeter M. & Braun W. (1984). Polypeptide secondary structure determination by nuclear magnetic resonance observation of short proton-proton- distances. *J. Mol. Biol.*, **180**, 715-740.
- [71] Wishart D.S., Bigam C.G., Holm A., Hodges R.S. & Sykes B.D. (1995). ^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*, **5**, 67-81.
- [72] Geyer M., Neidig K.-P. & Kalbitzer H.R. (1995). Automated peak integration in multidimensional NMR spectra by an optimized iterative segmentation procedure. *Journal of Magnetic Resonance*, **B 109**, 31-38.
- [73] Neidig K.-P., Geyer M., Görler A., Antz C., Saffrich R., Beneicke W. & Kalbitzer H.R. (1995). AURELIA: a program for computer-aided analysis of multidimensional NMR-spectra. *J. Biomol. NMR*, **6**, 255-270.
- [74] Karplus M. (1959). Contact electron-spin coupling of nuclear magnetic moments. *J. Chem. Phys.*, **30**, 11-15.
- [75] Bystrov V.F. (1976). Spin-spin coupling and the conformational states of peptide systems. *Progr. Nucl. Magn. Reson.*, **10**, 41-82.
- [76] de Marco A., Llinás M. & Wüthrich, K. (1978). Analysis of the ^1H NMR spectra of ferrichrome peptides I: the non-amide protons. *Biopolymers*, **17**, 617-636.
- [77] Pardi A., Billeter M. & Wüthrich, K. (1996). Calibration of the angular dependence of the amide proton-C α proton coupling constants, $^3J_{\text{HN}\alpha}$, in a globular protein: use of $^3J_{\text{HN}\alpha}$ for identification of helical secondary structure. *J. Mol. Biol.*, **180**, 741-751.
- [78] Basus V.J. (1989). Proton nuclear magnetic resonance assignments. *Methods in Enzymology*. **177**, 132-149.

- [79] Rajan R. & Balaram P. (1996). A model for the interaction of trifluoroethanol with peptides and proteins. *Int. J. Peptide Protein Res.*, **48**, 328-336.
- [80] Blundell T., Barlow D., Borkakoti, Nivedita & Thornton J. (1983). Solvent-induced distortions and the curvature of α -helices. *Nature*, **306**, 281-283.
- [81] Görler A. & Kalbitzer H.R. (1997). RELAX: A flexible program for the backcalculation of NOESY-spectra based on the complete relaxation matrix formalism. *Journal of Magnetic Resonance*, **124**, 177-188.
- [82] Brunger A.T., Adams P.D., Clore G.M., Delano W.L., Gros P., Grosse-Kunstleve R.W., Jiang J.-S., Kuszewski J., Nilges M., Pannu N.S., Read R.J., Rice L.M., Simonson T. & Warren G.L. (1998). Crystallography & NMR System: A new software suite for macromolecular structure determination. *Acta Cryst. D*, **54**, 905-921.
- [83] Güntert, P., Mumenthaler, C. & Wüthrich, K. (1997). Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.*, **273**, 283-298.
- [84] Koradi R., Billeter M., Engeli M., Guntert P. & Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.*, **14**, 51-55, 29-32.
- [85] Nilges M., Clore G.M. & Gronenborn A.M. (1988). Determination of three-dimensional structures of proteins from interproton distance data by hybrid distance geometry-dynamical simulated annealing calculations. *FEBS Lett.* **229**, 317-324.
- [86] Nilges M., Clore G.M. & Gronenborn A.M. (1988). Determination of three-dimensional structures of proteins from interproton distance data by dynamical simulated annealing from a random array of atoms. Circumventing problems associated with folding. *FEBS Lett.*, **239**, 129-136.

- [87] Nilges M., Clore G.M. & Gronenborn A.M. (1988). Determination of three-dimensional structures of proteins from interproton distance data by hybrid distance geometry-dynamical simulated annealing calculations. *FEBS Lett.*, **229**, 317-324.
- [88] Tjandra N., Marquardt J. & Clore G.M. (2000). Direct refinement against proton-proton dipolar couplings in NMR structure determination of macromolecules. *Journal of Magnetic Resonance*, **142**, 393-396.
- [89] Laskowski R.A., Rullmann J.A.C., MacArthur M.W., Kaptein R. & Thornton J.M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR*, **8**, 477-486.
- [90] Zhao D. & Jardetzky O. (1994). An Assessment of the Precision and Accuracy of Protein Structure Determined by NMR – Dependence on Distance Errors. *Journal of Molecular Biology*, **239**, 601-607.
- [91] Morris A.L., MacArthur M.W., Hutchinson E.G., Thornton J.M. (1992). Stereochemical quality of protein structure coordinates. *Proteins: Struct. Func. Gen.*, **12**, 345-364.
- [92] Ramachandran G.N., Ramakrishnan C. & Sasisekharan V. (1963). Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.*, **7**, 95-99.
- [93] Janin J. & Chothia C. (1980). Packing of alpha-helices onto beta-pleated sheets and the anatomy of alpha-beta-proteins. *J. Mol. Biol.*, **143**, 95-128.
- [94] Aurora R., Srinivasan R. & Rose, G.D. (1994). Rules for α -helix termination by glycine. *Science*, **264**, 1126-1130.
- [95] Richardson J.S., Getzoff E.D. & Richardson, D.C. (1978). The β bulge: A common small unit of nonrepetitive protein structure. *Proc. Natl. Acad. Sci. USA.*, **75**, 2574-2578.

