

JANELLE SAUVAGEAU

**ATTRIBUTION ET CARACTÉRISTIQUES DE  
LIAISON DU DOMAINE TANDEM PDZ2/3  
DE PTP-BL PAR RMN**

Mémoire présenté  
à la Faculté des études supérieures de l'Université Laval  
dans le cadre du programme de maîtrise en chimie  
pour l'obtention du grade de Maître ès sciences (M.Sc.)

FACULTÉ DES SCIENCES ET DE GÉNIE  
UNIVERSITÉ LAVAL  
QUÉBEC

2006

© Janelle Sauvageau, 2006

**À mon grand-père, Normand Sauvageau,  
et ma grand-mère, Alma Sauvageau,  
pour m'avoir enseigné ce qu'est le courage**

## Résumé

La protéine PTP-BL possède, entre autres, 5 domaines PDZ dont deux ont été étudiés en détail. Il s'agit des domaines PDZ2 et PDZ3. Les signaux RMN de ces domaines, formant une chaîne de 252 acides aminés et possédant un poids moléculaire de 26,4 kDa, ont été attribués. L'attribution du squelette des deux domaines a été effectuée à 88%. De plus, les caractéristiques de liaison du domaine tandem PDZ2/3 ont été étudiées plus en profondeur. PDZ2 se lie, entre autres, à la protéine APC tandis que PDZ3 se lie à la protéine PRK2. Deux peptides dérivés de ces protéines ont été utilisés afin de titrer le domaine tandem PDZ2/3. Les deux domaines liaient les peptides choisis. Des perturbations ont été observées dans PDZ3 et dans le lieu lors de la titration du domaine tandem PDZ2/3 avec APC et d'autres perturbations ont été observées pour ce qui est de PDZ2 et du lieu lors de la titration avec le peptide dérivé de PRK2. Cela indiquerait une certaine coopérativité entre les deux domaines et le lieu lors de la liaison avec les peptides.

## Remerciements

Je voudrais remercier chaleureusement ma directrice de maîtrise, Michèle Auger, pour son soutien et ses encouragements constants. Sans elle, ce projet n'aurait jamais eu lieu. Les membres de son groupe m'ont aussi très bien accueillie malgré mon arrivée tardive à Québec et je tenais aussi à leur faire part de mes remerciements. Je tenais aussi à remercier le groupe à la Ruhr-Universität Bochum en Allemagne, où les recherches ont été effectuées, qui m'ont aussi très bien accueillie et mon superviseur Raphael Stoll. Mes reconnaissances vont aussi à Stéphane Gagné et son groupe pour les bons conseils qu'ils m'ont prodigués. Ma famille et mes amis m'ont aussi beaucoup soutenue, en particulier mon père, André-Guy Sauvageau.



## Table des matières

<b>Résumé</b>	<b>iii</b>
<b>Remerciements</b>	<b>iv</b>
<b>Table des matières</b>	<b>v</b>
<b>Liste des figures</b>	<b>vii</b>
<b>Liste des tableaux</b>	<b>ix</b>
<b>Abréviations</b>	<b>x</b>
Acides aminés communs	x
Protéines	xi
Autres	xii
<b>Chapitre 1 : Introduction</b>	<b>1</b>
<b>1.1 PTP-BL</b>	<b>1</b>
<b>1.2 Rôle des phosphatases dans les systèmes biologiques</b>	<b>4</b>
<b>1.3 Description des domaines PDZ étudiés</b>	<b>4</b>
1.3.1 Le domaines tandem PDZ2/3	6
1.3.2 PDZ2	8
1.3.2.1 APC	9
1.3.3 PDZ3	10
1.3.3.1 PRK2	10
<b>1.4 Objectifs des recherches effectuées</b>	<b>11</b>
<b>Chapitre 2 : Matériel et Méthodes</b>	<b>13</b>
<b>2.1 Matériel</b>	<b>13</b>
<b>2.2 Spectres 2D et 3D</b>	<b>14</b>
<b>2.3 WATERGATE</b>	<b>16</b>
<b>2.4 HSQC</b>	<b>17</b>
2.4.1 INEPT	18
<b>2.5 HNCA</b>	<b>19</b>
<b>2.6 HN(CO)CA</b>	<b>22</b>
<b>2.7 CBCA(CO)NH</b>	<b>23</b>
<b>2.8 HNCO</b>	<b>24</b>
<b>2.9 NOESY-HSQC</b>	<b>25</b>
<b>2.10 Programmes utilisés</b>	<b>26</b>
2.10.1 Attribution semi-automatique	26
2.10.1.1 Monte	26
2.10.1.2 Mars	27
<b>2.11 Titration</b>	<b>28</b>
<b>2.12 Acides aminés sélectivement marqués</b>	<b>30</b>
<b>Chapitre 3 : Attribution du squelette du domaine tandem PDZ2/3 de PTP-BL</b>	<b>31</b>
<b>3.1 Le repliement des protéines</b>	<b>31</b>
<b>3.2 Attribution</b>	<b>32</b>
3.2.1 Spectre 1D du proton	32
3.2.2 Spectres 2D	34
3.2.2.1 HSQC du domaine tandem PDZ2/3	35
3.2.2.2 HSQC de PDZ2	38

3.2.2.3	HSQCs d'acides aminés spécifiquement marqués	41
3.2.3	Spectres 3D	44
3.2.3.1	HNCA	44
3.2.3.2	HN(CO)CA	46
3.2.3.3	CBCA(CO)NH	46
3.3	Attribution semi-automatique	47
3.4	Détermination de la structure du domaine tandem PDZ2/3 par RMN	47
<b>Chapitre 4 : Caractéristiques de liaison du domaine tandem PDZ2/3 de PTP-BL</b>		<b>49</b>
4.1	Interactions protéines-ligands	49
4.2	La bioinformatique	50
4.3	Titration	52
4.4	Prédiction de la structure quaternaire du domaine tandem PDZ2/3	60
<b>Chapitre 5 : Conclusions</b>		<b>64</b>
<b>Bibliographie</b>		<b>67</b>

## Liste des figures

<b>Figure 1.1</b>	Illustration schématique de PTP-BL	2
<b>Figure 1.2</b>	Ligands de PTP-BL	3
<b>Figure 1.3</b>	Motif de liaison de PDZ2	5
<b>Figure 1.4</b>	Séquence d'acides aminés du domaine tandem PDZ2/3	7
<b>Figure 1.5</b>	Alignement des domaines PDZ2 et PDZ3	7
<b>Figure 1.6</b>	Structure de PDZ2 déterminée par RMN	8
<b>Figure 2.1</b>	Schéma sommaire des séquences d'impulsions 2D et 3D	15
<b>Figure 2.2</b>	Exemple de séquence d'impulsions WATERGATE	16
<b>Figure 2.3</b>	Séquence d'impulsions HSQC	17
<b>Figure 2.4</b>	Corrélations observées dans un HNCA	19
<b>Figure 2.5</b>	Séquence d'impulsions HNCA	19
<b>Figure 2.6</b>	Corrélation observée dans un HN(CO)CA	22
<b>Figure 2.7</b>	Séquence d'impulsions HN(CO)CA	27
<b>Figure 2.8</b>	Corrélation observée pour le CBCA(CO)NH	23
<b>Figure 2.9</b>	Séquence d'impulsions CBCA(CO)NH	23
<b>Figure 2.10</b>	Corrélation observée dans le HNCO	24
<b>Figure 2.11</b>	Séquence d'impulsions HNCO	24
<b>Figure 2.12</b>	Séquence d'impulsions NOESY-HSQC	25
<b>Figure 3.1</b>	Spectre 1D du domaine tandem PDZ2/3 de PTP-BL	33
<b>Figure 3.2</b>	Spectres 1D du domaine tandem PDZ2/3 et de PDZ2	34
<b>Figure 3.3</b>	Résidus attribués sur la séquence du domaine tandem PDZ2/3	36
<b>Figure 3.4</b>	Attribution du HSQC du domaine tandem PDZ2/3 de PTP-BL	37
<b>Figure 3.5</b>	Spectre HSQC de PDZ2	39
<b>Figure 3.6</b>	Comparaison du domaine PDZ2 avec PDZ2 dans le domaine tandem PDZ2/3	40
<b>Figure 3.7</b>	Changements observés sur le domaine PDZ2 lors de la comparaison des déplacements chimiques avec le domaine tandem PDZ2/3	41
<b>Figure 3.8</b>	Spectres contenant des acides aminés spécifiquement marqués	43
<b>Figure 3.9</b>	Bandes peu chevauchées représentant l'attribution séquentielle à partir du HNCA et du HNCOCA	45
<b>Figure 3.10</b>	Exemple de raies très chevauchées	45
<b>Figure 3.11</b>	Exemple de raies pour le CBCACONH	46
<b>Figure 3.12</b>	Exemple du spectre NOESY-HSQC	48
<b>Figure 4.1</b>	Prédiction de la structure 2D de PDZ3	51
<b>Figure 4.2</b>	Structure 3D prédite de PDZ3 incluant les structures secondaires prédites	52
<b>Figure 4.3</b>	Exemple de la première étape de la titrage du domaine tandem PDZ2/3 marqué à $^{15}\text{N}$ par le ligand dérivé d'APC non marqué	53
<b>Figure 4.4</b>	Perturbations observées sur le domaine tandem PDZ2/3 de PTP-BL lors de l'ajout du peptide dérivé d'APC	54
<b>Figure 4.5</b>	Schéma des perturbations observées sur PDZ2	55

<b>Figure 4.6</b>	Changements observés sur PDZ3 lors de l'ajout du peptide dérivé d'APC dans le milieu	57
<b>Figure 4.7</b>	Perturbations observées lors de l'ajout du peptide dérivé de PRK2 sur le domaine tandem PDZ2/3	58
<b>Figure 4.8</b>	Perturbations induites par le peptide dérivé de PRK2 sur PDZ3	59
<b>Figure 4.9</b>	Perturbations induites par le peptide dérivé de PRK2 sur PDZ2	60
<b>Figure 4.10</b>	Comparaison de quelques signaux de la titration du domaine PDZ2 marqué à $^{15}\text{N}$ avec PDZ3 non-marqué	61
<b>Figure 4.11</b>	Structure secondaire prédite du lieu	62

## Liste des tableaux

<b>Tableau 1.1</b>	Ligands utilisés lors des titrations	9
<b>Tableau 2.1</b>	Déplacements chimiques moyens des $C_\alpha$ et $C_\beta$	21
<b>Tableau 3.1</b>	Résumé du métabolisme des acides aminés choisis pour effectuer des spectres spécifiquement marqués	42
<b>Tableau 4.1</b>	Tableau récapitulatif des titrations	56

**Abréviations**  
**Acides aminés communs**

<b>Nom</b>	<b>Abréviation à trois lettres</b>	<b>Abréviation à une lettre</b>	<b>Caractère chimique</b>
Glycine	Gly	G	Hydrophobe
Alanine	Ala	A	Hydrophobe
Valine	Val	V	Hydrophobe
Phénylalanine	Phe	F	Hydrophobe
Proline	Pro	P	Hydrophobe
Leucine	Leu	L	Hydrophobe
Isoleucine	Ile	I	Hydrophobe
Arginine	Arg	R	Hydrophile
Acide aspartique	Asp	D	Hydrophile
Acide glutamique	Glu	E	Hydrophile
Sérine	Ser	S	Hydrophile
Thréonine	Thr	T	Hydrophile
Cystéine	Cys	C	Hydrophile
Asparagine	Asn	N	Hydrophile
Glutamine	Gln	Q	Hydrophile
Histidine	His	H	Hydrophile
Lysine	Lys	K	Amphipathique
Tyrosine	Tyr	Y	Amphipathique
Méthionine	Met	M	Amphipathique
Tryptophane	Trp	W	Amphipathique

### Protéines

Abréviation	Nom
PTP-BL	Protein-tyrosine phosphatase basophil-like
PTP L1	Protein tyrosine phosphatase L1
hPTP1e	Human protein tyrosine phosphatase 1e
FAP-1	Fas-associated protein-tyrosine phosphatase 1
BP75	Bromoprotein 75
Ikappa B $\alpha$ ,	Specific association protein IkB $\alpha$
PRK2	Protein kinase C-related kinase 2
RIL	Reversion-induced LIM
APC	Adenomatous polyposis coli-protein
TRIP6	Tyroid receptor interacting protein-6
CRIP2	Cysteine-rich protein 2
PARG	Poly (ADP-ribose) glycohydrolase
LIM	Lin-11/Isl-1/Mec-3
PTN13_Mouse	Tyrosine protein phosphatase non-receptor type 13
PSD-95	Postsynaptic density protein 95
Dlg	Drosophila septate junction protein discs-large tumor supressor
ZO-1	Epithelial tight junction protein

## Autres

Abréviation	Nom
RMN	Résonance magnétique nucléaire
PDZ	PSD-95; Postsynaptic density-95, disc-large and zonulin-1
FERM	Four-point one, ezrin, radixin, moesin
ADP	Adénosine diphosphate
ATP	Adénosine triphosphate
ARM	Armadillo repeat
DHR	Discs-large homology region
GLGF	Gly-Leu-Gly-Phe
HSQC	Heteronuclear Single Quantum Coherence
INEPT	Insensitive Nuclei Enhanced by Polarization Transfer
WATERGATE	WATER suppression by GrAdient Tailored Excitation
GST	Glutathione S-transferase



# Chapitre 1

## Introduction

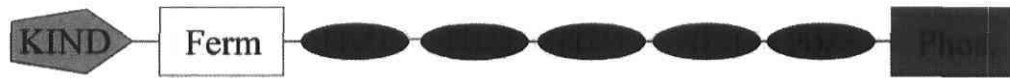
### 1.1 PTP-BL

Les recherches effectuées ont porté sur les domaines PDZ2 et PDZ3 de PTP-BL (aussi appelée PTN13\_Mouse, RIP,PTP36). Pour alléger le texte, le terme utilisé sera : le domaine tandem PDZ2/3. PTP-BL est une large phosphatase de 270 kDa. Ses homologues humains sont PTP L1 (Saras et al., 1994), hPTP1e (Banville et al., 1994), FAP-1 (Sato et al., 1995) et PTP-BAS (Maekawa et al., 1994). Son gène, Ptpn13, code pour 2453 acides aminés. Elle contient un domaine N-terminal, un domaine FERM, 5 domaines PDZ et un domaine C-terminal phosphatase. Les domaines PDZ sont notés de PDZ1 à PDZ5. Le domaine N-terminal consiste en l'extrémité possédant le groupe amine libre tandis que le groupe C-terminal contient le groupe carboxyle libre. La Figure 1.1 illustre sommairement PTP-BL.



**Figure 1.1** Illustration schématique de PTP-BL.

PTP-BL est impliquée dans des phénomènes aussi divers que la cytokinèse (Herrmann et al., 2003) et le réarrangement cytosquelettique (Banville et al., 1994). Les recherches effectuées sur PTP-BL ont surtout porté sur ses cinq domaines PDZ. En effet, l'identification des ligands des domaines PDZ permettrait de mieux comprendre le rôle biologique de PTP-BL. Les domaines PDZ se lient à plusieurs protéines, tel que décrit abondamment. En effet, PDZ1 interagit avec BP75 (Cuppen et al., 1999) et Ikappa B $\alpha$  (Maekawa et al., 1999). PDZ2 se lie avec TRIP6 (Murthy et al., 1999; Cuppen et al., 2000), RIL (Cuppen et al., 1998), Fas (Sato et al., 1995) et APC (Erdmann et al., 2000). PDZ3 s'attache avec PRK2 (Gross et al., 2001) et PDZ4 avec ephrin B1 (Palmer et al., 2002), CRIP2 (van Ham et al., 2003), TRIP6, RIL (Cuppen et al., 1998) et PARG (Saras et al., 1997). La Figure 1.2 représente un résumé des protéines interagissant avec les domaines PDZ de PTP-BL. Les protéines cochées se lient avec les différents domaines de PTP-BL. Dans cette figure, des ligands dérivés des protéines utilisés afin de faire les titrations sont marqués en rouge (APC et PRK2). Des peptides dérivés de ces protéines seront utilisés.



IκBα		×					
Bp75		×					
TRIP 6/ZRP-1			×		×		
RIL			×		×		
Ephrin B					×		
CRIP2					×		
FAS			×				
APC			×				
IRS2				×			
PARG					×		
PIP2	×						×
p75 <sup>NTR</sup>			×				
F-actin	×						

**Figure 1.2** Ligands de PTP-BL.

## 1.2 Rôle des phosphatases dans les systèmes biologiques

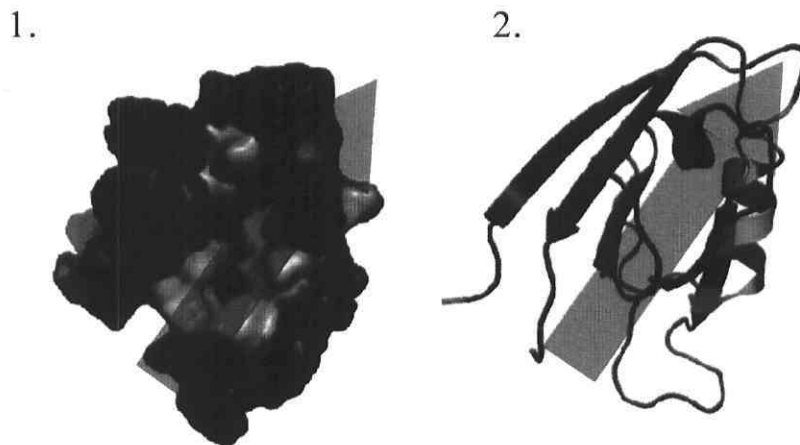
La phosphorylation consiste en la liaison covalente d'un ou de plusieurs groupes phosphate à une protéine. Ceci constitue un moyen très efficace afin de réguler l'activité des protéines cibles. La phosphorylation change la forme et la charge de la protéine modifiant ainsi son activité. Ce mécanisme a été observé dans presque tous les processus métaboliques des cellules eucaryotes. Une partie du travail est effectuée par une protéine kinase. Elle enlève un groupe phosphate de l'ATP, qui devient ADP, et le place sur une protéine. Les propriétés de la protéine sont alors altérées. Par la suite, la protéine phosphatase va déplacer ce groupe, ce qui consiste en la déphosphorylation. Les protéines tyrosines phosphatases, comme PTP-BL, déphosphorylent spécifiquement les tyrosines. PTP-BL peut déphosphoryler RIL (Cuppen et al., 1997),  $\beta$ -caténine, Src et ephrin B (Palmer et al., 2002).

## 1.3 Description des domaines PDZ étudiés

Les domaines PDZ, aussi appelés DHR ou GLGF, sont des domaines globulaires contenant habituellement de 80 à 90 résidus. Un domaine globulaire est formé de structures secondaires reliées par des boucles changeant abruptement de direction. Les domaines PDZ se composent habituellement de 5 ou 6 feuillets  $\beta$  et de deux hélices  $\alpha$ . Ils tiennent leur nom de trois protéines, c'est-à-dire PSD-95 (Cho et al., 1992), Dlg (Woods et al., 1991) et ZO-1 (Willott et al., 1993). Les domaines PDZ sont souvent impliqués dans des interactions protéines-protéines (Petsko et al., 2004). Par exemple, ils interagissent avec les canaux ioniques des membranes en se liant à leur groupement C-terminal (Doyle et al., 1996).

Des peptides sont souvent utilisés afin d'étudier les interactions protéines-protéines. En effet, les peptides sont dérivés de la protéine se liant à la protéine cible. Certains motifs de liaison se répètent pour les domaines PDZ. Le peptide se lie habituellement dans une fissure hydrophobe composée d'un feuillet  $\beta$ , d'une hélice  $\alpha$  et d'une boucle qui lie le peptide à son groupement carboxyle. Cette liaison s'effectue dans un mode antiparallèle

au feuillet  $\beta$  avec le résidu C-terminal qui occupe la pochette hydrophobe. Un exemple de ce motif de liaison est présenté à la Figure 1.3. Le domaine PDZ2 est représenté en surface de van der Waals (1.) et en rubans (2.). La flèche rouge indique la direction, c'est-à-dire à partir du C terminal jusqu'au N terminal, avec laquelle le ligand dérivé d'APC se lie par rapport à PDZ2. Les résidus en orange indiquent le site de liaison et les résidus en vert et en rouge représentent respectivement le début du domaine et la fin du domaine.



**Figure 1.3** Motif de liaison de PDZ2.<sup>1</sup> 1. Représentation de Van der Waals de PDZ2.  
2. Représentation en rubans de PDZ2.

Les domaines PDZ se lient à 4 ou 5 acides aminés de leur protéine cible. Souvent il s'agit de récepteurs transmembranaires ou de canaux ioniques. Le résidu du C-terminal est habituellement Val ou Ile. Les résidus à la 2<sup>ème</sup> et 3<sup>ème</sup> positions sont importants à la spécificité de la liaison.

<sup>1</sup> Voir banque de donnée PDB, identité 1GM1

### 1.3.1 Le domaine tandem PDZ2/3

Les domaines PDZ2 et PDZ3 ont été étudiés par RMN. Dans la séquence entière de PTP-BL, le domaine PDZ2 se trouve entre les acides aminés 1357 et 1442 tandis que PDZ3 se situe de 1491 à 1579. Le domaine tandem PDZ2/3 contient en tout 252 acides aminés. PDZ2 et PDZ3 sont séparés par un lieur de 45 acides aminés. Le poids moléculaire total de ces deux domaines est de 26,4 kDa. À la Figure 1.4, la séquence d'acides aminés du domaine tandem PDZ2/3 est présentée. Les résidus marqués en noir représentent une partie du lieur entre PDZ1 et PDZ2 ainsi qu'entre PDZ3 et PDZ4. Les acides aminés indiqués en orange constituent PDZ2, ceux en turquoise PDZ3 et ceux en magenta le lieur entre PDZ2 et PDZ3.

On constate que PDZ2 et PDZ3 sont très similaires en les comparant par ClustalW (Higgins et al., 1994). ClustalW<sup>2</sup> est un outil de comparaison de séquences. Le résultat obtenu est montré à la Figure 1.5. Les résidus indiqués en rouge sont hydrophobes et petits, ceux en bleus sont acides, ceux en magentas sont basiques et ceux en vert sont amphotères. Les acides aminés possédant une étoile sont identiques dans les deux séquences, ceux qui ont deux points ont des caractéristiques chimiques semblables et ceux qui possèdent un point sont totalement différents. Le pourcentage d'identité est de 33%, ce qui est élevé. Les domaines sont donc homologues.

---

<sup>2</sup> <http://www.ebi.ac.uk/clustalw/>

GSPGSASPPKPGDTFEVELAKTDGSLGISVTGGVNTSVRHGGIYVKAIIP  
 50  
 KGAAESDGR IHKGDRVLAVNGVSLEGATHKQAVETLRNTGQVVHLLLEKQ  
 100  
 QVPTSRRERDPAGPQSPPPDQDAQRQAPEKQVAKQTPHVKDYSFVTEENTFE  
 150  
 VKLFFKNSSGLGFSFSREDNLIPEQINGSIVRVKCLFPGQPAAESGKIDVG  
 200  
 DVILKVNAGAPLKGLSQQDVISALRGTAPVSLLLCRPAPGVLPEIDTPGN  
 250  
 SS

PDZ2 : 89 acides aminés  
 PDZ3 : 92 acides aminés  
 Lieur: 45 acides aminés  
 Poids moléculaire : 26,4 Kda

**Figure 1.4** Séquence d'acides aminés du domaine tandem PDZ2/3.

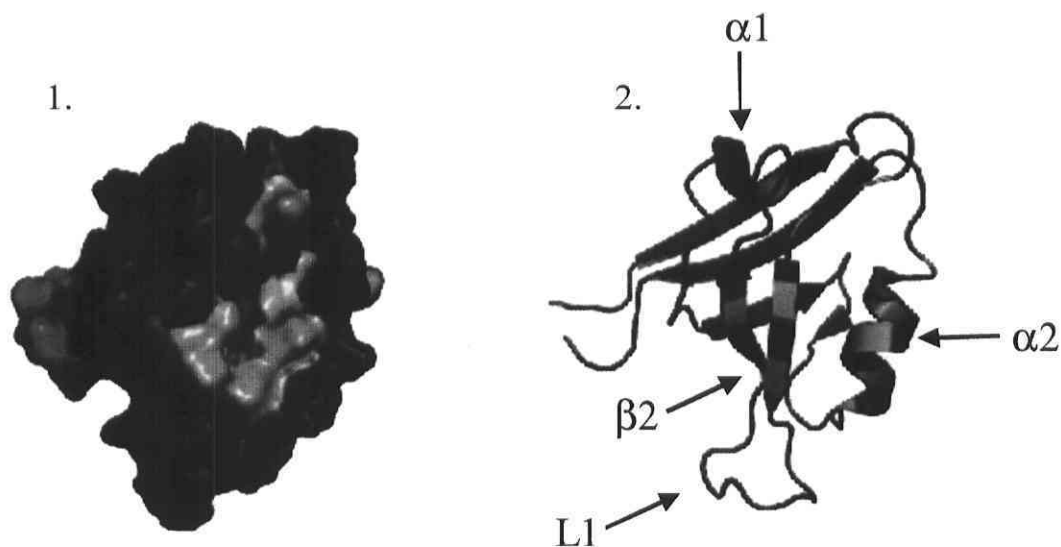
DTFEVELAKTDGSLGISVTGG---VNTSVRHGGIYVKAIIPKGAAESDGR  
 NTFEVKLFKNSSGLGFSFSREDNLIPEQINGSIVRVKCLFPGQPAAESGK  
 :\*\*\*\*:\* \*...\*\*:\*.: : .:. . : \*\* ::\* .\* ..\*:  
 IHKGDRVLAVNGVSLEGATHKQAVETLRNTGQVVHLLLEKQ  
 IDVGDVILKVNAGAPLKGLSQQDVISALRGTAPVSLLLCRPA  
 \*. \*\* :\* \*\*\*..\*:\* :::::.\*.\* \* \*\*\* :

**Figure 1.5** Alignement des domaines PDZ2 et PDZ3.

### 1.3.2 PDZ2

PDZ2 contient 89 acides aminés. Sa structure et son attribution ont été déterminées par RMN (Walma et al., 2002). PDZ2 contient 2 hélices  $\alpha$  ainsi que six feuillet  $\beta$ . Cette structure est présentée à la Figure 1.6. À droite, on retrouve la structure de PDZ2 en rubans, et à gauche la structure en surface de Van der Waals. Les changements induits par APC sont indiqués en orange sur la surface de van der Waals.

La pochette de liaison de PDZ2 se situe entre l'hélice  $\alpha 2$  et le feuillet  $\beta 2$ . Cette pochette est séparée par une boucle (L1) flexible qui se replie dans le squelette de la protéine. Cette boucle est un trait particulier différenciant PDZ2 des autres domaines PDZ. Elle se présente aussi dans les domaines PDZ2 des orthologues humains de PTP-BL. La boucle L1 semble être importante à la liaison car son extension de 5 acides aminés réduit à néant la liaison avec le ligand dérivé d'APC (Walma et al., 2004).



**Figure 1.6** Structure de PDZ2 déterminée par RMN<sup>3</sup>. 1. Représentation de Van der Waals de PDZ2. 2. Représentation en rubans de PDZ2.

<sup>3</sup> Voir banque de donnée PDB, identité 1GM1



### 1.3.2.1 APC

Tel qu'indiqué à la Figure 1.2, PDZ2 se lie avec APC. Le gène humain d'APC, un gène suppresseur de tumeurs, code pour une protéine de 2843 acides aminés. La structure d'APC consiste en un domaine N-terminal suivi de 7 ARM. APC a été reliée au cancer du côlon familial et sporadique. En effet, une mutation de son gène peut causer le désordre FAP (Familial adenomatous polyposis). FAP est un désordre héréditaire caractérisé par la présence de centaine de milliers de polypes adénomateux dans tout le côlon. Les patients atteints de ce syndrome et n'étant pas traités développent le cancer du côlon. Les protéines APC normales favorisent l'apoptose (suicide cellulaire) dans les cellules du côlon. Sa fonction la plus importante est de séquestrer les effets stimulateurs de croissance de la  $\beta$ -caténine, une protéine qui active les gènes associés à la croissance cellulaire. Ainsi la perte de fonction d'APC va empêcher l'apoptose. Cela va donc permettre à la  $\beta$ -caténine de stimuler la croissance cellulaire. Une croissance cellulaire rapide augmente la possibilité de mutations. En effet, une bonne quantité de ces événements se produisent chez les patients souffrant du FAP. Ces personnes développent souvent des tumeurs malignes. Un peptide de 12 acides aminés dérivant de APC a été utilisé afin d'effectuer une titration sur le domaine tandem PDZ2/3. Ce peptide est présenté au Tableau 1.1.

Domaine	Ligand utilisé	Dérivé de
PDZ2	H-KRHSGSYLVTSV-OH	APC
PDZ3	H-MFRDFDYIADWC-OH	PRK2

**Tableau 1.1** Ligands utilisés lors des titrations

### **1.3.3 PDZ3**

PDZ3 contient 92 acides aminés. Sa structure n'a pas encore été déterminée et l'attribution des signaux RMN n'a pas été faite dans la littérature. En fait, ce domaine n'a pas encore beaucoup été étudié. Dans la Figure 1.2, il est possible de remarquer que PDZ3 ne se lie qu'avec PRK2. Le ligand dérivé de PRK2 sera donc utilisé pour effectuer la titration.

#### **1.3.3.1 PRK2**

PRK2 est une kinase sérine/thréonine (elle phosphoryle les thréonines et les sérines des protéines cibles) qui est régulée par la protéine-G Rho. PRK2 (Vincent et al., 1997; Flynn et al., 2000) est un effecteur rho. Un effecteur est une molécule qui altère l'activité de la protéine. Il peut avoir un effet activateur ou inhibiteur. L'interaction de PRK2 avec PDZ3 dépend de trois acides aminés et une cystéine C-terminale est indispensable pour cette interaction. D'autres domaines PDZ montrent une préférence pour les acides aminés volumineux et hydrophobes comme valine, leucine ou isoleucine dans la position C-terminale (Songyang et al., 1997). Cette spécificité était expliquée par l'existence d'une pochette hydrophobe dans le domaine PDZ qui peut être remplie seulement par des acides aminés volumineux et hydrophobes (Doyle et al., 1996; Morais et al., 1996; Tochio et al., 1999). Dans le Tableau 1.1, le peptide dérivé de PRK2 peut être observé.

## 1.4 Objectifs des recherches effectuées

Long et Tochio ont démontré que le domaine tandem PDZ1/2 de PSD-95 était davantage efficace pour lier certains ligands que chacun des deux domaines séparés (Long et al., 2003). Ils ont en fait décrit que le lieu, beaucoup plus court que celui entre le domaine tandem PDZ2/3, rigidifie la structure du domaine tandem PDZ1/2. Ces deux domaines auraient donc un degré de rotation limité l'un par rapport à l'autre. Cela aurait pour effet de faciliter l'interaction avec des protéines membranaires multimériques. Ce phénomène était pour la première fois observé pour des domaines PDZs.

Des arrangements de domaines tandem similaires ont déjà été observés pour d'autres types de domaines que PDZ. Par exemple, la structure du domaine tandem SH2 de ZAP-70 a permis d'illustrer que les domaines SH2, arrangés par paires, permettent à la protéine d'interagir avec un peptide phosphotyrosique avec beaucoup de spécificité (Hatada et al., 1995).

L'hypothèse que les deux domaines seraient accolés par d'autres interactions a été esquissée dans l'article de Long et Tochio. En effet, le lieu ne serait pas le seul facteur à l'attachement des deux domaines. Des expériences effectuées sur le domaine tandem PDZ2/3 permettrait de vérifier cette hypothèse. Il est possible de remarquer à la Figure 1.4 que le lieu entre le domaine tandem PDZ2/3 contient 45 acides aminés. Il est donc beaucoup plus long que celui entre le domaine tandem PDZ1/2 de PSD-95 qui contient 4 acides aminés. Long et Tochio ont aussi constaté que lors de l'insertion dans le lieu de 3 acides aminés (SGG), la liaison avec le ligand était moins efficace. Avec deux domaines séparés d'un très long lieu, il est donc possible de vérifier si les domaines PDZ interfèrent entre eux.

Plusieurs hypothèses peuvent être posées à partir des faits établis par Long et Tochio. En premier lieu, les domaines PDZ2 et PDZ3 pourraient exister indépendamment en solution sans aucune interaction entre eux avant et après l'addition des peptides. Ensuite, ils

pourraient être associés dès le départ sans aucune interaction spécifique de la part du lieur et rester ainsi lors de l'ajout du peptide. Une autre possibilité serait qu'ils soient séparés en solution mais qu'à l'ajout du ligand ils forment une superstructure. Finalement, les deux domaines pourraient interagir ensemble en partie. Dans ce dernier cas, le lieur pourrait aussi être impliqué dans le rapprochement des deux domaines.

Afin de vérifier ces différentes hypothèses, le but des recherches effectuées était de déterminer si le domaine tandem PDZ2/3 forme une superstructure en milieu aqueux. Pour ce faire, il est nécessaire de comparer la structure des domaines libres en solution et liés. L'attribution spectrale du domaine tandem PDZ2/3 a été effectuée à 88%. Cependant, comme les spectres obtenus afin de déterminer les contraintes structurales n'ont pas été de qualité suffisante, la structure du domaine PDZ3 n'a pas pu être déterminée par spectroscopie RMN. D'autres techniques ont par contre été utilisées. Par exemple, des perturbations dans les déplacements chimiques ont été observés lors des titrations. Cela permet de tirer des conclusions sur le positionnement d'un domaine par rapport à l'autre. De plus, la simple comparaison des spectres du domaine PDZ2 avec le domaine tandem PDZ2/3 montre clairement que le domaine PDZ2 est fortement perturbé par la présence du lieur ainsi que du domaine PDZ3.

## Chapitre 2

### Matériel et Méthodes

#### 2.1 Matériel

La première étape à la détermination de la structure d'une protéine est la préparation des échantillons. Pour les spectres à une dimension, les échantillons ne doivent pas nécessairement être marqués. Le domaine tandem PDZ2/3 recombinant de PTP-BL de la souris a donc été préparé dans le laboratoire de Raphael Stoll par Christian Fetzer dans le département de biochimie à la Ruhr-Universität Bochum en Allemagne. Il a été obtenu à partir du système d'expression *E. coli* BL21(DE3) et contient les résidus 1354 à 1588 de PTP-BL. Le domaine tandem PDZ2/3 a été cloné dans un vecteur pGEXT2 (Amersham, Uppsala). Les échantillons de protéines isotopiquement marquées à  $^{13}\text{C}/^{15}\text{N}$  et  $^{15}\text{N}$  ont été préparés en faisant croître les bactéries dans un milieu minimal contenant  $^{15}\text{NH}_4\text{Cl}$  avec ou sans  $^{13}\text{C}_6$ -glucose. Le système de fusion de gène GST (Gluthathione S-transferase) (2002) a été utilisé afin d'effectuer l'expression et la purification du domaine tandem PDZ2/3.

Tous les spectres RMN ont été enregistrés à 25°C sur un spectromètre Bruker DRX600 (Rheinstetten, Allemagne). Les échantillons RMN contenaient environ 1 mM de protéine dans 50 mM  $\text{KH}_2\text{PO}_4$ , 50 mM  $\text{Na}_2\text{HPO}_4$ , 150 mM NaCl, pH 7.4, 5 mM DTT, 0.02 %  $\text{NaN}_3$ , et le mélange d'inhibiteurs de protéase Mini, sans EDTA, de Roche (Mannheim, Allemagne).

Le squelette de PDZ2/3 a été assigné à l'aide des séquences RMN CT-HNCO, CT-HNCA, HN(CO)CA, CBCA(CO)NH en utilisant la séquence WATERGATE, pour la suppression de l'eau. Ces séquences sont décrites plus en détail dans la section suivante.

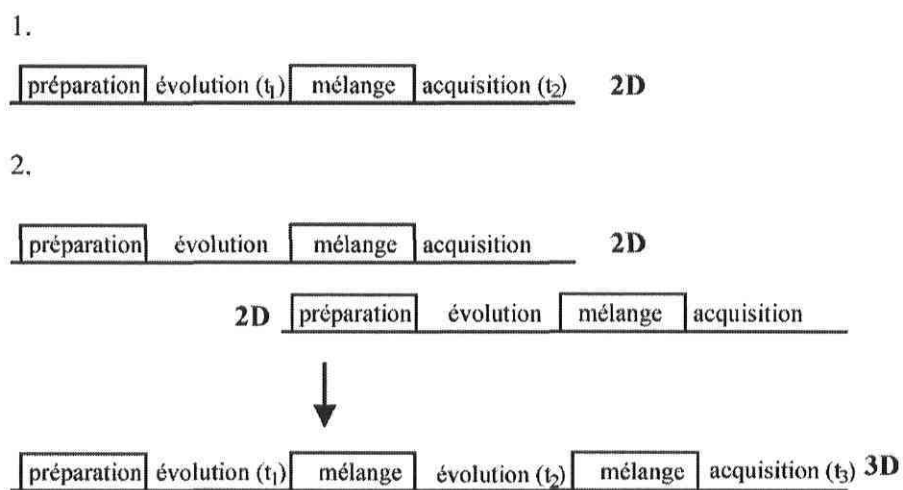
## 2.2 Spectres 2D et 3D

Dans la plus simple des expériences RMN pulsée, les spins nucléaires sont excités par une impulsion de radiofréquence (rf). La magnétisation transverse est enregistrée à des intervalles réguliers durant la période d'acquisition afin de générer une représentation digitale du signal de précession libre (FID). Par la suite, une transformée de Fourier est effectuée afin d'obtenir le spectre 1D conventionnel. Dans des expériences RMN plus complexes, les perturbations (prenant habituellement la forme de champs pulsés) sont imposées au système de spins durant la période précédant l'acquisition du spectre (voir Figure 2.1). De plus, comme les spectres 1D de protéines sont très chevauchés, des spectres à plusieurs dimensions doivent donc être mesurés.

Les spectres à 2 dimensions possèdent deux étapes de plus que les spectres à une dimension. Ces deux étapes additionnelles sont la période d'évolution et la période de mélange. Quatre étapes sont donc nécessaires à la mesure d'un spectre 2D. Ces étapes sont : les périodes de préparation, d'évolution, de mélange et d'acquisition. Durant la période de préparation l'état de non équilibre du système de spin est préparé à partir de l'état d'équilibre tandis que dans la période d'évolution ( $t_1$ ), un temps de délai variable augmente d'une valeur initiale à une valeur finale. Pour chaque valeur de  $t_1$ , un signal FID de  $n$  points digitalisés est enregistré. Pour ce qui est de la période de mélange, la cohérence est transférée d'un spin à

un autre, établissant ainsi l'information contenue dans le spectre. Finalement, le FID est enregistré durant la période d'acquisition ( $t_2$ ).