- [96] Milnerwhite E.J., Ross B.M., Ismail R., Belhad J., Mostefa K. & Poet R. (1988) One type of gamma-turn, rather than the other gives rise to chain-reversal in proteins. *J. Mol. Biol.*, **204**, 777-782.
- [97] Sibanda B.L., Blundell T.L. & Thornton J.M. (1989) Conformation of beta-hairpins in proteins structure – a systematic classification with applications to modeling by homology, electron-density fitting and protein engineering. *J. Mol. Biol.*, **206**, 759-777.
- [98] Bradford, M.M. (1976). A rapid and sensitive for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, **72**, 248-254.
- [99] Docherty K., Rhodes C.J., Taylor N.A., Shennan K.I.J. & Hutton J.C. (1989). Proinsulin endopeptidase substrate specificities defined by site-directed mutagenesis of proinsulin. *J. Biol. Chem.*, **264**, 18335-18339.
- [100] Gross D.J., Villa-Komaroff L., Kahn C.R. Weir G.C. & Halban P.A. (1989). Deletion of a highly conserved tetrapeptide sequence of the proinsulin connecting peptide (C-peptide) inhibits proinsulin to insulin conversion by transfected pituitary corticotroph (AtT20) cells. *J. Biol. Chem.*, **264**, 21486-21490.
- [101] Steiner D., Bell G. & Tager H. (1995). Chemistry and biosynthesis of pancreatic protein hormones. In: *Endocrinology*, edited by L. DeGroot. Philadelphia, PA: Saunders, p 1296-1328.

LIVROS UTILIZADOS PARA CONSULTA

Abragam A. (1961). *Principles of Nuclear Magnetism*, Clarendon Press, Oxford.

Cavanagh J., Fairbrother W., Palmer III A.G. & Skelton N. (1995). *Protein NMR Spectroscopy: Principles and Practice*, Academic Press.

Claridge T.D.W. (1999). *High-Resolution NMR Techniques in Organic Chemistry*, Pergamon.

Ernst R.R., Bodenhausen G. & Wokaun A. (1987) *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon Press, Oxford.

Friebolin H. (1992). *Basic One and Two Dimensional NMR Spectroscopy*, GCH Germany.

Goldman M. (1988). *Quantum description of high-resolution NMR in liquids*, Clarendon.

Hausser K.H. & Kalbitzer H.R. (1991). *NMR in Medicine and Biology: Structure Determination, Tomography, in-vivo Spectroscopy*, Springer-Verlag.

Macomber R.S. (1998). *A Complete Introduction to Modern NMR Spectroscopy*, Wiley-Interscience.

Slichter C.P. (1978). *Principles of Magnetic Resonance*, Springer-Verlag.

Apêndice A

Deslocamentos químicos ^1H em HPr(H15A) de *S. aureus* em água, relativos ao DSS, à temperatura de 303K e pH=7,0. (*) Ressonâncias não encontradas.

Resíduo	$^1\text{H}_\text{N}$	$^1\text{H}_\alpha$	$^1\text{H}_\beta$	$^1\text{H}_\gamma$	$^1\text{H}_\delta$	$^1\text{H}_\epsilon$	$^1\text{H}_\eta$	$^1\text{H}_\zeta$
Met1		4.71	2.20	2.63		2.07		
			2.13	2.54				
Glu2	9.39	4.69	2.03	2.28				
			1.66	2.18				
Gln3	8.68	5.70	1.94	2.03		7.39		
			1.49	1.92		6.74		
Asn4	8.57	4.79	2.47		6.79			
			2.47		6.21			
Ser5	7.90	5.57	3.90					
			3.51					
Tyr6	8.89	4.84	2.97		6.87	6.57		
			2.28		6.87	6.57		
Val7	8.28	4.93	1.86	0.94				
				0.76				
Ile8	8.92	4.69	2.36	1.69	0.49			
				1.06				
				0.87				
Ile9	8.68	4.57	2.14	1.24	0.73			
				0.84				
				0.94				
Asp10	6.59	4.57	2.78					
			2.70					
Glu11	8.74	3.92	2.12	2.36				
			2.12	2.36				
Thr12	9.13	4.33	4.34	1.20				
Gly13	8.11	4.26						
		3.07						
Ile14	7.63	3.94	1.83	1.36	0.61			
				0.66				
				0.61				
Ala15	7.53	4.49	1.46					
Ala16	8.31	4.09	1.47					
Arg17	8.79	4.35	2.01	1.73	3.32	(*)	(*)	
			1.95	1.52	3.17			
Pro18		4.21	2.24	2.40	3.74			
			1.78	2.24	3.67			
Ala19	8.21	3.71	1.31					
Thr20	8.05	3.79	4.42	1.25				