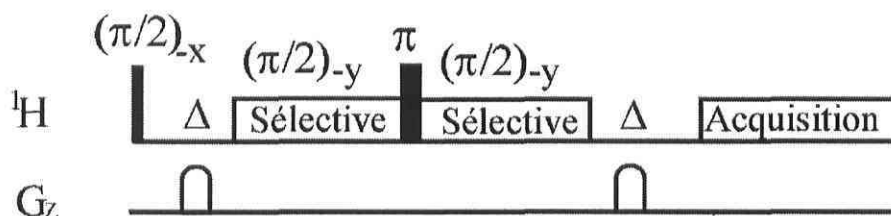
Les séquences d'impulsions des spectres 3D sont dérivés de celle des spectres 2D. En effet, le signal est enregistré durant le temps d'acquisition, notée  $t_3$ , en fonction des deux temps d'évolution  $t_1$  et  $t_2$ . Un schéma sommaire de la structure des séquences d'impulsions des spectres 2D et 3D est présenté à la Figure 2.1.



**Figure 2.1** Schéma sommaire des séquences d'impulsions 2D et 3D. 1. Séquence 2D, 2. Séquences 2D superposées résultant en un spectre 3D.

### 2.3 WATERGATE

Les mesures effectuées sur PTP-BL ont été faites dans un milieu aqueux comme la majorité des études portant sur les macromolécules biologiques. Cependant, comme la concentration en protons de l'eau est beaucoup plus élevée que celle des macromolécules, on observe majoritairement une bande intense représentant le solvant. WATERGATE (WATER suppression by Gradient Tailored Excitation) est donc utilisé comme technique de suppression d'eau. Un exemple de WATERGATE (Piotto et al., 1992; Grzesiek et al., 1993; Kay et al., 1994) est donné à la Figure 2.2. L'impulsion initiale non sélective est suivie par un fort gradient de champ pulsé. Celui-ci déphase la magnétisation du solvant et de la macromolécule. Un gradient de champ pulsé représente une période durant laquelle le champ  $B_0$  est délibérément inhomogène. Dans un champ inhomogène, la magnétisation transverse et les autres cohérences sont déphasés à travers l'échantillon. La cohérence peut être refocussée par un autre gradient pour générer des échos de gradient. La magnétisation de la macromolécule n'est pas affectée par les impulsions sélectives (les résonances d'intérêt doivent être en dehors de la largeur de ces impulsions). Une impulsion non sélective de  $180^\circ$  inverse l'ordre de la cohérence de la macromolécule. Le second gradient de champ pulsé rephase la magnétisation de la macromolécule pour former un écho de gradient. La combinaison d'impulsion sélective de  $90^\circ$  et de celle non sélective de  $180^\circ$  laisse l'ordre de cohérence du solvant inchangée, donc le second gradient de champ pulsé continue à déphaser la magnétisation du solvant, et aucun écho de gradient n'est formé.

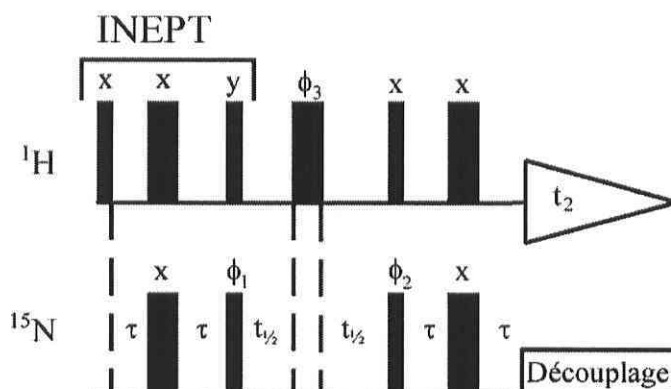


**Figure 2.2** Exemple de séquence d'impulsions WATERGATE



## 2.4 HSQC

Le HSQC (Heteronuclear Single Quantum Coherence) (Bodenhausen et al., 1980; Piotto et al., 1992; Sklenar et al., 1993) est le spectre de base de l'attribution d'une protéine. Il forme la base 2D de plusieurs spectres 3D. Un signal par acide aminé est observé dans un HSQC. Il est habituellement mesuré avec un échantillon marqué seulement à  $^{15}\text{N}$ . La corrélation NH de l'amine de chaque acide aminé est visible. Les prolines ne sont donc pas visibles dans ce type de spectre. En effet, la proline ne possède pas de  $\text{H}_\text{N}$  suite au lien peptidique. Pour ce qui est du premier résidu, il ne sera pas observé car son amine est protonnée. Une liste de signaux est effectuée et elle constitue la base utilisée pour les spectres 3D. Cette liste est utilisée par la suite pendant toute la durée de l'attribution et elle contient les informations relatives à la position du signal en ppm ainsi que son intensité et sa phase. La séquence HSQC est illustrée à la Figure 2.3.



**Figure 2.3** Séquence d'impulsions HSQC.

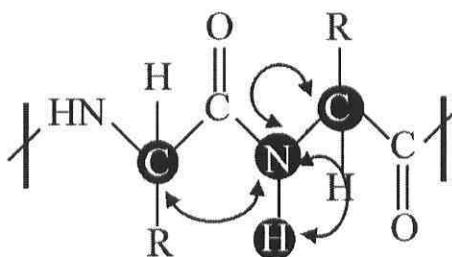
La séquence d'impulsion consiste en deux transferts de type INEPT (Morris et al., 1979; Morris, 1980; Morris, 1980). En effet, la polarisation du proton est transférée sur le  $^{15}\text{N}$ . Une seconde séquence INEPT est utilisée pour transférer la polarisation au proton pour la détection.

### 2.4.1 INEPT

La séquence INEPT (Insensitive Nuclei Enhanced by Polarization Transfer) (Maudsley et al., 1977; Morris et al., 1979) est une composante importante de plusieurs expériences RMN multidimensionnelles. Son but est de transférer la magnétisation d'un noyau sensible avec un rapport gyromagnétique élevé (habituellement les protons) à un noyau moins sensible avec un rapport gyromagnétique plus faible (l'azote ou le carbone). Le signal du  $^{13}\text{C}$  ou de  $^{15}\text{N}$  va être ainsi plus élevé. La séquence INEPT est illustrée à la Figure 2.3.

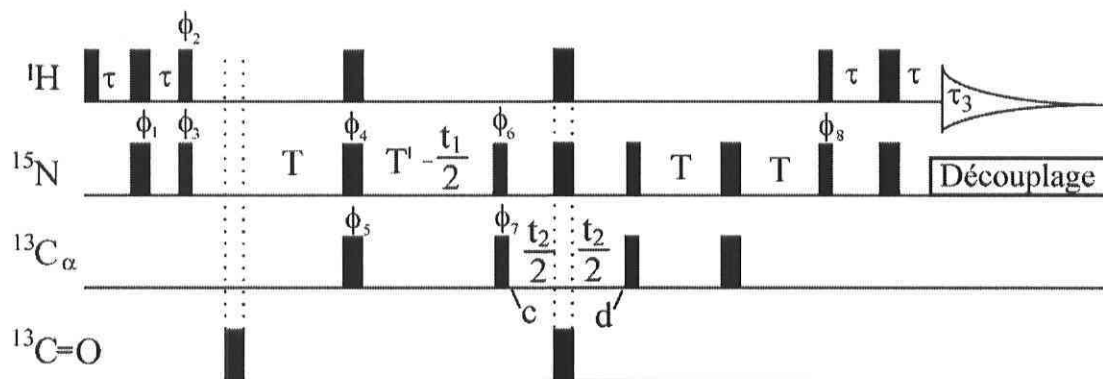
## 2.5 HNCA

Le HNCA (Ikura et al., 1990; Grzesiek et al., 1992) est une séquence très utilisée pour l'attribution des signaux de protéines. En effet, les signaux obtenus correspondent aux corrélations suivantes :  $^1\text{H}_i^{\text{N}}-^{15}\text{N}_i-^{13}\text{C}_i^{\alpha}$  et  $^1\text{H}_i^{\text{N}}-^{15}\text{N}_i-^{13}\text{C}_{i-1}^{\alpha}$ . Ces corrélations sont présentées à la Figure 2.4.



**Figure 2.4** Corrélations observées dans un HNCA.

On obtient donc deux signaux par acide aminé. Habituellement le signal étant le plus intense provient du  $\text{C}_\alpha$  et le moins intense du  $\text{C}_{\alpha-1}$ . La séquence est présentée à la Figure 2.5.



**Figure 2.5** Séquence d'impulsions HNCA.

Le HNCA utilise l'INEPT pour transférer la polarisation du  $^1\text{H}$  à celle du  $^{15}\text{N}$ . La première dimension spectrale mesurée consiste au signal représentant la magnétisation transférée du

proton amide à l'atome d'azote directement attaché à celui-ci. Dans la deuxième dimension spectrale, la magnétisation est transférée au noyau  $C_{\alpha}$ . Finalement la magnétisation est transférée au proton de l'amide ce qui constitue la troisième dimension. L'expérience a été mesurée à CT (constant time). Ceci indique que l'expérience inclut une période durant laquelle l'évolution du déplacement chimique de  $^{15}\text{N}$  et l'évolution de l'interaction du couplage scalaire  $^{15}\text{N}$ - $^{13}\text{C}_{\alpha}$  se produit simultanément. Le principal avantage de l'évolution à temps constant est une réduction des pertes de signaux dus à la relaxation et une concomitante amélioration de la sensibilité. CT peut parfois ne pas être approprié pour les grosses protéines.

En théorie, il est donc possible d'attribuer les signaux du HSQC seulement avec un spectre HNCA par l'attribution séquentielle. Chaque acide aminé est représenté habituellement par deux signaux : un signal représentant le  $C_{\alpha}$  du résidu précédent et un signal représentant le  $C_{\alpha}$  du résidu cible. La même fréquence de  $C_{\alpha}$  est donc retrouvée deux fois dans le spectre à des valeurs HN- $^{15}\text{N}$  différentes. En reliant ces signaux, il est possible de « marcher » le long de la chaîne d'acides aminés. Cependant, davantage de spectres sont requis car plusieurs atomes possèdent des  $C_{\alpha}$  ayant un déplacement chimique similaire. Au Tableau 2.1, une table des déplacements chimiques observées pour les  $C_{\alpha}$  ainsi que pour les  $C_{\beta}$  est présentée. Cette table est un résumé de celle présente sur le site Internet du BMRB (Biological Magnetic Resonance Data Bank)<sup>4</sup>.

---

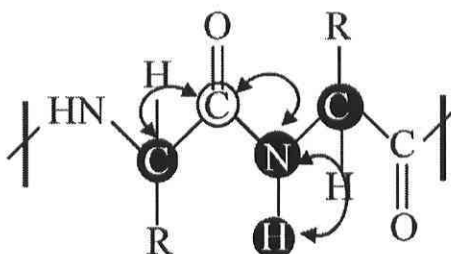
<sup>4</sup> <http://www.bmrwisc.edu>

Acides aminés	Déplacements chimiques moyens (ppm)	
	$C_{\alpha}$	$C_{\beta}$
Ala	(51.15 à 55.21) $\pm$ 2.01	(16.92 à 20.98) $\pm$ 1.85
Arg	(54.45 à 59.23) $\pm$ 2.37	(28.78 à 32.44) $\pm$ 1.83
Asp	(52.53 à 56.71) $\pm$ 2.09	(39.16 à 42.48) $\pm$ 1.68
Asn	(51.53 à 55.47) $\pm$ 1.95	(36.86 à 40.6) $\pm$ 1.75
Cys	(54.12 à 61.00) $\pm$ 3.40	(27.62 à 41.28) $\pm$ 6.53
Glu	(55.3 à 59.58) $\pm$ 2.13	(28.22 à 31.8) $\pm$ 1.75
Gln	(54.35 à 58.73) $\pm$ 2.15	(27.26 à 31.1) $\pm$ 1.89
Lys	(54.64 à 59.16) $\pm$ 2.23	(30.9 à 34.62) $\pm$ 1.83
Phe	(55.53 à 60.89) $\pm$ 2.66	(37.84 à 42.02) $\pm$ 2.09
Pro	(61.64 à 64.92) $\pm$ 1.62	(30.53 à 33.01) $\pm$ 1.27
Gly	(44.00 à 46.70) $\pm$ 1.33	-
His	(53.96 à 58.96) $\pm$ 2.45	(28.06 à 32.38) $\pm$ 2.15
Ile	(58.79 à 64.35) $\pm$ 2.75	(36.52 à 40.66) $\pm$ 2.07
Leu	(53.44 à 57.84) $\pm$ 2.17	(40.35 à 44.17) $\pm$ 1.88
Ser	(56.45 à 60.83) $\pm$ 2.18	(62.23 à 65.41) $\pm$ 1.57
Thr	(59.41 à 64.89) $\pm$ 2.70	(67.57 à 71.61) $\pm$ 1.82
Tyr	(55.44 à 60.68) $\pm$ 2.61	(37.11 à 41.45) $\pm$ 2.18
Val	(59.51 à 65.43) $\pm$ 2.93	(30.84 à 34.46) $\pm$ 1.82

**Tableau 2.1** Intervalle des déplacements chimiques moyens des  $C_{\alpha}$  et  $C_{\beta}$

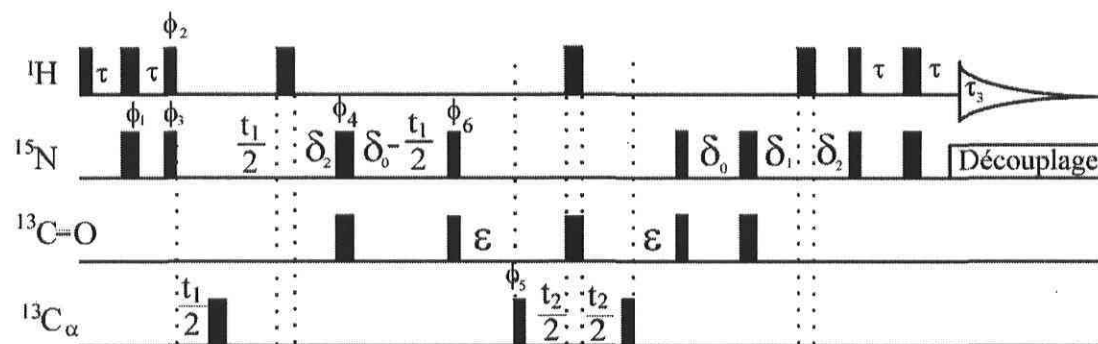
## 2.6 HN(CO)CA

Le HN(CO)CA (Grzesiek et al., 1992) permet de confirmer quel signal dans le HNCA est le  $C_{\alpha-1}$  dans ce spectre. On obtient donc une seule corrélation;  ${}^1\text{H}_i\text{N}-{}^{15}\text{N}_i-{}^{13}\text{C}_{i-1}\alpha$ . Cette corrélation est représentée à la Figure 2.6. Ce spectre permet parfois d'observer des signaux étant invisibles sur le HNCA ou de différencier de façon sûre le  $C_\alpha$  du  $C_{\alpha-1}$ .



**Figure 2.6** Corrélation observée dans un HN(CO)CA.

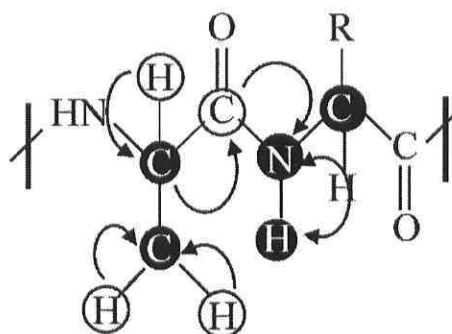
La séquence est présentée à la Figure 2.7. La magnétisation est transférée du proton de l'amide via le  ${}^{15}\text{N}$  au  $\text{C}^\alpha$  via le spin du carbonyle.



**Figure 2.7** Séquence d'impulsions HN(CO)CA.

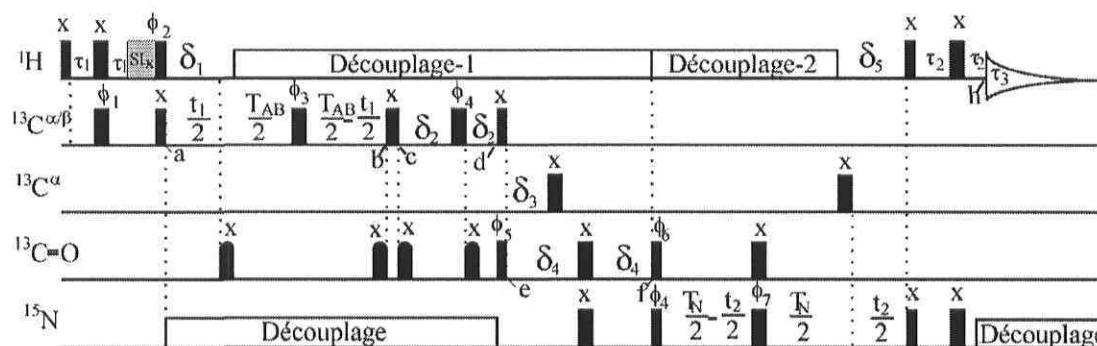
## 2.7 CBCA(CO)NH

Le CBCA(CO)NH (Grzesiek et al., 1993; Muhandiram et al., 1994) est aussi beaucoup utilisé pour aider à l'attribution des spectres de protéines. Dans ce spectre, il est possible de mesurer un signal pour le  $C_\beta$  du résidu précédent. La corrélation observée est celle-ci :  $^{13}C_i^\beta - ^{13}C_i^\alpha - ^{15}N_{i+1} - ^1H_{i+1}^N$ . Cette corrélation est présentée à la Figure 2.8. Comme plusieurs acides aminés possèdent des  $C_\alpha$  similaires, ce spectre permet d'identifier le  $C_\beta$  de l'acide aminé précédent.



**Figure 2.8** Corrélation observée pour le CBCA(CO)NH.

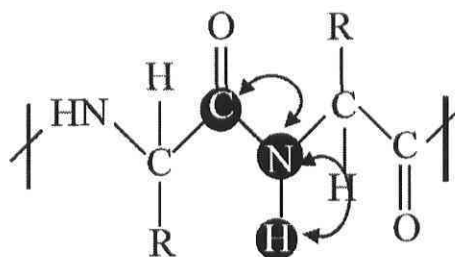
La séquence est illustrée à la Figure 2.9. Comme pour le HN(CO)CA, la corrélation entre les résidus est établie en transférant la cohérence via le spin du  $^{13}CO$ .



**Figure 2.9** Séquence d'impulsions CBCA(CO)NH.

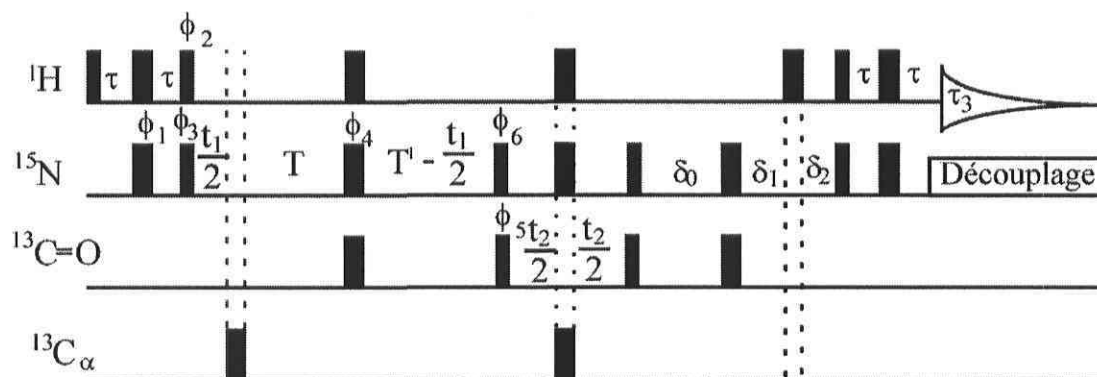
## 2.8 HNCO

Le HNCO (Grzesiek et al., 1992; Schleucher et al., 1993; Kay et al., 1994) permet d'obtenir la fréquence du carbonyle de l'acide aminé précédent. La corrélation, représentée à la Figure 2.10, est celle-ci :  ${}^1\text{H}_i^\alpha - {}^{15}\text{N}_i - {}^{13}\text{CO}_{i-1}$ .