Resíduo	$^1\text{H}_\text{N}$	$^1\text{H}_\alpha$	$^1\text{H}_\beta$	$^1\text{H}_\gamma$	$^1\text{H}_\delta$	$^1\text{H}_\epsilon$	$^1\text{H}_\eta$	$^1\text{H}_\zeta$
Met21	7.77	4.34	2.12	2.66		2.12		
			1.98	2.55				
Leu22	8.55	3.92	1.83	1.68	0.69			
			1.61		0.69			
Val23	8.28	4.19	2.05	1.02				
				0.84				
Gln24	8.58	4.03	2.26	2.57		7.39		
			2.13	2.39		6.75		
Thr25	7.90	3.92	4.34	1.25				
Ala26	8.63	3.98	1.36					
Ser27	8.09	4.13	4.05					
			4.05					
Lys28	7.38	4.01	1.47	1.29	1.57	2.85		
			1.39	1.29	1.57	2.85		
Phe29	7.45	4.54	3.44		7.40	7.19		7.09
			2.78		7.40	7.19		
Asp30	11.31	4.52	2.66					
			2.63					
Ser31	9.18	4.33	3.48	5.61				
			3.29					
Asp32	8.19	4.67	2.80					
			2.80					
Ile33	9.19	5.42	1.48	1.68	0.87			
				0.64				
				0.82				
Gln34	9.12	5.36	2.06	2.29		7.90		
			1.76	2.15		6.31		
Leu35	9.16	5.16	1.61	1.54	0.82			
			1.46		0.75			
Glu36	9.89	5.77	2.40	2.14				
			1.88	1.96				
Tyr37	8.84	5.15	3.14		7.26	6.99		
			2.70		7.26	6.99		
Asn38	9.44	4.15	2.76		7.10			
			1.53		6.56			
Gly39	8.74	4.07						
		3.67						
Lys40	7.82	4.60	1.86	1.45	1.74	3.00		
			1.86	1.37	1.74	3.00		
Lys41	8.33	5.55	1.74	1.41	1.41	2.57		
			1.58	1.22	1.41	2.57		
Val42	9.13	4.74	2.10	0.87				
				0.79				
Asn43	8.44	4.65	2.99		7.68			
			2.99		7.05			

Resíduo	$^1\text{H}_\text{N}$	$^1\text{H}_\alpha$	$^1\text{H}_\beta$	$^1\text{H}_\gamma$	$^1\text{H}_\delta$	$^1\text{H}_\epsilon$	$^1\text{H}_\eta$	$^1\text{H}_\zeta$
Leu44	8.97	4.07	1.47	1.93	0.93			
			1.10		0.85			
Lys45	7.96	4.29	2.20	1.38	1.92	3.49		
			1.68	0.95	1.79	3.19		
Ser46	7.60	4.82	4.26					
			3.79					
Ile47	9.04	4.04	1.90	1.39	0.90			
				1.39				
				1.02				
Met48	8.22	4.21	2.02	2.67		2.01		
			2.02	2.58				
Gly49	8.16	3.82						
		3.82						
Val50	8.28	3.50	2.03	1.07				
				0.79				
Met51	8.35	4.37	2.15	2.76		2.07		
			2.15	2.54				
Ser52	7.90	4.36	4.13					
			4.05					
Leu53	7.19	4.26	1.94	1.96	0.98			
			1.60		0.98			
Gly54	7.50	3.89						
		3.78						
Val55	8.20	2.73	1.66	0.46				
				0.46				
Gly56	7.65	4.41						
		3.64						
Lys57	7.94	3.64	1.80	1.34	1.65	3.03		
			1.54	1.09	1.56	3.03		
Asp58	8.81	4.26	3.12					
			2.88					
Ala59	7.57	4.35	1.32					
Glu60	8.52	5.25	2.17	2.31				
			2.02	2.31				
Ile61	8.88	5.52	1.95	1.69	0.64			
				1.03				
				0.90				
Thr62	8.64	5.15	3.82	0.52				
Ile63	8.99	5.17	1.75	1.61	0.69			
				1.13				
				0.80				
Tyr64	8.87	5.19	3.04		7.00	6.72		
			2.72		7.00	6.72		
Ala65	9.40	5.50	1.19					

Resíduo	$^1\text{H}_\text{N}$	$^1\text{H}_\alpha$	$^1\text{H}_\beta$	$^1\text{H}_\gamma$	$^1\text{H}_\delta$	$^1\text{H}_\epsilon$	$^1\text{H}_\eta$	$^1\text{H}_\zeta$
Asp66	8.87	5.57	2.95 2.55					
Gly67	10.30	4.83 3.98						
Ser68	9.00	4.12	3.97 3.97					
Asp69	7.53	5.13	3.51 2.79					
Glu70	7.16	3.79	2.40 1.77	2.76 2.17				
Ser71	8.25	4.18	3.92 3.85					
Asp72	7.65	4.31	2.76 2.76					
Ala73	8.30	2.11	0.72					
Ile74	7.59	3.67	1.81	1.45 1.27 0.93	0.63			
Gln75	7.33	4.01	2.13 2.13	2.46 2.41		7.55 6.84		
Ala76	7.96	4.21	1.60					
Ile77	8.24	3.60	1.93	1.53 1.10 0.69	0.71			
Ser78	8.61	4.11	4.20 4.20					
Asp79	8.03	4.45	2.85 2.72					
Val80	7.86	3.85	2.16	1.15 0.95				
Leu81	8.01	3.74	1.75 1.36	1.49	0.49 -0.17			
Ser82	7.53	4.37	3.96 3.96					
Lys83	8.85	4.07	2.03 1.91	1.58 1.54	1.67 1.67	3.00 3.00		
Glu84	8.18	4.31	2.37 1.85	2.46 2.11				
Gly85	7.85	4.07 3.98						
Leu86	8.12	4.23	1.98 1.43	1.68	0.64 0.49			
Thr87	7.33	4.91	4.57	1.07				
Lys88	8.50	4.16	1.82 1.82	1.49 1.49	1.74 1.74	3.02 3.02		

Deslocamentos químicos ^{13}C em HPr(H15A) de *S. aureus* em água, relativos ao amônio líquido conforme calculado da frequência de ressonância de próton no DSS; T=303K, pH=7,0.