**Figure 2.10** Corrélation observée dans le HNCO.

La séquence d'impulsions est présentée à la Figure 2.11. En effet, le proton de l'amide est corrélé avec l'azote ainsi qu'avec le carbonyle du résidu précédent en utilisant le couplage de la liaison simple  ${}^{15}\text{N}-{}^{13}\text{CO}$  (15 Hz). Ce spectre est surtout utilisé lors du dépôt de l'attribution du squelette sur le site du BMRB. Les fréquences du carbonyle ne sont pas vraiment caractéristiques d'un acide aminé à un autre. Il est donc difficile d'attribuer des acides aminés avec l'aide de ces renseignements.

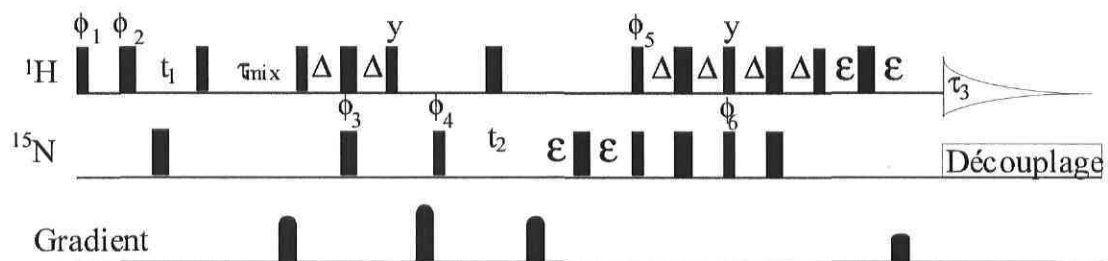


**Figure 2.11** Séquence d'impulsions HNCO.



## 2.9 NOESY-HSQC

Le NOESY-HSQC est habituellement utilisé afin de déterminer les distances interprotons. Il est ainsi possible de savoir si par exemple, un proton d'un acide aminé est très proche d'un autre proton d'un acide aminé éloigné. Ainsi, il sera possible de déterminer selon la distance, si cet endroit est structuré (feuillet  $\beta$  ou hélice  $\alpha$ ). Un exemple de séquence d'impulsions NOESY-HSQC est donné à la Figure 2.12.



**Figure 2.12** Exemple de la séquence d'impulsions du NOESY-HSQC.

## **2.10 Programmes utilisés**

Les expériences RMN de protéines génèrent beaucoup de données qui doivent être analysées. À cet effet, on utilise NMRPipe afin de traiter les données obtenues par RMN (Delaglio et al., 1995). Nmrview est utilisé, par la suite, pour visualiser et analyser les données tandis que MARS (Jung et al., 2004) et Monte (Hitchens et al., 2003) ont été utilisés afin d'effectuer une attribution semi-automatique. Pymol (DeLano, 1998) a été utilisé afin de générer les figures de structures présentées dans la discussion.

### **2.10.1 Attribution semi-automatique**

Afin d'attribuer les signaux obtenus dans les spectres mentionnés plus hauts de façon semi-automatique, plusieurs programmes ont été mis au point. Des essais ont été faits sur deux de ces programmes, c'est-à-dire Monte et MARS.

#### **2.10.1.1 Monte**

Monte (Hitchens et al., 2003) est un programme qui utilise une approche Monte Carlo basée sur le recuit simulé. C'est une approche qui permet d'obtenir et d'identifier une solution unique pour le problème d'attribution de déplacements chimiques. La méthode débute par une attribution aléatoire des systèmes de spins des résidus dans la séquence primaire. Dans ce programme, un système de spins est associé à toutes les informations reliées au signal amide. Cela inclut les connections entre les pics ainsi que les signaux NOE et les déplacements chimiques dans les différents résidus. Après les choix initiaux, deux systèmes de spins sont choisis aléatoirement et par la suite permutés. La nouvelle solution est corrigée et basée sur les mêmes critères. Si le résultat de la permutation améliore le score, la permutation est acceptée. Si le résultat du score est plus bas, la différence est comparée à la « température » du système et la permutation peut ou ne pas être acceptée. À une température infinie (élevée) toutes les permutations sont acceptées. Quand la température du système est basse (recuit simulé), les permutations qui résultent en une amélioration du score général seront préférentiellement acceptées. Monte est

habituellement exécuté de 10 à 50 cycles débutant de positions de départ aléatoires. La meilleure solution de chaque cycle est comparée afin d'évaluer la fiabilité d'une attribution.

### 2.10.1.2 MARS

MARS (Jung et al., 2004) optimise la qualité globale et locale d'une attribution pour minimiser la propagation d'erreurs initiales et pour extraire les attributions fiables. Il ne requiert pas de seuil étroit pour établir des connectivités séquentielles ou un ajustement détaillé de ces seuils. Il peut fonctionner avec une grande variété d'expériences RMN et est robuste contre le manque d'information sur les déplacements chimiques. Ce programme peut être utilisé pour des protéines de plus de 15 kDa en utilisant seulement des informations sur les déplacements chimiques des  $C_\alpha$  et  $C_\beta$ . En premier lieu, l'optimisation simultanée de la qualité globale et locale de l'attribution est effectuée. Ensuite, une recherche complète de la longueur des fragments compris jusqu'à PRs (pseudo-résidu) pendant la liaison et l'application est effectuée. L'application constitue en l'action de comparer deux ensembles de données afin d'en voir les correspondances. Les meilleurs éléments de l'application et de la liaison sont conservés. Par la suite, la combinaison de la structure secondaire du programme PSIPRED (McGuffin et al., 2000) avec les distributions de déplacements chimiques effectue une amélioration dans l'identification des meilleures positions dans la séquence primaire. PSIPRED est un serveur qui permet aux usagers de soumettre une séquence afin d'en prédire la structure secondaire. L'évaluation de la fiabilité de l'attribution par l'application de fragments en effectuant des attributions multiples est effectuée.

## 2.11 Titration

La titration du domaine tandem PDZ2/3 avec les deux ligands mentionnés a été effectuée afin de vérifier les hypothèses posées dans l'introduction. De plus, le domaine PDZ2 marqué à  $^{15}\text{N}$  a aussi été titré avec le domaine PDZ3 non marqué. La première étape lors d'une titration est de calculer la constante de dissociation ( $K_d$ ) (Harding et al., 2001). La liaison d'un ligand L avec une protéine P peut être considérée comme une réaction de second ordre;



Cet équilibre est décrit par la constante de dissociation;

$$K_d = K_{\text{Dis}} / K_{\text{Ass}} \quad (2)$$

$K_{\text{Dis}}$  est la constante de vitesse de dissociation du ligand du complexe ligand-protéine tandis que  $K_{\text{Ass}}$  est la constante de vitesse d'association du ligand avec la protéine. Les temps de vie  $\tau_L$ ,  $\tau_P$  et  $\tau_{\text{PL}}$  sont donnés respectivement par les équations suivantes :

$$\tau_L = 1 / (K_{\text{Ass}} [P]) \quad (3)$$

$$\tau_P = 1 / (K_{\text{Ass}} [L]) \quad (4)$$

$$\tau_{\text{PL}} = 1 / (K_{\text{Dis}}) \quad (5)$$

[P] et [L] représentent respectivement la concentration de la protéine libre en solution et du ligand (peptide) libre en solution. Un seul temps  $\tau$  est utilisé pour caractériser le processus défini par l'équation 1 :

$$1 / \tau = 1 / \tau_{\text{PL}} + 1 / \tau_L = K_{\text{Dis}} (1 + X_{\text{PL}} / X_L) \quad (6)$$

$X_{\text{PL}}$  et  $X_L$  sont respectivement la fraction du ligand dans les états liés et libres.

$$X_{\text{PL}} = [\text{PL}] / [\text{L}_{\text{total}}] \quad (7)$$

$$X_L = [L] / [\text{L}_{\text{total}}] \quad (8)$$

$$[\text{L}_{\text{total}}] = [\text{PL}] + [L] \quad (9)$$

[PL] représente la concentration du complexe en solution. Une approximation de la vitesse d'échange peut être effectuée à partir des spectres 1D et 2D. Si des résonances séparées sont observées pour le ligand libre et lié alors la vitesse d'échange est lente. Cela indique que la vitesse d'équilibre  $1/\tau$  est plus lente que la séparation du déplacement chimique  $\Delta\delta$  de la résonance du ligand lié et du ligand libre en solution :

$$1 / \tau \ll \Delta\delta \quad (10)$$

Quand  $1/\tau$  est plus rapide que la séparation du déplacement chimique et plus rapide que les différences dans les vitesses de relaxation:

$$1 / \tau \gg \Delta\delta \quad (11)$$

Il y a alors échange rapide et une seule résonance est observée sur le spectre. Il y a échange intermédiaire lorsque la vitesse est presque égale à la séparation du déplacement chimique.

$$1 / \tau > \Delta\delta \quad (12)$$

L'intérêt principal est d'observer si PDZ2 se lie bien avec le ligand A et PDZ3 avec le ligand B. Un changement significatif des valeurs de déplacement chimique permet de savoir où se trouve le site de liaison. Afin de comparer les résultats obtenus pour chaque signal, il est possible de calculer un  $\Delta\delta_{\text{obs}}$  dont la formule est indiquée ci-dessous.

$$\Delta\delta_{\text{obs}} = \sqrt{(\Delta\delta_{\text{HN}})^2 + \left(\frac{\Delta\delta_{\text{15N}}}{5}\right)^2} \quad (13)$$

Cette formule tient en compte le fait que les déplacements observés dans l'échelle du proton sont situés entre 5 et 11 ppm, tandis que pour l'azote, ils sont situés entre 100 et 130 pp. Un facteur de 5 est donc apporté à l'azote afin de normaliser son déplacement chimique.  $\Delta\delta$  représente le déplacement chimique après la titration moins le déplacement chimique avant la titration.

## 2.12 Acides aminés sélectivement marqués

Afin de faciliter l'attribution, des spectres d'échantillons contenant des acides aminés spécifiquement marqués ont été mesurés (Gly, Val, Phe, Leu) (McIntosh et al., 1990). En effet, les spectres RMN contiennent certains artéfacts, comme du chevauchement et des signaux manquants. Les échantillons ont été préparés selon le protocole donné dans l'article de Muchmore et al. (Muchmore et al., 1989). Dans ces expériences, un média minimal a été utilisé. Afin d'obtenir un acide aminés spécifiquement marqué, celui-ci a été ajouté dans le milieu. Parfois il est possible d'observer de la dilution isotopique. En effet, le métabolisme peut parfois inverser l'endroit où l'azote devrait être marquée. Par exemple, en ajoutant spécifiquement des glycines marquées, souvent on observe aussi des sérines marquées.

## **Chapitre 3**

### **Attribution du squelette du domaine tandem PDZ2/3 de PTP-BL**

#### **3.1 Le repliement des protéines**

Comme il a déjà été mentionné dans le texte, la structure des protéines est directement reliée à sa fonction. Les interactions entre deux molécules dépendent aussi de la structure des protéines. Quelques interactions reliées aux domaines PDZs seront analysées au Chapitre 4. Le repliement d'une protéine, sous sa conformation native, s'effectue habituellement spontanément dans les conditions physiologiques. Dans ce cas-ci, la protéine est déjà repliée car elle a été synthétisée dans des conditions similaires aux conditions physiologiques. Cela impliquerait que la structure primaire dicterait la structure tridimensionnelle de la protéine. En biochimie, il est donc important d'avoir des conditions proches du système dans lequel la protéine croît habituellement si on veut obtenir un bon repliement.

## 3.2 Attribution

L'attribution est une étape clé à la détermination de la structure des protéines. En effet, lorsque chaque signal du HSQC est attribué, il est aisé d'obtenir des mesures rapides sur les caractéristiques de liaison. Cependant l'attribution est très longue à effectuer. Les différentes étapes de l'attribution sont les suivantes : mesure des spectres 1D, 2D et 3D, traitement des données, attribution semi-automatique et attribution des chaînes latérales. Ces étapes seront décrites dans ce chapitre avec les résultats obtenus pour chacune d'elles.

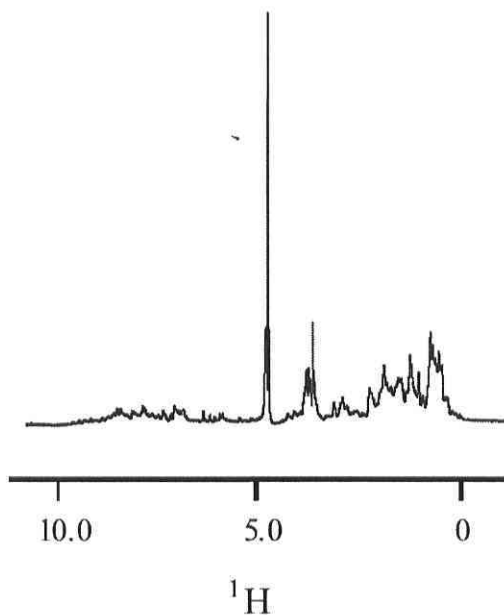
### 3.2.1 Spectre 1D du proton

Les spectres à une dimension sont des outils très pratiques aux travaux préliminaires. Par exemple, le rapport signal sur bruit (S/B) peut être sommairement évalué. Ce rapport dépend de la concentration de la protéine ainsi que des largeurs des raies spectrales. Les protéines de haut poids moléculaire possèdent des largeurs de bandes plus grandes ainsi qu'un S/B plus faible. La largeur de raie est reliée au temps de corrélation rotationnel de la protéine, et celui-ci dépend de la viscosité de la solution et de la température. Si le rapport S/B est trop faible, il est vain de mesurer des spectres 2D ou 3D. Les conditions de mesures de l'échantillon doivent être changées afin d'obtenir un S/B plus satisfaisant.

La dispersion des résonances de la protéine indique l'intégrité de la protéine dans les conditions choisies. Les protéines dénaturées ou qui sont nativement « non-structurées » montrent des déplacements chimiques similaires à de courts peptides (Cavanagh et al., 1996). Les déplacements chimiques des protéines repliées vont varier dus aux champs magnétiques anisotropes de groupes carbonyle ou aromatiques à proximité. La protéine peut ne pas être repliée ou avoir une structure peu stable si une faible dispersion de signaux est observée. L'examen de la dispersion des déplacements chimiques indique aussi la facilité avec laquelle l'attribution des résonances va être effectuée par la méthode du séquençage. Cette méthode sera expliquée à la section 3.1.3. Si le chevauchement dans la région HN est significatif, l'attribution séquentielle sera difficile.

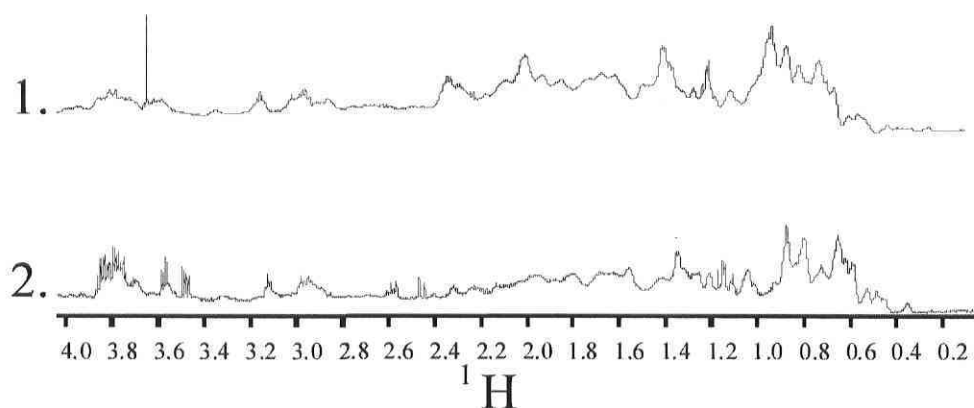


La pureté de l'échantillon peut aussi être évaluée à partir du spectre 1D. Si des impuretés organiques sont présentes, des signaux intenses et étroits seront observés. Finalement, ce spectre démontre bien le fait que des spectres à plusieurs dimensions doivent être mesurés. En effet, aucun signal ne peut être attribué dans un spectre 1D contenant plusieurs centaines de protons. Le spectre 1D du domaine tandem PDZ2/3 de PTP-BL non marqué est montré à la Figure 3.1 les signaux semblent bien répartis et il ne semble pas avoir de dégradation.



**Figure 3.1** Spectre 1D du proton du domaine tandem PDZ2/3 de PTP-BL non marqué.

La Figure 3.2 montre la comparaison entre les spectres 1D du domaine tandem PDZ2/3 de PTP-BL avec le domaine PDZ2. Il est possible d'observer que le spectre du domaine PDZ2 est beaucoup mieux défini que le spectre du domaine tandem PDZ2/3. Le domaine PDZ2/3 de PTP-BL est environ deux fois plus gros que le domaine PDZ2. Il est donc normal qu'il tourne plus lentement en solution et que les bandes soient plus larges, se perdant plus facilement dans le bruit. Dans le cas où les domaines seraient séparés par un lieur et évolueraient seuls en solution, les domaines tourneraient environ à la même vitesse que les deux domaines séparés, et la largeur de raie serait donc environ la même pour le domaine tandem PDZ2/3 que pour le domaine PDZ2. La comparaison entre les deux spectres confirme que cette hypothèse n'est pas la bonne.



**Figure 3.2** Spectres 1D 1. Du domaine tandem PDZ2/3 de PTP-BL. 2. Du domaine PDZ2.

### 3.2.2 Spectres 2D

Comme les spectres 1D ne fournissent pas assez d'informations pour déterminer la structure complète des protéines, des spectres 2D doivent être mesurés. Cette section sur les spectres 2D comporte 3 sous-sections dont le HSQC du domaine tandem PDZ2/3, le HSQC du domaine PDZ2 et des HSQCs d'échantillons contenant des acides aminés spécifiquement marqués.