Resíduo	$^{13}\text{C}_\alpha$	$^{13}\text{C}_\beta$	$^{13}\text{C}_\gamma$	$^{13}\text{C}_\delta$	$^{13}\text{C}_\epsilon$	$^{13}\text{C}_\eta$	$^{13}\text{C}_\zeta$	$^{13}\text{C}_\text{O}$
Met1	54,48	33,23	31,21		16,86			172,23
Glu2	55,37	33,64	35,42					173,83
Gln3	53,54	33,96	32,75					174,28
Asn4	52,24	42,13						172,58
Ser5	57,50	66,32						171,96
Tyr6	56,12	42,78						174,37
Val7	59,10	33,56	21,50					175,97
			20,78					
Ile8	57,42	35,10	26,20	8,28				177,40
			17,86					
Ile9	60,78	39,06	26,52	13,93				175,97
			19,00					
Asp10	54,70	44,32						176,87
Glu11	58,72	30,08	36,07					177,40
Thr12	62,20	68,57	21,99					174,81
Gly13	45,34							173,12
Ile14	61,31	36,07	27,01	12,04				173,65
			17,62					
Ala15	50,95	19,73						175,53
Ala16	56,05	18,51						179,10
Arg17	58,07	26,52	27,41	42,94				
Pro18	65,23	31,37	27,90	49,41				178,12
Ala19	56,05	18,67						178,65
Thr20	67,05	68,55	21,10					176,60
Met21	57,59	31,70	31,78		16,82			179,63
Leu22	59,28	42,30	26,60	24,42				177,58
				24,50				
Val23	65,19	31,86	22,07					179,99
			20,94					
Gln24	59,12	28,46	34,37					179,01
Thr25	66,32	69,06	21,02					175,35
Ala26	55,48	18,51						178,39
Ser27	61,31	63,41						174,01
Lys28	57,18	32,02	24,50	29,19	41,97			175,80
Phe29	58,32	40,35						174,64
Asp30	56,25	41,73						178,65
Ser31	61,47	65,44						172,32
Asp32	53,78	42,30						176,60
Ile33	60,17	40,68	27,01	15,52				174,81
			19,48					
Gln34	53,86	33,44	33,72					173,65
Leu35	53,86	45,77	27,90	25,95				173,30

Resíduo	$^{13}\text{C}_\alpha$	$^{13}\text{C}_\beta$	$^{13}\text{C}_\gamma$	$^{13}\text{C}_\delta$	$^{13}\text{C}_\epsilon$	$^{13}\text{C}_\eta$	$^{13}\text{C}_\zeta$	$^{13}\text{C}_\text{O}$
				26,28				
Glu36	53,86	35,18	36,39					174,55
Tyr37	55,97	42,21						174,64
Asn38	52,97	36,71						175,26
Gly39	45,53							173,65
Lys40	54,96	34,61	24,98	29,19	42,21			174,01
Lys41	55,08	35,58	25,63	29,51	41,57			176,69
Val42	58,71	35,90	20,94					173,83
			18,75					
Asn43	52,86	38,66						176,33
Leu44	55,00	43,02	27,81	25,55				173,56
				26,84				
Lys45	56,18	31,78	27,17	30,24	42,46			176,06
Ser46	54,96	64,14						174,37
Ile47	62,84	38,09	29,27	14,55				176,33
			17,86					
Met48	62,77	31,54	32,18		17,05			179,63
Gly49	47,07							176,33
Val50	67,05	31,94	24,26					178,39
			20,86					
Met51	57,18	30,73	31,94		15,82			178,92
Ser52	60,42	63,41						175,17
Leu53	56,29	43,99	27,81	26,36				177,85
				23,77				
Gly54	47,39							175,08
Val55	64,46	30,89	22,56					176,33
			21,26					
Gly56	43,72							171,51
Lys57	58,15	32,75	24,66	29,43	42,21			177,22
Asp58	56,86	38,57						174,73
Ala59	52,65	22,15						176,06
Glu60	55,16	30,81	36,96					176,51
Ile61	58,88	41,97	26,28	14,06				175,35
			19,08					
Thr62	61,95	71,66	21,02					172,14
Ile63	59,28	40,11	28,62	13,41				172,94
			17,38					
Tyr64	55,72	41,49						174,99
Ala65	50,39	23,04						174,64
Asp66	52,17	44,88						175,17
Gly67	45,93							175,26
Ser68	62,03	62,52						175,62
Asp69	51,36	40,03						176,78
Glu70	59,61	27,81	34,04					176,78
Ser71	61,63	62,28						177,14

Resíduo	$^{13}\text{C}_\alpha$	$^{13}\text{C}_\beta$	$^{13}\text{C}_\gamma$	$^{13}\text{C}_\delta$	$^{13}\text{C}_\epsilon$	$^{13}\text{C}_\eta$	$^{13}\text{C}_\zeta$	$^{13}\text{C}_\text{o}$
Asp72	56,70	40,52						178,21
Ala73	54,03	18,19						178,74
Ile74	62,12	36,71	28,30 17,78	10,91				177,94
Gln75	58,96	28,46	33,88					177,94
Ala76	54,92	19,16						181,06
Ile77	64,06	36,96	29,43 17,62	13,17				177,05
Ser78	62,76	62,76						176,42
Asp79	57,59	40,92						177,85
Val80	67,54	31,37	23,37 21,42					177,05
Leu81	58,64	41,16	26,44	24,26 24,98				178,65
Ser82	60,83	63,09						179,19
Lys83	59,53	32,43	25,47	28,87	42,21			179,46
Glu84	55,64	29,43	36,96					176,60
Gly85	45,86							174,10
Leu86	57,18	43,02	28,14	24,42 26,12				176,51
Thr87	61,04	71,88	22,39					173,65
Lys88	58,39	33,32	24,98	29,19	42,21			

Deslocamentos químicos ^{15}N em HPr(H15A) de *S. aureus* em água, relativos ao amônio líquido conforme calculado da frequência de ressonância de próton no DSS; T=303K, pH=7,0.

Resíduo	^{15}NH (ppm)	$^{15}\text{N}_{\delta/\epsilon}$ (ppm)	Resíduo	^{15}NH (ppm)	$^{15}\text{N}_{\delta/\epsilon}$ (ppm)
Met1			Lys45	110,49	
Glu2	127,96		Ser46	114,15	
Gln3	118,19	110,75	Ile47	129,16	
Asn4	118,45	109,63	Met48	119,92	
Ser5	117,16		Gly49	108,00	
Tyr6	118,58		Val50	122,93	
Val7	121,25		Met51	116,65	
Ile8	127,83		Ser52	114,28	
Ile9	120,30		Leu53	121,08	
Asp10	122,93		Gly54	105,93	
Glu11	124,73		Val55	119,83	
Thr12	111,83		Gly56	118,11	
Gly13	109,85		Lys57	117,98	
Ile14	121,08		Asp58	121,64	
Ala15	128,65		Ala59	122,07	
Ala16	120,69		Glu60	121,51	
Arg17	118,24		Ile61	117,25	

Resíduo	^{15}NH (ppm)	$^{15}\text{N}_{\delta/\epsilon}$ (ppm)	Resíduo	^{15}NH (ppm)	$^{15}\text{N}_{\delta/\epsilon}$ (ppm)
Pro18			Thr62	117,76	
Ala19	118,62		Ile63	126,32	
Thr20	113,08		Tyr64	123,44	
Met21	120,52		Ala65	124,04	
Leu22	123,74		Asp66	122,32	
Val23	120,17		Gly67	118,28	
Gln24	120,69	111,35	Ser68	115,14	
Thr25	116,00		Asp69	118,28	
Ala26	122,02		Glu70	118,24	
Ser27	107,70		Ser71	115,79	
Lys28	119,53		Asp72	122,63	
Phe29	117,81		Ala73	123,40	
Asp30	125,68		Ile74	115,70	
Ser31	116,52		Gln75	120,69	113,94
Asp32	122,45		Ala76	121,55	
Ile33	128,52		Ile77	120,04	
Gln34	124,86	111,14	Ser78	116,77	
Leu35	124,22		Asp79	120,60	
Glu36	128,43		Val80	120,47	
Tyr37	125,25		Leu81	119,74	
Asn38	129,51	109,94	Ser82	111,74	
Gly39	103,27		Lys83	124,43	
Lys40	122,54		Glu84	114,24	
Lys41	122,15		Gly85	107,31	
Val42	116,13		Leu86	116,77	
Asn43	120,65	111,23	Thr87	106,32	
Leu44	129,42		Lys88	124,69	

Identificação estereoespecífica dos metilenos $\text{H}_{\beta 1}/\text{H}_{\beta 2}$ em HPr(H15A) de *S. aureus*.

Resíduo	Rotâmetro	$^1\text{H}_{\beta 1}$ (ppm)	$^1\text{H}_{\beta 2}$ (ppm)	Resíduo	Rotâmetro	$^1\text{H}_{\beta 1}$ (ppm)	$^1\text{H}_{\beta 2}$ (ppm)
Gln3	g^2g^3	1.49	1.94	Glu36	g^2t^3	1.88	2.40
Tyr6	t^2g^3	2.28	2.97	Tyr37	g^2t^3	3.14	2.70
Asp10	g^2t^3	2.78	2.70	Asn38	g^2g^3	2.76	1.53
Arg17	g^2t^3	1.95	2.00	Lys45	t^2g^3	1.68	2.20
Met21	g^2t^3	1.98	2.12	Lys57	g^2t^3	1.80	1.54
Leu22	g^2t^3	1.83	1.61	Tyr64	t^2g^3	2.71	3.04
Gln24	g^2t^3	2.13	2.26	Asp66	g^2t^3	2.55	2.95
Phe29	t^2g^3	2.78	3.44	Glu70	g^2g^3	1.77	2.40
Asp30	g^2t^3	2.66	2.63	Ser71	g^2t^3	3.92	3.85
Ser31	t^2g^3	3.29	3.48	Asp79	g^2t^3	2.71	2.85
Gln34	g^2t^3	2.06	1.76	Leu81	t^2g^3	1.75	1.35
Leu35	g^2t^3	1.46	1.61	Glu84	g^2t^3	2.37	1.85

Apêndice B

Deslocamentos químicos ^1H em HPr(wt) de *S. aureus* em água, relativos ao DSS, à temperatura de 303K e pH=7,0.