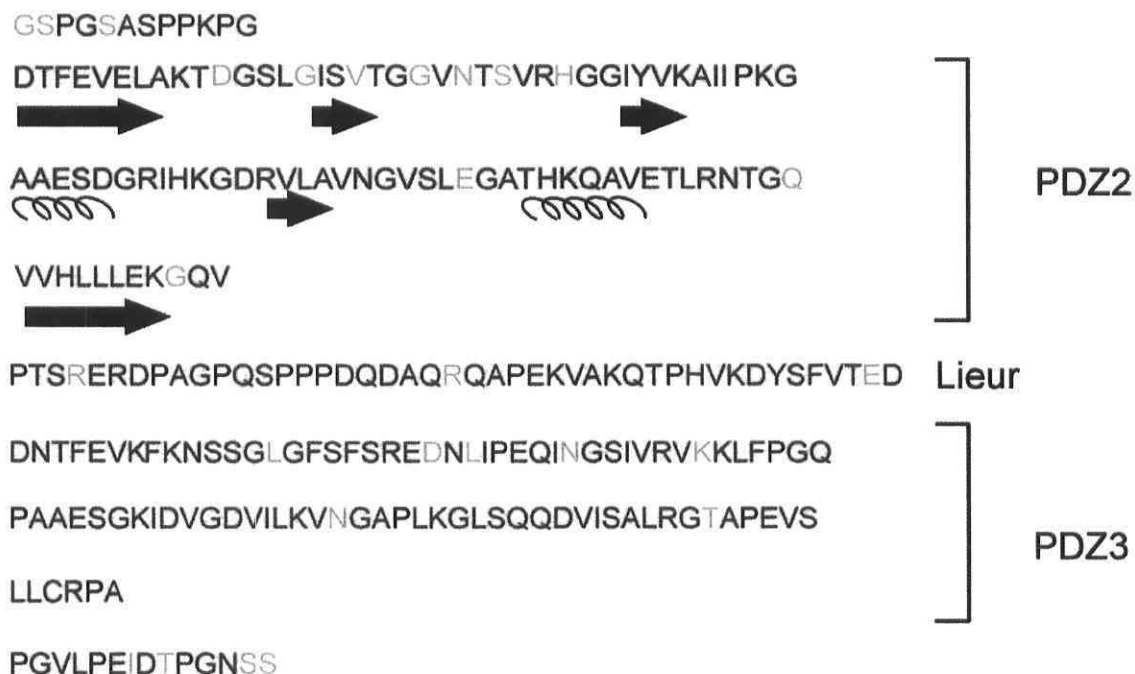
### 3.2.2.1 HSQC du domaine tandem PDZ2/3

Le spectre HSQC du domaine tandem PDZ2/3 contient en tout 227 signaux. De ces signaux, 25 représentent les chaînes latérales de l'asparagine et de la glutamine. Ils se situent dans le coin supérieur droit du spectre (indiqué en bleu) et ils ne sont pas attribués. Les signaux des chaînes latérales seront attribués lors de l'acquisition du HCCH-COSY (Bystrov, 1976; Pardi et al., 1984; Wagner et al., 1987) et HCCH-TOCSY (Montelione et al., 1992; Vuister et al., 1994). Les chaînes latérales de l'asparagine et de la glutamine possèdent un motif ressemblant à celui de l'acide aminé soit un carbonyle ainsi qu'une amine. Dans le spectre HSQC, la fréquence pulsée pour l'amine est de 7 à 11 Hz, ce qui inclut la fréquence du bout de la chaîne de l'asparagine et de la glutamine. Un changement de pH peut aussi faire apparaître des signaux dus à la lysine, l'histidine et le tryptophane (qui n'est pas présent dans la séquence). Les signaux des trois premiers acides aminés sont très difficiles à assigner car ils sont en dehors de l'échelle du spectre. Ils sont donc repliés en ce cas dans le milieu du spectre.

Comme l'attribution du domaine PDZ2 de PTP-BL avait déjà été effectuée (Walma et al., 2004), il a donc été possible d'utiliser les données se trouvant sur le site du BMRB afin de faciliter l'attribution. En effet, la moitié du domaine PDZ2 a été attribué de cette façon. Les signaux de PDZ2 n'ont été qu'à moitié attribués pour deux raisons : les déplacements chimiques indiqués dans la BMRB étaient différents du domaine libre en solution car il avait été mesuré avec le peptide dérivé d'APC. De plus, le domaine PDZ2 dont l'attribution avait été effectuée était perturbé par le lieu et PDZ3 dans le domaine tandem PDZ2/3 (voir section 3.1.3.2).

L'attribution a donc été effectuée à 88% et les déplacements chimiques sont montrés à l'Annexe. Les HN, C $\alpha$ , N, C et C $\beta$  sont attribués respectivement à 82,4%, 93,2%, 82,4%, 72% et 48,4%. Les prolines ne possédant pas de HN lorsqu'elles forment le lien peptidique n'apparaissent pas dans le spectre. De plus, le premier résidu est protonné et n'est aussi pas attribués. À la Figure 3.3 l'attribution effectuée a été indiquée sur la séquence. Les

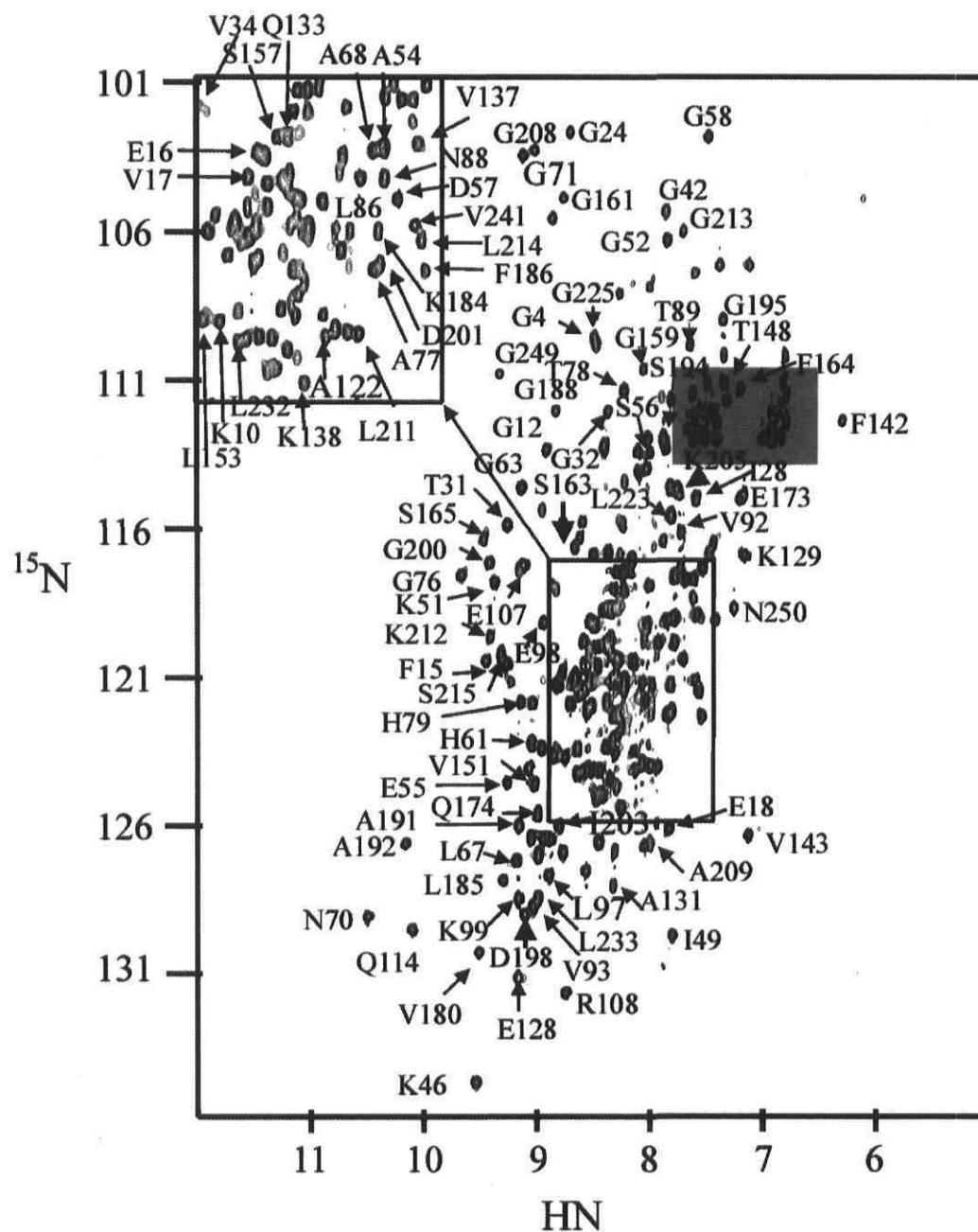
résidus qui n'ont pas été attribués ont été indiqués en orange, tandis que les résidus en noir représentent les prolines et les résidus en vert sont ceux qui ont été attribués. Il est intéressant de remarquer que la plupart des résidus non attribués sont situés dans les boucles. Les boucles sont réputées pour être moins définies dans les spectres RMN car elles sont plus mobiles et en contact avec le solvant.



**Figure 3.3** Résidus attribués sur la séquence du domaine tandem PDZ2/3.

Le domaine PDZ2 a été attribué à 83% alors que le domaine PDZ3 a été attribué à 93%. L'attribution n'a pas été faite à 100% pour plusieurs raisons. En premier lieu, les domaines PDZ2 et PDZ3 sont très similaires (33% d'identité) ce qui cause beaucoup de chevauchement. Cette attribution aurait peut-être pu être plus élevée si l'échantillon obtenu avait été plus concentré. Cependant, lorsqu'un échantillon est trop concentré, une certaine agrégation peut être observée, ce qui réduit la qualité des spectres. La concentration mécanique d'un échantillon comporte, de plus, un risque pour sa dégradation et sa détérioration. L'échantillon peut se contaminer par le matériel utilisé et chaque fois qu'il est manipulé, il se dégrade en partie. À la Figure 3.4 le HSQC attribué des domaines

tandem PDZ2/3 peut être observé. Il est à noter que ce spectre a été attribué à l'aide de tous les autres spectres mentionnés dans le Chapitre 2.



**Figure 3.4** Attribution du HSQC du domaine tandem PDZ2/3 de PTP-BL.

Il est à noter que certains signaux sont absents du spectre. Les acides aminés qui sont davantage en contact avec l'eau risquent de subir des processus d'échange et ainsi d'être de moindre intensité. Ainsi, certains acides aminés peuvent être situés à des endroits de la protéine où les mouvements moléculaires sont plus lents. Comme la largeur de raie est inversement proportionnelle au temps de rotation, ces signaux sont moins intenses.

Certains signaux n'ont pas été attribués car ils ne sont pas présents dans le spectre HNCA. Une des raisons possibles est que les spectres 3D sont moins sensibles que les spectres 2D. Plusieurs facteurs peuvent expliquer cette moindre sensibilité. Par exemple, une perte progressive de signal lorsque le délai variable augmente due à la relaxation durant les périodes de temps d'incrémentations se produit. D'autre part, l'intégration de l'intensité d'un signal dans le spectre 1D est associée avec plusieurs signaux dans le spectre multidimensionnel (le HNCA possède deux signaux). Plus d'incrémentations de  $t_1$  doivent être enregistrés pour augmenter la résolution ce qui fait que la relaxation durant la période d'évolution  $t_1$  cause une diminution du signal et une perte de sensibilité.

Comme il y a beaucoup de chevauchement dans ce spectre on doit mesurer des spectres 3D afin d'obtenir plus de renseignements et d'effectuer l'attribution. Effectivement le centre du spectre est très encombré.

### 3.2.2.2 HSQC de PDZ2

Comme les deux domaines étaient très similaires, il a été important de mesurer le HSQC du domaine PDZ2. Plusieurs signaux possédant des déplacements chimiques similaires ont pu être différenciés de cette façon. Le HSQC de PDZ2 est présenté à la Figure 3.5.

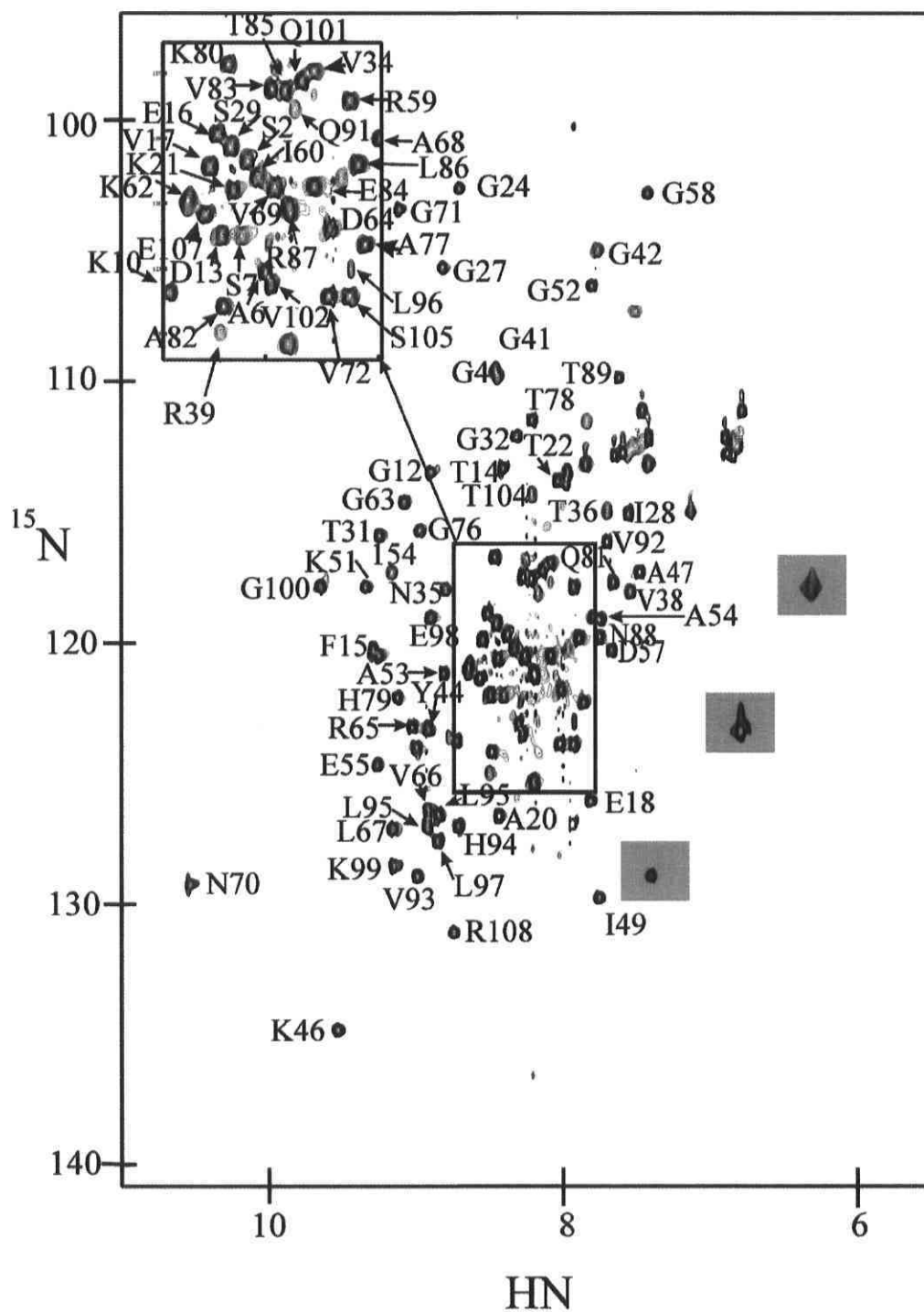
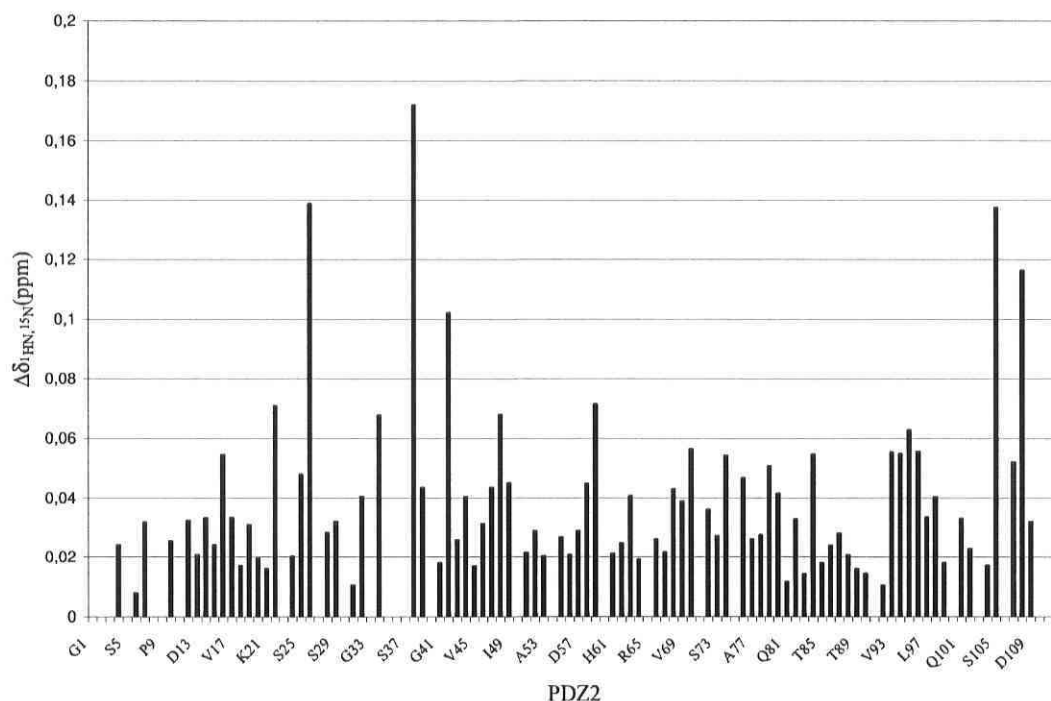


Figure 3.5 Spectre HSQC de PDZ2.

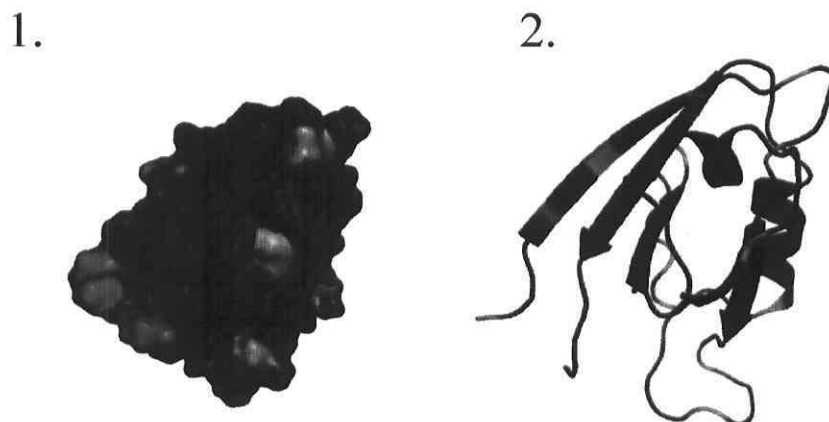
Les signaux indiqués en rouge ne sont pas présents sur le HSQC du domaine tandem PDZ2/3. Les différences de déplacement chimique observées peuvent être associées aux conditions de mesures ou par des perturbations induites par le lieu et PDZ3. À la Figure 3.6 un graphique représentant ces perturbations est observé. L'axe en « y » représente le  $\Delta\delta_{\text{HN},^{15}\text{N}}$  (voir section 2.10) tandis que l'axe en x représente les acides aminés présents dans le domaine PDZ2.



**Figure 3.6** Comparaison du domaine PDZ2 avec PDZ2 dans le domaine tandem PDZ2/3.

Ces perturbations sont aussi schématisées à la Figure 3.7. Il est possible d'observer qu'elles se situent dans le dos de la protéine, c'est-à-dire à l'opposé du début du domaine. Le premier acide aminé est indiqué en vert et le dernier en rouge.





**Figure 3.7** Changements observés sur le domaine PDZ2 lors de la comparaison des déplacements chimiques avec le domaine tandem PDZ2/3.1. Structure de PDZ2 en surfaces de Van der Waals. 2. Structure de PDZ2 en rubans.

### 3.2.2.3 HSQCs d'acides aminés spécifiquement marqués

Les spectres 2D d'acides aminés spécifiquement marqués servent à identifier des acides aminés n'ayant pas encore été attribués. Les acides aminés difficilement identifiables possédant des  $C_{\alpha}$  similaires sont difficiles à différencier (voir section 3.1.3.1) et parfois, pour certains, le CBCA(CO)NH n'est pas assez intense et ne donne pas les informations voulues (voir section 3.2.3.2).