Resíduo	$^1\text{H}_\text{N}$	$^1\text{H}_\alpha$	$^1\text{H}_\beta$	$^1\text{H}_\gamma$	$^1\text{H}_\delta$	$^1\text{H}_\epsilon$	$^1\text{H}_\eta$	$^1\text{H}_\zeta$
Met1		4.68	2.19	2.61		2.05		
			2.12	2.51				
Glu2	9.36	4.74	2.01	2.26				
			1.62	2.18				
Gln3	8.66	5.69	1.93	2.02		7.37		
			1.47	1.91		6.72		
Asn4	8.54	4.77	2.46		6.77			
			2.46		6.19			
Ser5	7.88	5.56	3.88					
			3.49					
Tyr6	8.88	4.84	2.97		6.85	6.55		
			2.26		6.85	6.55		
Val7	8.24	4.93	1.84	0.93				
				0.76				
Ile8	8.90	4.67	2.37	1.69	0.48			
				1.08				
				0.90				
Ile9	8.72	4.57	2.12	1.20	0.71			
				0.79				
				0.93				
Asp10	6.54	4.53	2.81					
			2.74					
Glu11	8.72	3.89	2.08	2.35				
			2.08	2.35				
Thr12	9.28	4.29	3.97	1.06				
Gly13	8.20	4.38						
		3.12						
Ile14	7.56	3.88	1.87	1.41	0.56			
				0.65				
				0.57				
His15	6.75	4.75	3.34		7.26	7.80		
			3.04					
Ala16	8.46	4.04	1.50					
Arg17	8.92	4.29	2.00	1.74	3.32	7.68	6.89	
			1.94	1.50	3.22			
Pro18		4.18	2.15	1.69	3.14			
			1.60	1.69	3.14			
Ala19	8.05	3.67	1.23					
Thr20	8.07	3.75	4.38	1.22				

Resíduo	$^1\text{H}_\text{N}$	$^1\text{H}_\alpha$	$^1\text{H}_\beta$	$^1\text{H}_\gamma$	$^1\text{H}_\delta$	$^1\text{H}_\epsilon$	$^1\text{H}_\eta$	$^1\text{H}_\zeta$
Met21	7.51	4.34	2.05	2.62	2.05			
			1.92	2.55				
Leu22	8.50	3.89	1.73	1.66	0.68			
			1.61		0.68			
Val23	8.26	4.14	2.03	0.99				
				0.83				
Gln24	8.50	4.00	2.24	2.54		7.35		
			2.11	2.36		6.74		
Thr25	7.89	3.90	4.33	1.23				
Ala26	8.59	3.97	1.34					
Ser27	8.06	4.11	4.04					
			4.04					
Lys28	7.37	4.00	1.45	1.27	1.55	2.85		
			1.36	1.27	1.55	2.85		
Phe29	7.42	4.53	3.41		7.38	7.16		7.07
			2.76		7.38	7.16		
Asp30	11.29	4.50	2.65					
			2.60					
Ser31	9.15	4.30	3.45	5.58				
			3.26					
Asp32	8.18	4.65	2.78					
			2.78					
Ile33	9.17	5.39	1.45	1.66	0.86			
				0.62				
				0.80				
Gln34	9.08	5.34	2.04	2.28		7.89		
			1.75	2.12		6.28		
Leu35	9.13	5.14	1.57	1.53	0.79			
			1.45		0.73			
Glu36	9.89	5.76	2.39	2.13				
			1.86	1.95				
Tyr37	8.82	5.13	3.12		7.25	6.97		
			2.67		7.25	6.97		
Asn38	9.40	4.12	2.73		7.09			
			1.50		6.55			
Gly39	8.71	4.05						
		3.66						
Lys40	7.82	4.59	1.84	1.42	1.73	3.01		
			1.84	1.35	1.73	3.01		
Lys41	8.31	5.54	1.73	1.39	1.39	2.54		
			1.56	1.21	1.39	2.54		
Val42	9.12	4.72	2.09	0.86				
				0.76				
Asn43	8.42	4.64	2.96		7.66			
			2.96		7.03			

Resíduo	$^1\text{H}_\text{N}$	$^1\text{H}_\alpha$	$^1\text{H}_\beta$	$^1\text{H}_\gamma$	$^1\text{H}_\delta$	$^1\text{H}_\epsilon$	$^1\text{H}_\eta$	$^1\text{H}_\zeta$
Leu44	8.94	4.07	1.46	1.92	0.90			
			1.08		0.82			
Lys45	7.93	4.24	2.18	1.39	1.90	3.47		
			1.65	0.93	1.76	3.18		
Ser46	7.58	4.78	4.22					
			3.78					
Ile47	9.02	4.01	1.86	1.34	0.86			
				1.34				
				1.00				
Met48	8.17	4.19	2.01	2.65		2.00		
			2.01	2.55				
Gly49	8.14	3.81						
		3.81						
Val50	8.26	3.48	2.01	1.03				
				0.76				
Met51	8.25	4.38	2.13	2.74		2.05		
			2.13	2.48				
Ser52	7.90	4.34	4.12					
			4.03					
Leu53	7.15	4.24	1.90	1.93	0.95			
			1.58		0.95			
Gly54	7.51	3.88						
		3.77						
Val55	8.14	2.73	1.64	0.48				
				0.48				
Gly56	7.72	4.45						
		3.64						
Lys57	7.91	3.66	1.84	1.35	1.67	3.02		
			1.50	1.12	1.56	3.02		
Asp58	8.79	4.24	3.12					
			2.87					
Ala59	7.55	4.34	1.32					
Glu60	8.50	5.25	2.15	2.31				
			1.99	2.31				
Ile61	8.85	5.49	1.92	1.69	0.63			
				1.00				
				0.89				
Thr62	8.62	5.13	3.81	0.50				
Ile63	8.97	5.15	1.73	1.58	0.68			
				1.10				
				0.78				
Tyr64	8.85	5.17	3.03		6.98	6.69		
			2.70		6.98	6.69		
Ala65	9.38	5.47	1.16					

Resíduo	$^1\text{H}_\text{N}$	$^1\text{H}_\alpha$	$^1\text{H}_\beta$	$^1\text{H}_\gamma$	$^1\text{H}_\delta$	$^1\text{H}_\epsilon$	$^1\text{H}_\eta$	$^1\text{H}_\zeta$
Asp66	8.85	5.55	2.93 2.54					
Gly67	10.29	4.80 3.97						
Ser68	8.98	4.11	3.96 3.96					
Asp69	7.51	5.11	3.48 2.77					
Glu70	7.14	3.77	2.39 1.76	2.74 2.16				
Ser71	8.24	4.16	3.90 3.83					
Asp72	7.64	4.29	2.73 2.73					
Ala73	8.29	2.09	0.70					
Ile74	7.59	3.66	1.79	1.42 1.24 0.91	0.61			
Gln75	7.29	4.00	2.10 2.10	2.44 2.40		7.52 6.82		
Ala76	7.92	4.19	1.57					
Ile77	8.26	3.59	1.92	1.53 1.09 0.67	0.68			
Ser78	8.59	4.11	4.19 4.19					
Asp79	7.95	4.44	2.82 2.70					
Val80	7.86	3.85	2.13	1.10 0.94				
Leu81	8.06	3.68	1.71 1.36	1.46	0.46 -0.21			
Ser82	7.40	4.37	3.94 3.94					
Lys83	8.81	4.04	2.01 1.88	1.57 1.50	1.66 1.66	2.98 2.98		
Glu84	8.22	4.29	2.35 1.81	2.43 2.04				
Gly85	7.83	4.05 3.99						
Leu86	8.09	4.29	1.98 1.36	1.62	0.82 0.76			
Thr87	7.35	4.89	4.59	1.04				
Lys88	8.51	4.14	1.82 1.82	1.49 1.49	1.73 1.73	3.01 3.01		

Apêndice C

Deslocamentos químicos ^1H no peptídeo-C da proinsulina humana em água (600 MHz, 283K, pH=7,0), relativos ao DSS. (*) ressonância não identificada.

Resíduo	$^1\text{H}_\text{N}$	$^1\text{H}_\alpha$	$^1\text{H}_\beta$	$^1\text{H}_\gamma$	$^1\text{H}_\delta$	$^1\text{H}_\epsilon$	$^1\text{H}_\eta$	$^1\text{H}_\zeta$
Glu1		3.99	2.15	2.39				
			2.04	2.39				
Ala2	(*)	4.30	1.41					
Glu3	8.71	4.22	2.02	2.27				
			1.94	2.27				
Asp4	8.46	4.57	2.72					
			2.60					
Leu5	8.25	4.31	1.64	1.58	0.91			
			1.58		0.85			
Gln6	8.47	4.33	2.10	2.35		7.65		
			1.99	2.35		6.91		
Val7	8.22	4.08	2.08	0.94				
				0.94				
Gly8	8.62	3.95						
		3.95						
Gln9	8.31	4.35	2.09	2.34		7.66		
			1.98	2.34		6.92		
Val10	8.31	4.08	2.04	0.92				
				0.92				
Glu11	8.64	4.30	2.01	2.25				
			1.98	2.25				
Leu12	8.52	4.34	1.67	1.67	0.92			
			1.56		0.85			
Gly13	8.61	3.97						
		3.97						
Gly14	8.39	3.98						
		3.98						
Gly15	8.31	4.12						
		4.12						
Pro16		4.43	2.28	2.04	3.67			
			1.98	2.04	3.62			
Gly17	8.66	3.95						
		3.95						
Ala18	8.26	4.29	1.39					
Gly19	8.55	3.97						
		3.97						
Ser20	8.19	4.44	3.86					
			3.86					

Resíduo	$^1\text{H}_\text{N}$	$^1\text{H}_\alpha$	$^1\text{H}_\beta$	$^1\text{H}_\gamma$	$^1\text{H}_\delta$	$^1\text{H}_\epsilon$	$^1\text{H}_\eta$	$^1\text{H}_\zeta$
Leu21	8.38	4.37	1.63 1.58	1.63	0.91 0.85			
Gln22	8.39	4.60	2.09 1.92	2.38 2.38		7.61 6.93		
Pro23		4.40	2.30 1.89	2.01 2.01	3.77 3.65			
Leu24	8.43	4.27	1.56 1.56	1.56	0.93 0.88			
Ala25	8.35	4.31	1.36					
Leu26	8.32	4.33	1.64 1.47	1.64	0.92 0.86			
Glu27	8.54	4.25	2.05 1.98	2.27 2.27				
Gly28	8.56	3.98 3.98						
Ser29	8.24	4.45	3.88 3.88					
Leu30	8.43	4.40	1.65 1.65	1.65	0.92 0.86			
Gln31	7.97	4.14	2.12 1.91	2.29 2.29			7.58 6.87	

Deslocamentos químicos ^1H no peptídeo-C da proinsulina humana em 50% de água, 50% de TFE-d₂ (600 MHz, 283K, pH=7,0), relativos ao DSS.