Le spectre d'un échantillon contenant des glycines spécifiquement marquées peut être observé à la Figure 3.8.1. Le spectre de la glycine peut sembler inutile car elle n'est pas un acide aminé difficile à attribuer dû à la fréquence caractéristique de son  $C_{\alpha}$  (45 ppm) (voir section 3.2.3). Les glycines présentes dans le HSQC du domaine tandem PDZ2/3 et dans le spectre HNCA étaient donc déjà attribuées. Cependant, la sérine, qui est difficile à attribuer, apparaît aussi dans le spectre ce qui permet d'attribuer quelques sérines. 6 sérines ont été attribuées de cette façon. Ceci est causé par le métabolisme des acides aminés (McIntosh et al., 1990). En effet, plusieurs acides aminés sont la seule source disponible d'azote supportant la croissance bactérienne. Le métabolisme peut donc changer la position finale du  $^{15}\text{N}$  introduit dans un acide aminé dans un milieu minimal ne contenant

aucune autre source d'azote. Par exemple, en introduisant des glycines marquées à  $^{15}\text{N}$  dans le milieu, il sera possible d'observer des sérines spécifiquement marquées. Les informations contenues dans le Tableau 3.1 résumant le métabolisme des acides aminés ont été tirées de l'article de McIntosh et al. (1990).

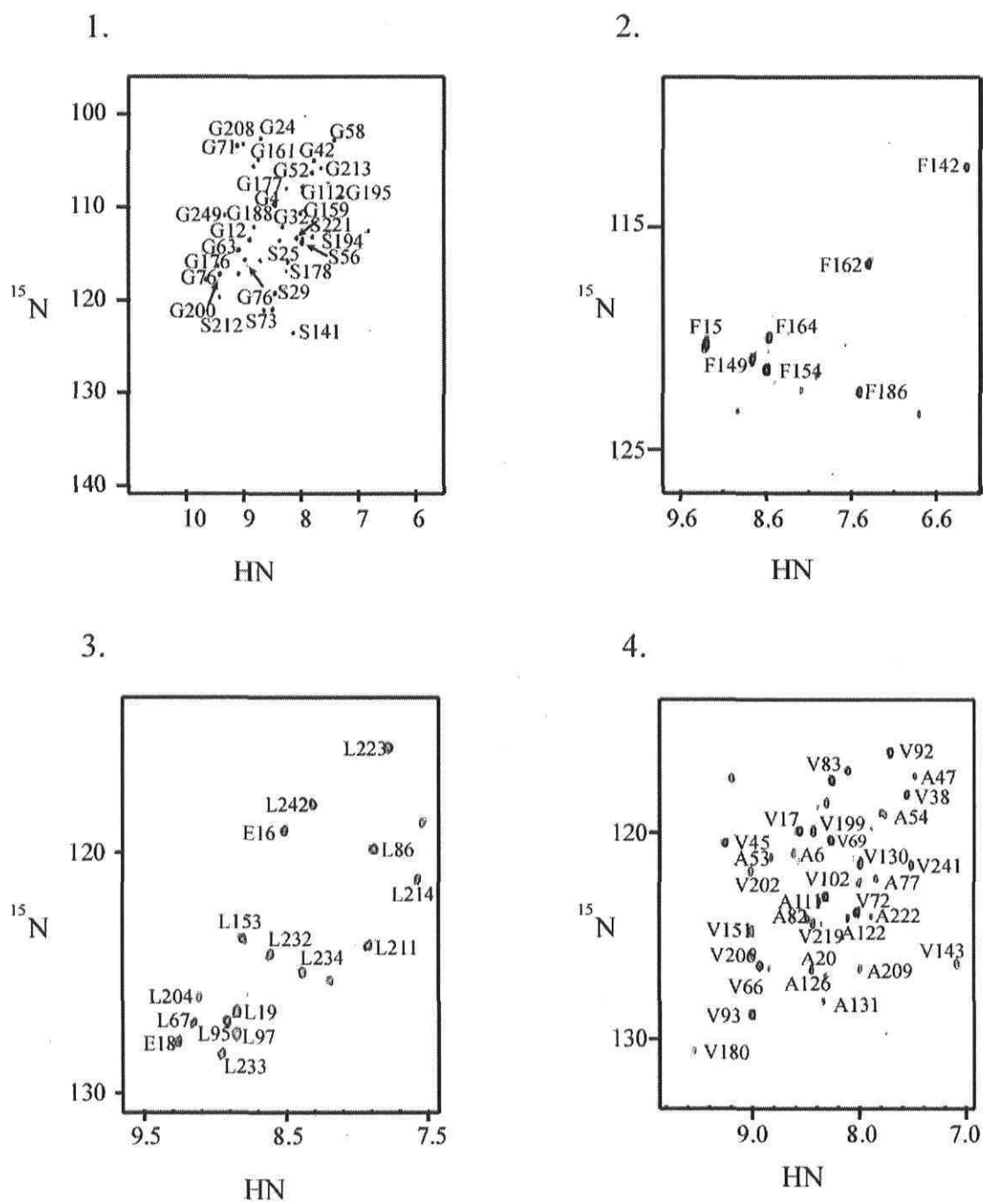
Acide aminé	Source d'azote
Gly	Ser <sup>5</sup>
Phe	Glu
Val	Glu, Ala
Leu	Glu

**Tableau 3.1** Résumé du métabolisme des acides aminés choisis pour effectuer des spectres spécifiquement marqués.

Le spectre des phénylalanines spécifiquement marquées peut être observé à la Figure 3.8.2. Ce spectre contient en tout 7 phénylalanines. À la Figure 3.8.3 et 3.8.4 on observe respectivement les spectres des leucines et des valines spécifiquement marquées. Pour ce qui est du spectre des leucines, il est possible d'observer de l'acide glutamique et pour la valine il est possible d'observer des alanines.

L'attribution des acides aminés est un processus itératif. En rassemblant des indices, il est possible d'affirmer que tel acide aminé est bien une alanine ou une valine. Cependant, la fréquence du résidu précédent est aussi très importante à l'attribution et donc un acide aminé peut être changé de position plusieurs fois. Plusieurs acides aminés déjà attribués ont été confirmés par ces spectres et de nouveaux ont été attribués. Par exemple, 3 valines et 2 alanines ont été attribuées grâce au spectre 3.8.4.

<sup>5</sup> Il est à noter que la glycine peut aussi être une source d'azote pour la sérine.

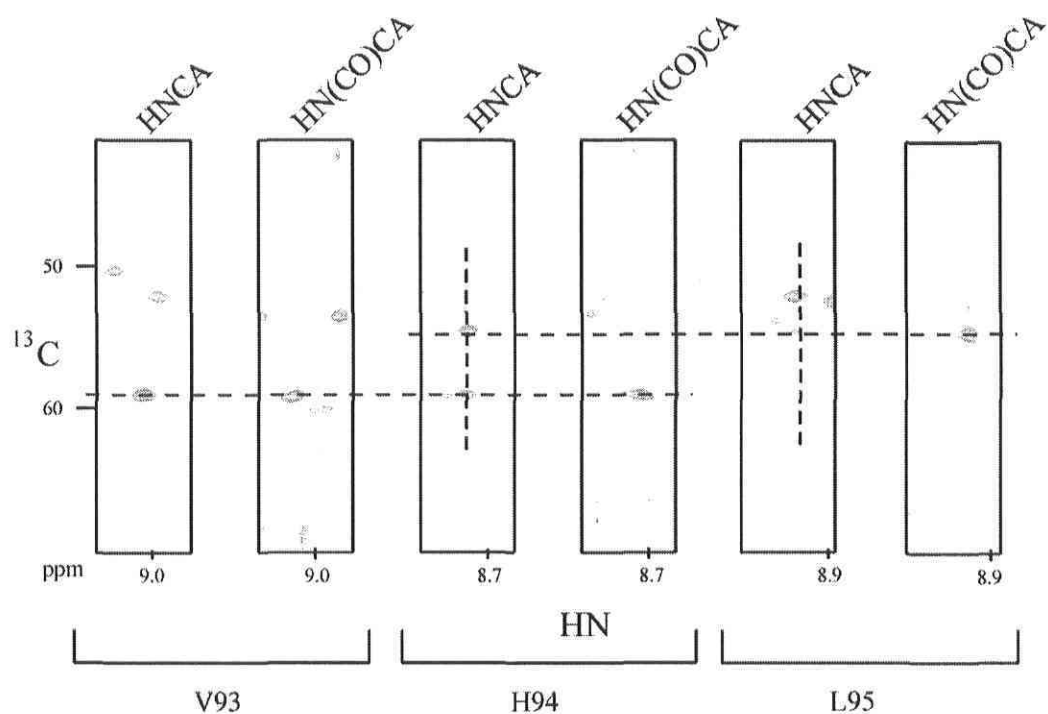


**Figure 3.8** Spectres contenant des acides aminés spécifiquement marqués : 1. glycines, 2. phénylalanines, 3. leucines et 4. valines.

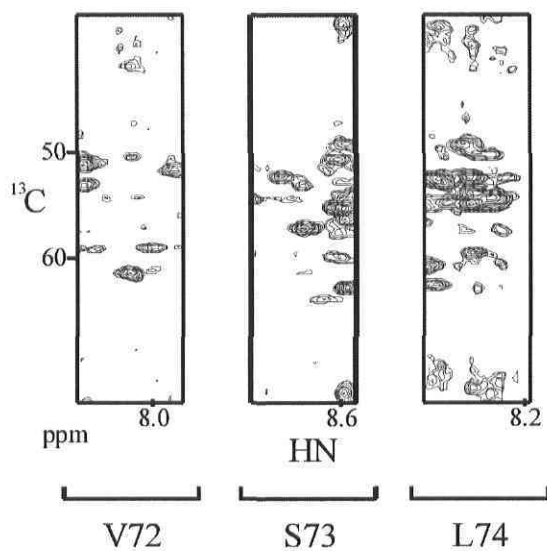
### 3.2.3 Spectres 3D

#### 3.2.3.1 HNCA

Les spectres 3D permettent d'obtenir des informations supplémentaires sur un acide aminé à partir d'une troisième dimension spectrale. Afin d'effectuer l'analyse, des tranches sont coupées du spectre 3D. Ces tranches, dont un exemple est présenté à la Figure 3.9, permettent d'observer, dans le cas du HNCA, les  $C_\alpha$  et  $C_{\alpha-1}$ . L'attribution séquentielle est effectuée à partir de ce spectre. L'attribution séquentielle consiste dans une tranche donnée du HNCA de démystifier quelle tranche est avant et après. Une tranche représenterait un acide aminé. Le  $C_\alpha$  est donc corrélé avec le  $C_\alpha$  du résidu précédent. C'est-à-dire que la fréquence de l'hydrogène est la même pour les deux  $C_\alpha$ . Le signal du  $C_\alpha$  est habituellement plus intense que celui du  $C_{\alpha-1}$ . En trouvant le  $C_\alpha$  correspondant au  $C_\alpha$  du résidu précédent il est possible de relier plusieurs tranches entre elles et donc de « marcher » ainsi à travers la séquence d'acides aminés comme il est schématisé à la Figure 3.9. Les bandes présentées dans la Figure 3.9 ne sont pas chevauchées et ont été faciles à attribuer. Dans la Figure 3.10 un exemple de bandes très chevauchées est présenté. Ces bandes démontrent bien que des spectres supplémentaires sont nécessaires comme des spectres d'acides aminés spécifiquement marqués, le HSQC de PDZ2 ainsi que les spectres HN(CO)CA et CBCA(CO)NH.



**Figure 3.9** Bandes peu chevauchées représentant l'attribution séquentielle à partir du HNCA et du HN(CO)CA. Les résidus V93, H94 et L95 sont présentés reliés entre eux.



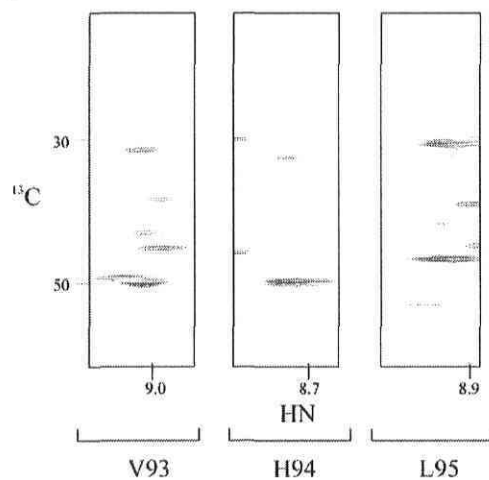
**Figure 3.10** Exemple de raies très chevauchées du spectre HNCA.

### 3.2.3.2 HN(CO)CA

Le HN(CO)CA permet de confirmer le  $C_{\alpha-1}$ . Parfois le  $C_{\alpha-1}$  n'est pas présent ou le  $C_{\alpha}$  et le  $C_{\alpha-1}$  sont chevauchés. Un exemple où le  $C_{\alpha}$  et le  $C_{\alpha-1}$  sont chevauchés est présenté à la Figure 3.9. La Val93 ne possède qu'un seul signal. Cela est dû au fait que le résidu précédent est très similaire à Val93. En effet, il s'agit de Val92. Ces cas sont observés lorsque l'environnement chimique des atomes observés est très similaire. Cela est confirmé par le HN(CO)CA. À la Figure 3.9 des exemples de HN(CO)CA sont montrés.

### 3.2.3.3 CBCA(CO)NH

Le CBCA(CO)NH sert à mieux démystifier le résidu précédent un autre résidu. On observe donc le  $C_{\beta-1}$  du résidu cible. Le  $C_{\beta}$  de 123 résidus a été trouvé à l'aide du CBCA(CO)NH. Le  $C_{\beta}$  est davantage caractéristique que le  $C_{\alpha}$  (Voir Tableau 2.1). À la Figure 3.11, un exemple de raies est indiqué.



**Figure 3.11** Exemple de raies pour le CBCACONH.

### 3.3 Attribution semi-automatique

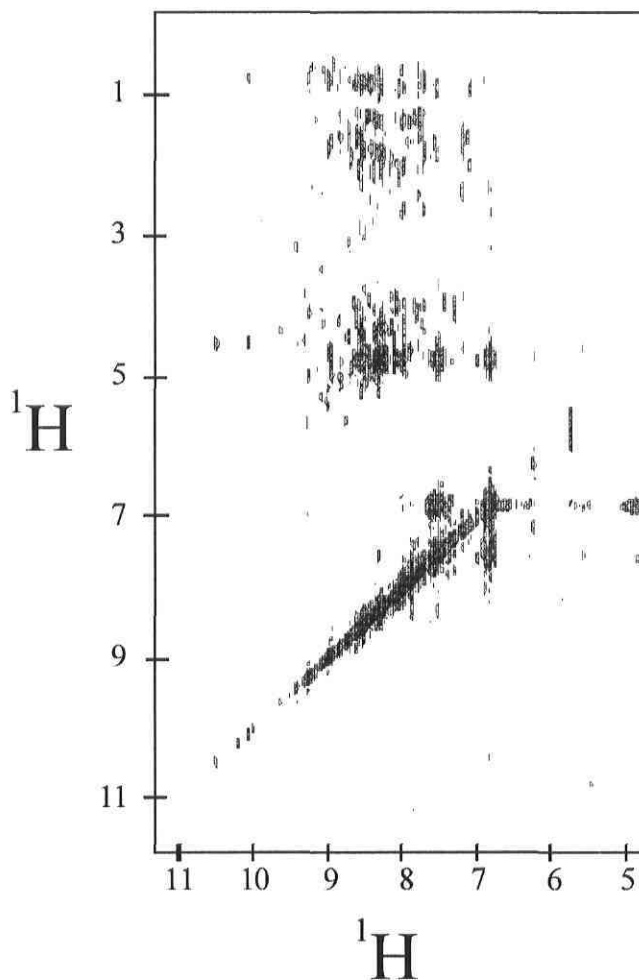
Afin de simplifier l'attribution des spectres obtenus par RMN, des programmes ont été développés afin d'effectuer l'attribution semi-automatique. Deux de ces programmes ont été essayés, c'est-à-dire Monte et Mars. Cependant, cette attribution a posé certains problèmes. Dans les cas comme celui-ci, c'est-à-dire de deux domaines étant très similaires, les résultats n'ont pas été très satisfaisants. L'algorithme de ces programmes a été fait pour des domaines seuls et non pour deux domaines similaires séparés. Comme plusieurs signaux sont identiques d'un domaine à l'autre et même dans les mêmes domaines, les algorithmes n'ont pas été très utiles dans ce cas.

### 3.4 Détermination de la structure du domaine tandem PDZ2/3 par RMN

La détermination de la structure du domaine tandem PDZ2/3 de PTP-BL n'a pas pu être effectuée pour plusieurs raisons. En premier lieu, des spectres NOESY satisfaisants à l'obtention de contraintes structurales (entre autres les distances interprotons) n'ont pas pu être obtenus. Plusieurs HSQC-NOESY ont été mesurés (6 spectres) et ils n'étaient pas de qualité souhaitable. À la Figure 3.12, un exemple de spectre NOESY-HSQC est montré. Cependant, seules quelques corrélations interprotons dans l'acide aminé cible sont observables. Habituellement, dans ce type de spectre, il est possible de voir des corrélations avec les protons d'autres acides aminés. D'autres solutions ont été trouvées afin de déterminer si les domaines PDZ2 et PDZ3 s'associaient. Ces solutions seront vues plus en détail au Chapitre 4.

La détermination de la structure des protéines par RMN pose souvent plusieurs problèmes. Les échantillons sont extrêmement coûteux dus au marquage isotopique (3000\$ pour 2 ml, 1 mM), la durée de la mesure des spectres RMN est très longue ce qui cause des problèmes pour obtenir du temps de mesure et les échantillons de protéines se dégradent souvent trop rapidement. Tous ces problèmes font que la détermination de la structure d'une protéine peut prendre de plusieurs mois à plusieurs années.

Depuis plusieurs années, les données sur les structures déterminées par RMN et par cristallographie ont été entreposées dans des banques de données. Plusieurs bioinformaticiens tentent d'utiliser ces mesures afin de prédire la structure des protéines afin de pouvoir seulement avec la structure primaire déterminer la structure secondaire, tertiaire et quaternaire des protéines.



**Figure 3.12** Exemple du 3<sup>e</sup> spectre NOESY-HSQC.



## **Chapitre 4**

### **Caractéristiques de liaison du domaine tandem PDZ2/3 de PTP-BL**

#### **4.1 Interactions protéines-ligands**

Les interactions entre deux molécules sont très importantes dans plusieurs processus de reconnaissance. Par exemple, on retrouve des complexes protéine-protéine, protéine-peptide, protéine-médicament et protéine-acide nucléique. Le contrôle de ces interactions vient des sensibles spécificités entre les partenaires de liaison. Il a été démontré que les interactions protéine-protéine sont conduites énergétiquement seulement par un petit sous-ensemble de résidus (Clackson et al., 1995). Ces observations impliquent qu'il peut être possible d'imiter les interactions de grosses protéines par de petits peptides et ultimement à les convertir en médicaments avec des caractéristiques de liaison hautement spécifiques. Comme les complexes protéines-protéines sont plus difficiles à étudier par RMN à cause du temps de corrélation rotationnel plus élevé (ce qui augmente la largeur de raie), seulement de courts peptides dérivés de ces protéines sont en général utilisés.

## 4.2 La bioinformatique

La bioinformatique connaît une très grande popularité de nos jours. En passant par la phylogénie (une application est de trouver l'origine d'une espèce) jusqu'à la prédiction de structure d'une protéine, elle permet des applications qui étaient jusqu'alors insoupçonnées. Par exemple, grâce à la phylogénie, les bioinformaticiens ont pu découvrir l'origine du syndrome respiratoire aigu sévère (SRAS) (He et al., 2004) et même remonter jusqu'à la personne qui a emmené le SRAS au Canada. Dans le cadre de ce mémoire, la bioinformatique va servir comme outil. Comme les structures du domaine tandem PDZ2/3 et de PDZ3 libre n'ont pas encore été élucidés, la bioinformatique a permis de prédire ces structures. Afin de mieux schématiser les perturbations induites par les titrations ainsi que de poser une hypothèse sur la structure tertiaire et la structure quaternaire du domaine tandem PDZ2/3, la structure de PDZ3 a été prédite par la bioinformatique.

La structure secondaire a été, en premier lieu, prédite. Trois serveurs ont été choisis à partir de EVA<sup>6</sup>. EVA est un serveur qui permet d'évaluer les serveurs de prédiction de structure 2D et 3D. Les calculs sont basés à partir de l'exactitude de la prédiction pour chaque résidu, pour chaque segment et pour la classe de protéine prédite. Trois serveurs ont donc été choisis : Sable2 (Adamczak et al., 2005), PSIPred (Jones, 1999) et Scratch (Baldi et al., 1999). La structure secondaire a été prédite pour le domaine PDZ3 à partir de la séquence primaire (la chaîne polypeptidique) par les trois programmes et finalement une séquence consensus a été choisie. Une séquence, représentant le consensus des trois séquences plus ou moins différentes déterminées à partir des programmes mentionnés, est choisie. Cette séquence consensus (voir Figure 4.1) contient 5 feuillets  $\beta$  et une hélice  $\alpha$ . Habituellement les domaines PDZ contiennent deux hélices  $\alpha$ . Comme le domaine PDZ3 contient beaucoup plus de prolines que le domaine PDZ2 (respectivement 6 pour 1), il est possible que la structure de la deuxième hélice soit détruite. Les prolines sont réputées pour détruire la structure car leur angle  $\phi$  est rigide, ce qui ne leur permet pas d'adopter la structure en hélice. Cette hélice aurait pu être située à l'endroit marqué en orange dans la Figure 4.1

---

<sup>6</sup> <http://cubic.bioc.columbia.edu/eva/>

comme certains serveurs l'ont indiqué (les deux alanines) mais comme une proline se situe juste avant les deux alanines, cette hélice pourrait être détruite.

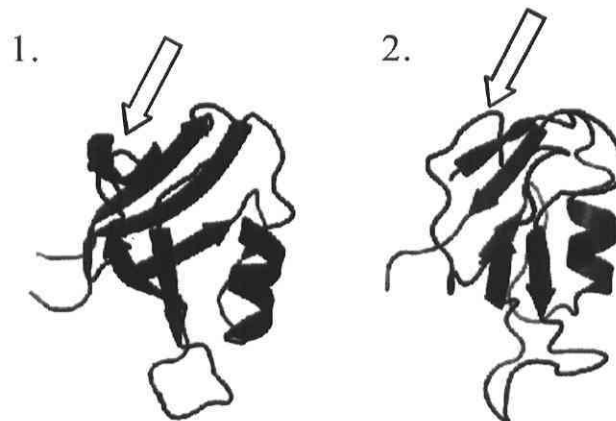
```

NTFEVKLKFKNSSGLGFSFREDNLIPEQINGSIVRVKKLFPGQPAAE
BBFFFFFFBBBBBBFFFFFFBBBBBBBBBBBBBBFFFFFFBBBBBB
SGKIDVGDVILKVNGAPLKGLSQQDVISALRGTAPEVSLLLCRPA
BBBBBBBBBBFFFFFFBBBBBBBBBBHHHHHHHHHHBBBBFFFFFFBB

```

**Figure 4.1** Prédiction de la structure 2D de PDZ3. Les lettres B, F et H sont utilisées pour boucle, feuillet et hélice, respectivement.

La structure tertiaire a ensuite été prédite. Le serveur EasyPred3D (Lambert et al., 2002) a été choisi à partir de EVA. La méthode utilisée est le modelage par homologie. La méthode de modelage par homologie est la meilleure méthode de prédiction de structure. En effet, il y a plus d'applications disponibles que de développements. À la Figure 4.2 cette prédiction est présentée. Cette structure est comparée avec celle de PDZ2. Il est intéressant de remarquer que PDZ2 contient 2 hélices  $\alpha$  alors que PDZ3 en contiendrait une seule. Cette structure sera utilisée afin de schématiser les perturbations induites dans PDZ3 lors des titrations. Il est cependant important de mentionner qu'il s'agit seulement d'une prédiction et que la structure de PDZ3 devra être déterminée dans le futur.

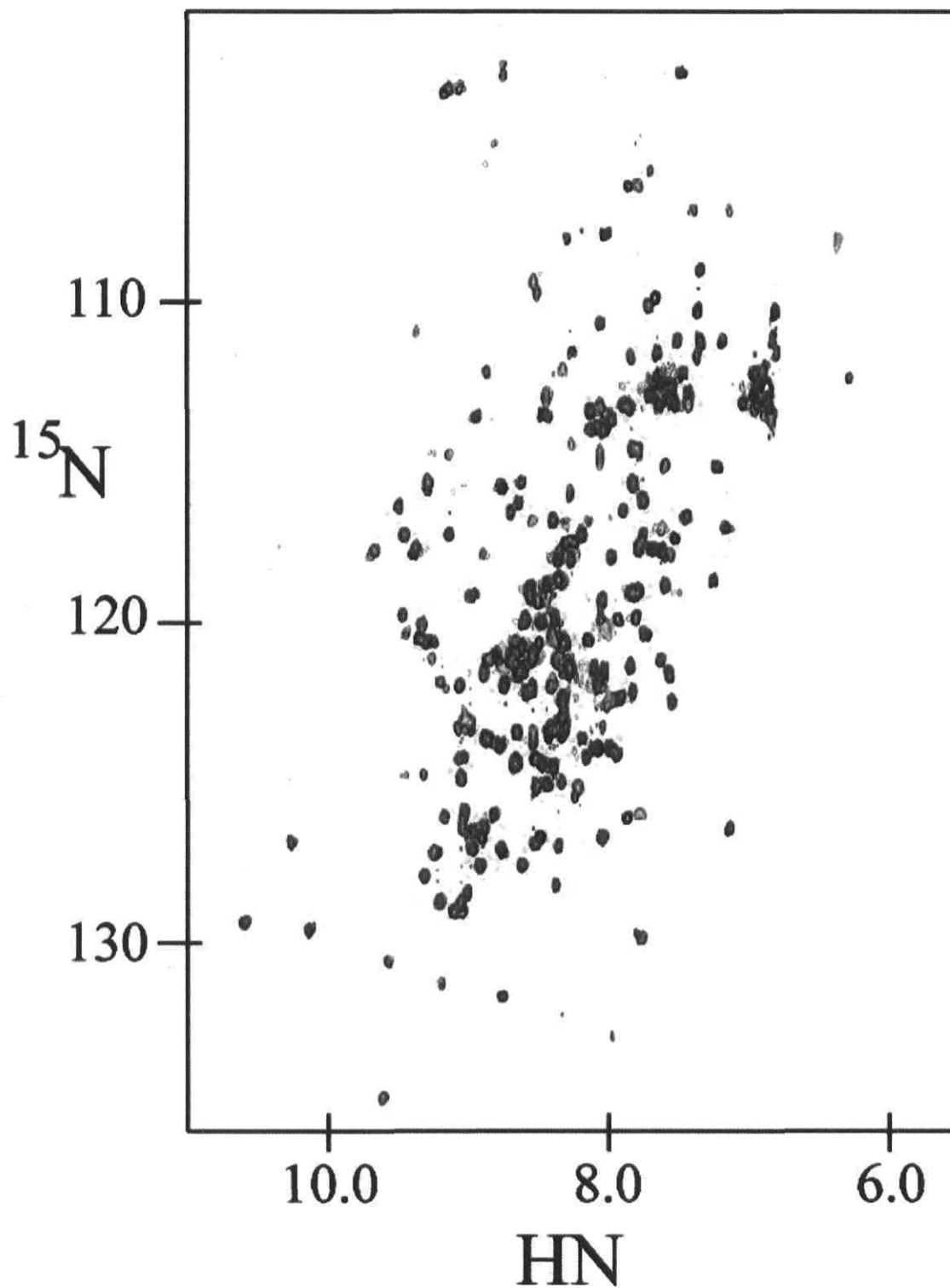


**Figure 4.2** Structures 3D de 1. PDZ2 2. PDZ3

### 4.3 Titrations

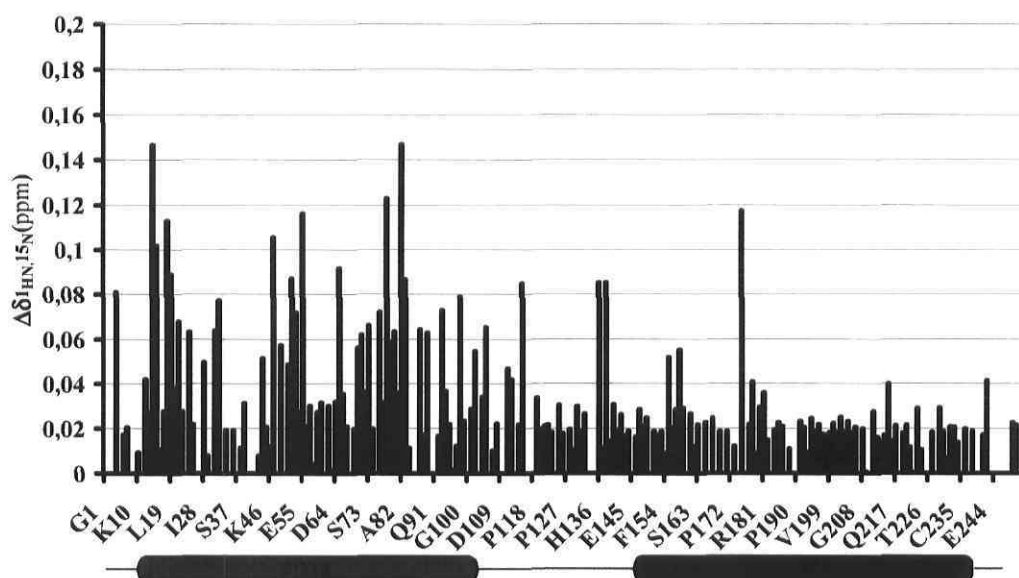
Afin d'étudier les caractéristiques de liaison du domaine tandem PDZ2/3 de PTP-BL, il a été possible d'effectuer plusieurs titrations. Deux de ces titrations seront présentées dans le cadre de ce mémoire.

Dans la première titration, le domaine tandem PDZ2/3 a été titré avec le peptide dérivé d'APC (voir Tableau 1.1) dans une première étape. Cette titration permettait de vérifier si le peptide dérivé d'APC liait bien le domaine PDZ2 dans le domaine tandem PDZ2/3. De plus, il était intéressant de noter les perturbations obtenues pour le lieur ainsi que pour le domaine PDZ3. À la Figure 4.3 un exemple de cette étape de la titration est montré. Il est possible d'observer en rouge le domaine tandem PDZ2/3 lié avec le peptide et en noir le domaine tandem PDZ2/3 libre en solution. Lorsque les signaux ne sont pas superposés, un changement est observé.



**Figure 4.3** Exemple de la première étape du domaine tandem PDZ2/3 marqué à  $^{15}\text{N}$  par le ligand dérivé d'APC non marqué.

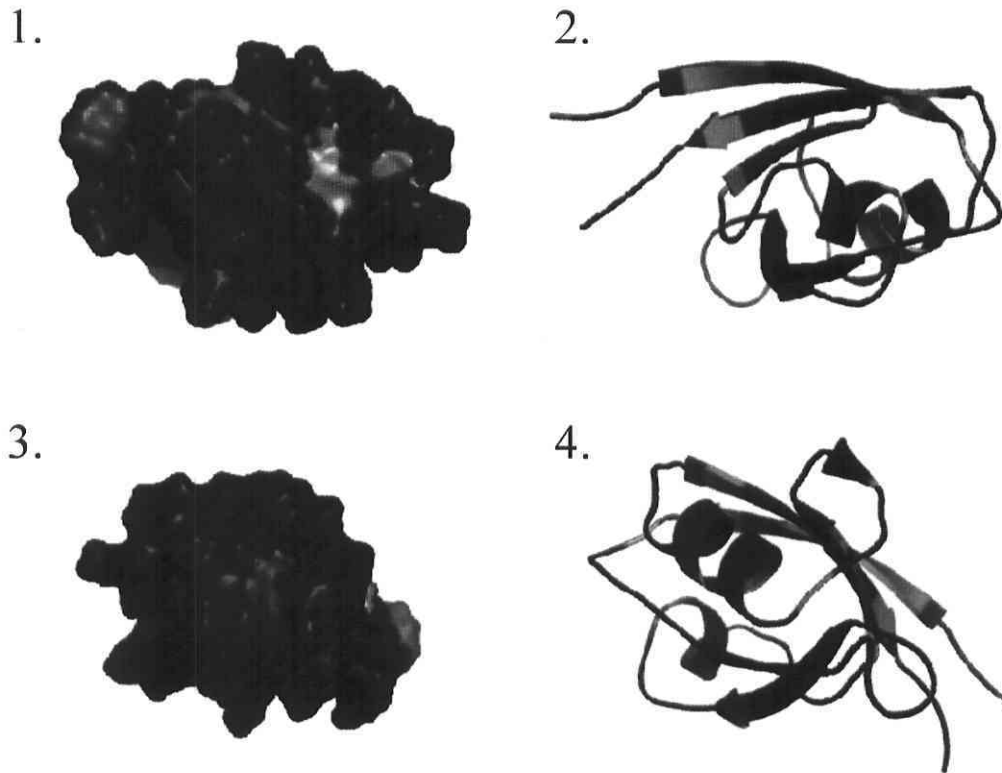
L'observation du spectre de la forme liée nous permet d'évaluer sommairement le  $K_d$ . En effet, la vitesse d'échange est rapide ( $K_d > 10^{-3}$  M) car un seul signal moyen est observé pour la forme libre et la forme liée (voir section 2.10). Des mesures dynamiques plus précises devront par contre être effectuées afin de trouver le  $K_d$  exact de la liaison. Le  $\Delta\delta$  est ensuite mesuré à partir des changements observés sur l'axe des protons et de l'azote. La Figure 4.4 montre graphiquement les perturbations observées dans le domaine tandem PDZ2/3 lors de l'addition du peptide dérivé d'APC.



**Figure 4.4** Perturbations observées sur le domaine tandem PDZ2/3 de PTP-BL lors de l'ajout du peptide dérivé d'APC.

Cette figure démontre deux points importants. En premier lieu, le peptide dérivé d'APC se lie bien à PDZ2 dans le domaine tandem PDZ2/3. De plus, ce peptide perturbe aussi PDZ3 et le lieu. Le fait que les deux domaines et le lieu soient perturbés indique qu'il y a une certaine coopérativité entre les deux domaines afin de lier le peptide. Dans la Figure 4.5, les perturbations observées ont été schématisées sur la structure de PDZ2. On observe que le site de liaison de PDZ2 est bien conservé. Les acides aminés contenus dans le site de

liaison de la littérature sont résumés au Tableau 4.1 ainsi que les perturbations dues aux peptides sur les deux domaines dans le domaine tandem PDZ2/3.



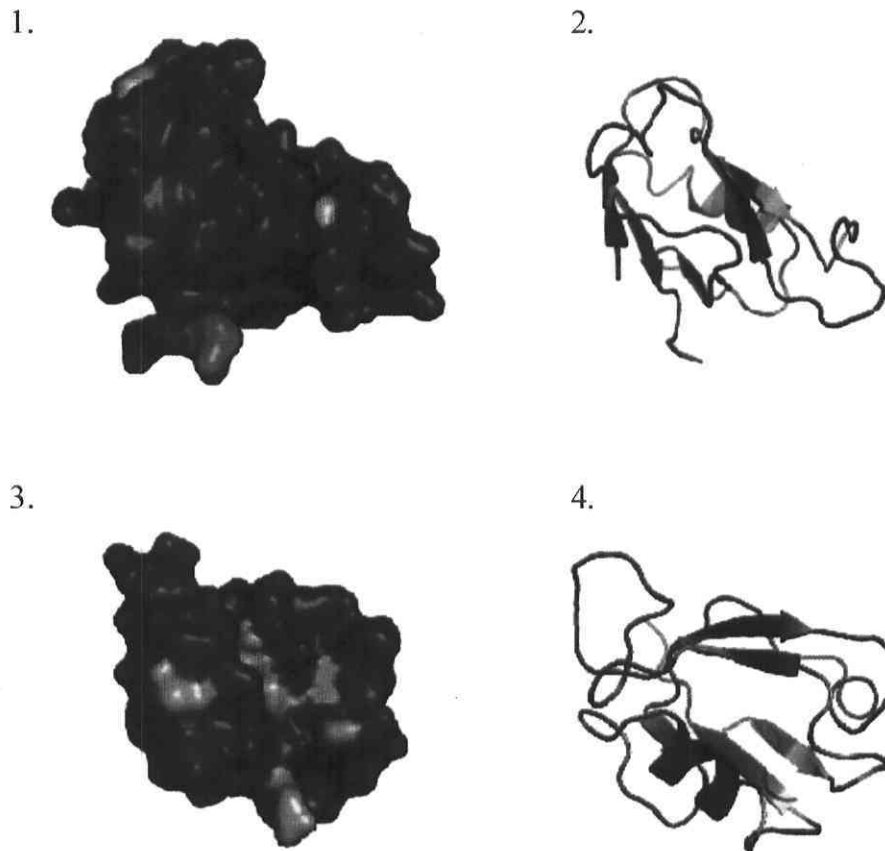
**Figure 4.5** Schéma des perturbations observées sur PDZ2. Les perturbations observées sont en orange. 1. Structure de PDZ2 avec surface de Van der Waals. 2. Structure de PDZ2 en rubans. 3. Structure 1, basculée de 90°. 4. Structure 2., basculée de 90°.

Peptide dérivé d'APC		Littérature	Peptide dérivé de PRK2	
PDZ2	PDZ3		PDZ2	PDZ3
G4	S115	K21	R39	L153
T14	H136	S25	K46	N156
F15	K138	G27	A53	G159
L19	K155	S29	S56	G161
A47	S158	K46		S163
A53	I175	T31		S165
T78	S215	H79		N169
A82		V83		F186
V83		L86		L211
T89		N88		G213
V93				R224
E98				E229

**Tableau 4.1** Tableau récapitulatif des titrations

Le fait que les acides aminés obtenus soient quelques peu différents de ceux de la littérature pour ce qui est du site de liaison est une preuve supplémentaire que les deux domaines coopèrent dans la liaison. Le site de liaison est dit conservé car si les acides aminés ne sont pas les mêmes pour la liaison, des acides aminés voisins sont impliqués. Par exemple, K21 est impliqué dans le site de liaison de la littérature. Pour ce qui est du site de liaison de PDZ2 dans le domaine tandem PDZ2/3, L19 est impliqué. Ces changements sont plutôt rassurants car le système étudié est différent de celui de la littérature. Si les deux domaines coopèrent dans la liaison, il est possible que d'autres acides aminés soient impliqués que ceux qui le sont seulement lorsque PDZ2 est libre en solution. De plus, le peptide utilisé contient 12 acides aminés au lieu de 10 comme le peptide ayant été utilisé dans la littérature. Lorsque des acides aminés sont ajoutés aux peptides, la spécificité de la liaison augmente. La liaison n'est vraiment efficace que lorsque la protéine entière est présente. Cependant des complexes aussi volumineux sont plus difficiles à analyser par RMN. À la Figure 4.6, les changements observés sur les acides aminés de PDZ3 sont schématisés. Il est possible de remarquer que les changements sont marquants dans le dos de la protéine. Le dos de la protéine est représenté par les résidus colorés en rouge et vert.