Resíduo	$^1\text{H}_\text{N}$	$^1\text{H}_\alpha$	$^1\text{H}_\beta$	$^1\text{H}_\gamma$	$^1\text{H}_\delta$	$^1\text{H}_\epsilon$	$^1\text{H}_\eta$	$^1\text{H}_\zeta$
Glu1		4.07	2.31 2.14	2.52 2.52				
Ala2	8.99	4.23	1.50					
Glu3	9.13	4.16	2.06 2.06	2.43 2.33				
Asp4	8.25	4.55	2.78 2.74					
Leu5	7.82	4.32	1.79 1.66	1.67	0.97 0.90			
Gln6	8.11	4.35	2.21 2.11	2.44 2.40		7.41 6.71		
Val7	7.99	3.99	2.15	1.05 1.00				
Gly8	8.31	3.97 3.97						
Gln9	8.07	4.29	2.15 2.15	2.42 2.42		7.43 6.66		

XXXIV

Resíduo	$^1\text{H}_\text{N}$	$^1\text{H}_\alpha$	$^1\text{H}_\beta$	$^1\text{H}_\gamma$	$^1\text{H}_\delta$	$^1\text{H}_\epsilon$	$^1\text{H}_\eta$	$^1\text{H}_\zeta$
Val10	8.02	3.95	2.17	1.03 0.97				
Glu11	8.55	4.26	2.07 2.07	2.34 2.34				
Leu12	8.12	4.37	1.78 1.65	1.72	0.95 0.91			
Gly13	8.20	4.01 4.01						
Gly14	8.17	4.02 4.02						
Gly15	8.13	4.22 4.03						
Pro16		4.43	2.30 2.03	2.11 2.06	3.72 3.59			
Gly17	8.48	3.97 3.97						
Ala18	8.08	4.28	1.47					
Gly19	8.42	4.01 3.98						
Ser20	8.01	4.49	3.98 3.92					
Leu21	7.98	4.46	1.75 1.69	1.75	0.95 0.90			
Gln22	8.03	4.50	2.17 2.09	2.43 2.43		7.35 6.63		
Pro23		4.35	2.36 1.93	2.13 2.01	3.76 3.70			
Leu24	7.76	4.25	1.81 1.63	1.70	0.98 0.92			
Ala25	7.97	4.23	1.48					
Leu26	7.97	4.24	1.76 1.70	1.70	0.94 0.92			
Glu27	8.31	4.14	2.14 2.14	2.36 2.36				
Gly28	8.33	3.99 3.99						
Ser29	8.04	4.51	4.03 3.99					
Leu30	7.85	4.46	1.77 1.67	1.77	0.95 0.90			
Gln31	7.61	4.20	2.21 2.00	2.40 2.40		7.52 6.67		

Deslocamentos químicos ^1H no peptídeo-C da proinsulina humana em 20% de água, 80% de TFE- d_2 (600 MHz, 283K, pH=7,0), relativos ao DSS.

Resíduo	$^1\text{H}_\text{N}$	$^1\text{H}_\alpha$	$^1\text{H}_\beta$	$^1\text{H}_\gamma$	$^1\text{H}_\delta$	$^1\text{H}_\epsilon$	$^1\text{H}_\eta$	$^1\text{H}_\zeta$
Glu1		4.04	2.31	2.56				
			2.12	2.56				
Ala2	8.87	4.16	1.51					
Glu3	9.44	4.07	2.05	2.50				
			2.05	2.34				
Asp4	8.19	4.47	2.76					
			2.76					
Leu5	7.64	2.83	1.85	1.70	0.97			
			1.68		0.90			
Gln6	7.96	4.27	2.22	2.45		7.19		
			2.16	2.38		6.51		
Val7	7.94	3.90	2.16	1.07				
				1.00				
Gly8	8.17	3.94						
		3.94						
Gln9	7.96	4.23	2.19	2.47		7.27		
			2.19	2.40		6.45		
Val10	7.94	3.86	2.21	1.06				
				0.97				
Glu11	8.56	4.20	2.10	2.36				
			2.10	2.36				
Leu12	8.05	4.35	1.81	1.75	0.94			
			1.64		0.91			
Gly13	8.06	4.01						
		4.01						
Gly14	8.06	4.00						
		4.00						
Gly15	8.01	4.23						
		3.99						
Pro16		4.41	2.29	2.10	3.72			
			2.00	2.03	3.57			
Gly17	8.36	3.95						
		3.95						
Ala18	7.96	4.27	1.47					
Gly19	8.31	3.97						
		3.97						
Ser20	7.95	4.48	4.00					
			3.92					
Leu21	7.48	4.47	1.75	1.75	0.96			
			1.69		0.90			
Gln22	7.92	4.49	2.18	2.44		7.26		
			2.09	2.44		6.45		

Resíduo	$^1\text{H}_\text{N}$	$^1\text{H}_\alpha$	$^1\text{H}_\beta$	$^1\text{H}_\gamma$	$^1\text{H}_\delta$	$^1\text{H}_\epsilon$	$^1\text{H}_\eta$	$^1\text{H}_\zeta$
Pro23		4.35	2.36	2.13	3.75			
			1.93	2.01	3.71			
Leu24	7.63	4.24	1.81	1.70	0.98			
			1.64		0.91			
Ala25	7.90	4.22	1.49					
Leu26	7.95	4.22	1.77	1.70	1.06			
			1.70		0.98			
Glu27	8.37	4.12	2.14	2.37				
			2.14	2.37				
Gly28	8.28	4.01						
		3.93						
Ser29	7.94	4.48	4.06					
			4.00					
Leu30	7.72	4.43	1.77	1.77	0.95			
			1.66		0.90			
Gln31	7.54	4.19	2.19	2.40		7.36		
			1.99	2.40		6.44		