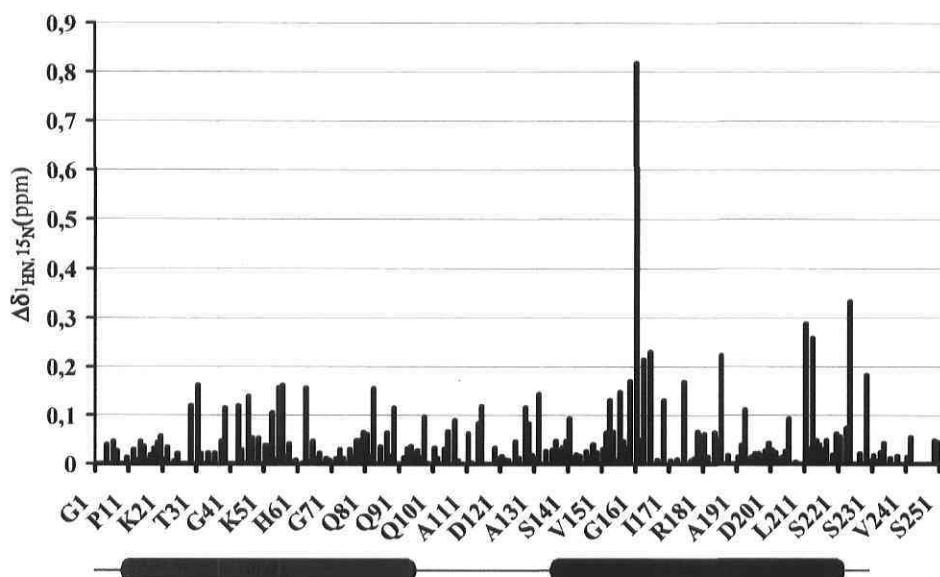


**Figure 4.6** Changements observés sur PDZ3 lors de l'ajout du peptide dérivé d'APC dans le milieu. 1. Structure de PDZ3 prédite avec surface de Van der Waals. 2. Structure de PDZ3 prédite en rubans. 3. Structure 1. basculée de 90°. 4. Structure 2. basculée de 90°.

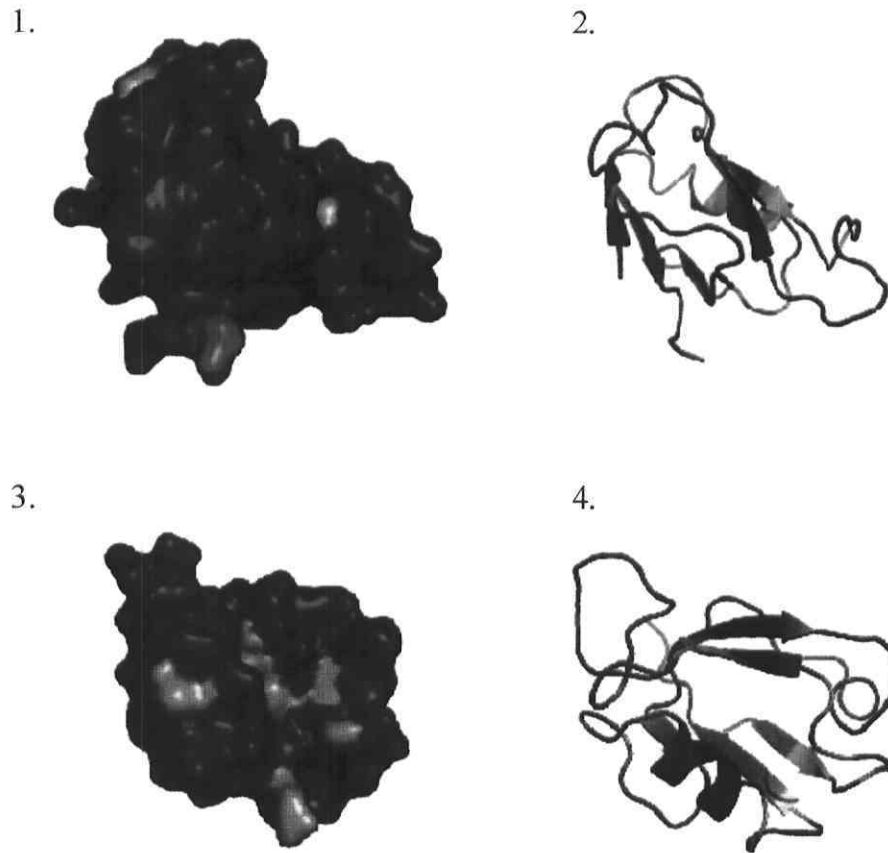
La deuxième étape de la titration, c'est-à-dire suite à l'ajout du peptide dérivé de PRK2, vérifie que le peptide se lie bien au domaine tandem PDZ2/3. Le site de liaison habituellement observé pour les domaines PDZ ne semble pas s'appliquer à PDZ3. Il comporte habituellement une hélice  $\alpha$ , un feuillet  $\beta$  et une boucle. Pour PDZ2 il s'agit de la première boucle, du deuxième feuillet et de la première hélice  $\alpha$ . Comme la première hélice  $\alpha$  de PDZ3 ne serait pas présente, considérant que seulement la deuxième hélice d'un domaine PDZ normal est présente, elle ne peut pas être impliquée dans la liaison. Le peptide dérivé de PRK2 se lie donc à PDZ3. De plus, des perturbations sont observées

pour PDZ2 et le lieu. Les changements observés sont de plus grande taille que pour la première étape de la titration (lors de l'ajout du peptide dérivé de PRK2). Cette observation vient du fait que ce peptide possède un tryptophane (voir Tableau 1.1). Le tryptophane possède un cycle aromatique qui induit des courants de cycle changeant drastiquement l'environnement chimique des acides aminés impliqués dans la liaison.

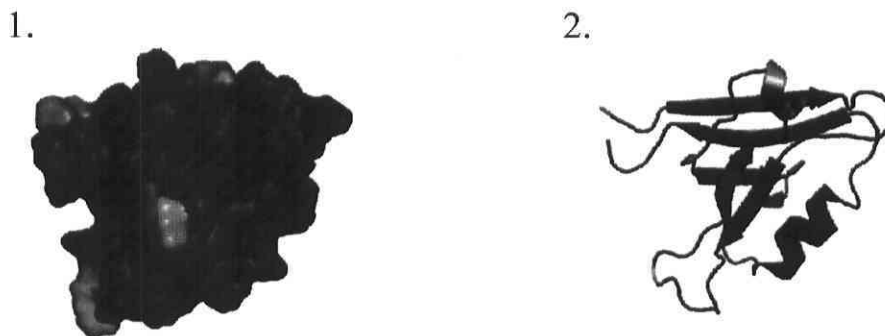
À la Figure 4.7, les perturbations induites par le peptide dérivé de PRK2 sont indiquées sur le graphique. Ces perturbations se positionnent encore dans le dos de la protéine, comme il est possible de le constater en observant ces mêmes perturbations sur la structure prédite de PDZ3 (Voir Figure 4.8). À la Figure 4.9, elles sont schématisées sur la structure de PDZ2.



**Figure 4.7** Perturbations observées lors de l'ajout du peptide dérivé de PRK2 sur le domaine tandem PDZ2/3.



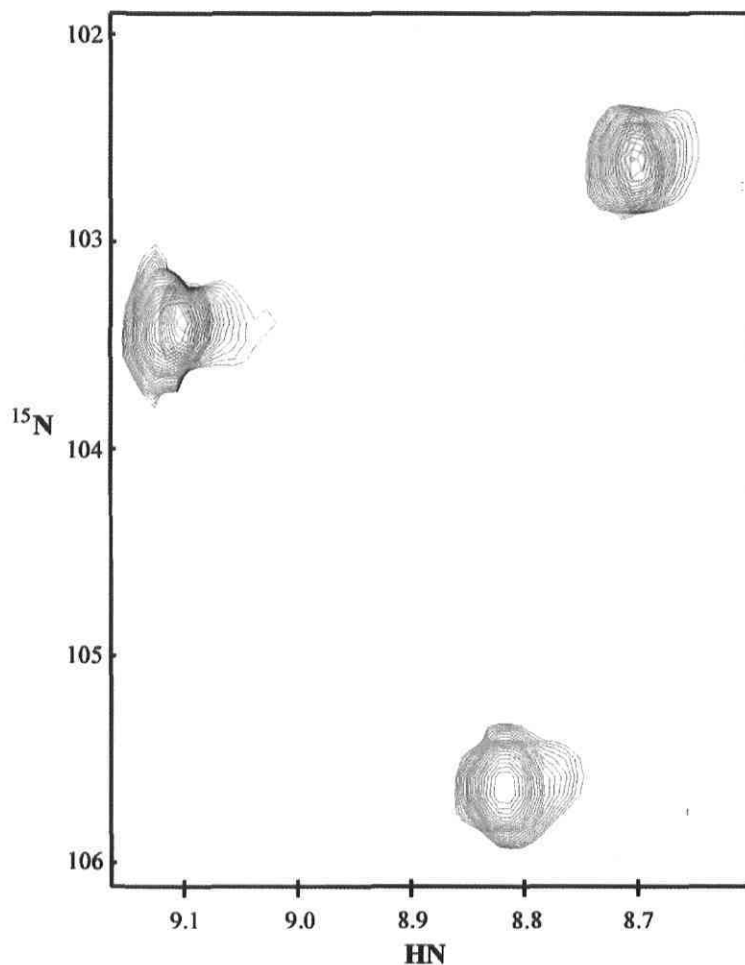
**Figure 4.8** Perturbations induites (en orange) par le peptide dérivé de PRK2 sur PDZ3. 1. Structure prédite de PDZ3 avec la surface de Van der Waals. 2. Structure prédite de PDZ3 en rubans. 3. Structure 1. basculée de 90°. 4. Structure 2. basculée de 90°.



**Figure 4.9** Perturbations induites par le peptide dérivé de PRK2 sur PDZ2. 1. Structure de PDZ2 avec la surface de Van der Waals. 2. Structure de PDZ2 en rubans.

#### 4.4 Prédiction de la structure quaternaire du domaine tandem PDZ2/3

Comme il a été mentionné dans la section sur les titrations, les domaines composant le domaine tandem PDZ2/3 semblent former une coopération afin de lier le peptide. Lors de l'ajout d'un peptide, les changements sont observés dans les deux domaines ainsi que dans le lieu. De plus, cette superstructure semble être formée dès le départ. En effet, aucun changement majeur n'est observé dans le spectre lors de l'ajout des peptides. Les perturbations observées le sont toujours à l'opposée du début des domaines. C'est aussi le cas pour les différences observées entre le domaine PDZ2 libre et le domaine tandem PDZ2/3. Il est donc possible que les deux domaines soient accolés et dos à dos. Les deux domaines réagissent lors du titrage de l'un avec l'autre. Les signaux observés disparaissent (voir Figure 4.10). Ces changements sont moins marquants que les titrations avec les peptides, cependant, il est impossible de nier que les deux domaines réagissent tout de même l'un avec l'autre. La liaison entre les deux domaines ne serait peut-être pas assez forte pour conserver la superstructure en solution. Le lieu pourrait aider les deux domaines à garder cette superstructure.



**Figure 4.10** Comparaison de quelques signaux de la titration du domaine PDZ2 marqué à  $^{15}\text{N}$  avec PDZ3 non-marqué. Le domaine PDZ2 libre en solution est indiqué en noir et la dernière étape de la titration est indiquée en mauve.

Le lieu se retrouve, en effet, seulement dans des protéines contenant des domaines PDZ. Après avoir fait une recherche sur BLAST (Basic Local Alignment Search Tool), disponible sur le site BMRB, des boucles similaires ont été retrouvées seulement dans des protéines ayant des domaines PDZ. Ces protéines sont : BA14\_bovin, PTN13\_H (orthologue humain de PTP-BL) et chez trois autres mutants de PTN13\_H. Cette boucle a donc peut-être une fonction biochimique liée à ces domaines PDZ. La prédiction de la structure secondaire du lieu a aussi été effectuée. Les lieux ne possèdent habituellement aucune structure. Dans le cas du lieu de PDZ2/3, un feuillet  $\beta$  semble être présent. Ce feuillet pourrait obliger le lieu à se replier et ainsi rapprocher les deux domaines ensemble.

À la Figure 4.11, la structure secondaire du lieur prédite peut être observée. Cependant, des études plus approfondies sur le lieur devront être effectuées.

PTSREERDPAGPQSPPPDQDAQRQAPEKVAKQTPHVKDYSFVTED  
BBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBHHBBBBBBBBBBFF'FF'BBB

**Figure 4.11** Structure secondaire prédite du lieur. Les lettres B, F et H sont utilisées pour boucle, feuillet et hélice, respectivement.

## Chapitre 5

### Conclusions

Le but des recherches effectuées était de déterminer si les domaines PDZ2 et PDZ3 s'associaient en milieu aqueux pour former une superstructure. Une des solutions afin d'atteindre ce but était de trouver la structure du domaine tandem PDZ2/3 de PTP-BL. Cependant, comme des spectres insatisfaisants à la détermination de structure ont été mesurés, d'autres solutions ont dû être trouvées. En effet, lorsque l'attribution a été terminée, des mesures simples ont pu être effectuées afin de savoir si les deux domaines s'associaient en solution et coopéraient afin de lier les peptides dérivés de leur ligand respectifs.

L'attribution du squelette du domaine tandem PDZ2/3 de PTP-BL a donc été, en premier lieu, effectuée à 88 %. Le domaine PDZ2 a été attribué à 83 % et le domaine PDZ3 à 93%. Cette attribution a été effectuée à partir de spectres RMN 2D et 3D comme il a été vu au Chapitre 3. Plusieurs difficultés ont été rencontrées lors de cette attribution. Ces difficultés étaient surtout reliées au fait que les deux domaines étaient très similaires. Elle aurait pu être plus élevée avec l'aide de spectres TROSY (Transverse Relaxation-Optimized Spectroscopy) ou d'échantillons plus concentrés qui augmenteraient ainsi l'intensité des signaux. En effet, l'utilisation de spectres TROSY diminue de beaucoup la

largeur de raies. Pour ce qui est de la concentration des échantillons, cette partie est plus difficile. Souvent l'agrégation des protéines est un problème avec des solutions de protéines trop concentrées. Cette agrégation augmente de beaucoup la largeur des raies. De plus, concentrer mécaniquement un échantillon comporte des risques. Chaque fois que la protéine est manipulée, elle se dégrade quelque peu. Il y a donc un choix à faire entre obtenir des spectres moins intenses et dégrader une protéine très chère. Un spectre 2D du domaine PDZ3 aurait été utile afin de bien différencier les signaux du lieur et de PDZ3.

Afin de déterminer si les deux domaines s'associaient ou non en solution, les déplacements chimiques de PDZ2 ont été comparés avec ceux du domaine tandem PDZ2/3. Des perturbations ont, en effet, été observées dans le domaine PDZ2. Ces perturbations, observées dans le dos de la protéine, ont démontré que l'ajout du lieur et du domaine PDZ3 changeait l'environnement chimique de PDZ2, donc que les deux domaines et le lieur étaient probablement rapprochés. Par la suite, le domaine PDZ2 marqué à  $^{15}\text{N}$  a été titré avec le domaine PDZ3 non-marqué. Des perturbations ont encore été observées. Cependant, celles-ci étaient moins marquées qu'avec l'ajout du lieur ainsi que du domaine PDZ3. Cela indiquerait que le lieur aurait probablement une utilité inconnue jusqu'à maintenant.

Afin de pouvoir mieux illustrer les perturbations observées lors des titrations ainsi que de prédire la structure du domaine tandem PDZ2/3 la bioinformatique a été utilisée. La structure secondaire et tertiaire du domaine PDZ3 a ainsi été prédite. Le niveau de confiance de ces prédictions est de 79% ce qui est un résultat satisfaisant. Cependant, cette structure devra être confirmée par des études RMN. La structure obtenue contient 5 feuillets  $\beta$  et une hélice  $\alpha$ . Habituellement les domaines PDZ contiennent 5 ou 6 feuillets  $\beta$  et deux hélices  $\alpha$ . Cette structure, inhabituelle aux domaines PDZ, pourrait être causée par le nombre de proline élevé. Les prolines sont réputées pour détruire la structure et 6 sont présentes dans la séquence de PDZ3. Le fait que PDZ3 ne contiendrait qu'une hélice  $\alpha$  expliquerait aussi le fait que le site de liaison soit différent.



Par la suite le domaine tandem a été titré avec les peptides dérivés des protéines APC et PRK2, APC étant connue pour se lier à PDZ2 et PRK2 à PDZ3. Ces titrations avaient pour but de vérifier si le domaine tandem liait vraiment ces protéines et de vérifier comment ils étaient liés. Le peptide dérivé d'APC se liait bien à PDZ2. De plus, certaines perturbations étaient aussi observées pour PDZ3 ainsi que pour le lieur. Cela démontrait que les deux domaines coopéraient afin de lier le peptide dérivé d'APC, que le lieur est bien impliqué dans la liaison et que les deux domaines formaient probablement une superstructure. Les changements observés dans les déplacements chimiques étaient pour la plupart observés dans le dos de la protéine appuyant ainsi l'hypothèse que les deux domaines seraient associés et probablement dos à dos. Le domaine tandem PDZ2/3 a ensuite été titré avec le peptide dérivant de PRK2. Le peptide se liait bien avec PDZ3 et des perturbations étaient aussi observées dans le lieur ainsi que dans PDZ2. Ces résultats appuyant d'autant plus les hypothèses ci-haut mentionnées.

Plusieurs travaux futurs seraient intéressants à la poursuite du projet. En premier lieu, la structure du domaine tandem avec et sans les peptides devrait être effectuée. Cela prouverait sans aucun doute le positionnement des domaines l'un par rapport à l'autre. Pour la détermination de la structure, plusieurs expériences RMN supplémentaires devraient être effectuées. Afin d'attribuer les chaînes latérales des acides aminés, des spectres HCCH-COSY et HCCH-TOCSY devraient être effectués. De plus, des HSQC-NOESY devraient aussi être mesurés afin de déterminer les contraintes structurales, c'est-à-dire les distances interprotons. Par la suite des RDCs (Residual dipolar coupling) devraient être mesurés. Finalement, il aurait été intéressant d'étudier la dynamique du domaine tandem par la mesure de NOE hétéronucléaires et de temps de relaxation  $T_1$  et  $T_2$ .

Plusieurs autres titrations pourraient être effectuées. Afin de mieux comprendre l'association des domaines avec leurs peptides respectifs, PDZ3 devrait être titré avec le peptide dérivé de PRK2 et PDZ2 devra être titré avec le peptide dérivé de APC. Ainsi, il

serait possible de calculer la différence de déplacement chimique observée et de pouvoir doser si le ligand se lie davantage avec le domaine tandem ou s'il se lie davantage avec le domaine seul. De plus, il serait possible de vérifier le site de liaison de PDZ3 qui ne semble pas suivre les motifs habituels de liaison pour les domaines PDZ. PDZ3 marqué à  $^{15}\text{N}$  devrait être titré avec PDZ2 afin de vérifier l'effet que PDZ2 a sur PDZ3. Afin de mieux comprendre le système de liaison en général, d'autres titrations pourraient être effectuées avec plusieurs autres peptides dérivés de d'autres protéines.

Finalement, le lieu pourrait être étudié davantage en profondeur. Plusieurs études sur le lieu peuvent être effectuées afin de mieux comprendre son rôle dans la liaison. En premier lieu, une expérience intéressante serait d'insérer un lieu comprenant seulement des prolines. Cela créerait un lieu rigide qui obligerait les deux domaines à être séparés. Il serait alors intéressant de voir ensuite la façon dont ces domaines mécaniquement séparés lient les peptides dérivés d'APC et de PRK2. Il serait aussi intéressant de séparer les deux domaines par un très court peptide afin de comparer aussi les capacités de liaison.

## Bibliographie

- Adamczak, R., Porollo, A. et Meller, J. (2005) Combining prediction of secondary structure and solvent accessibility in proteins. *Proteins: Struct., Funct. and Bioinf.* **59**: 467-475.
- Amersham Bioscience (2002). GST Gene fusion system Handbook, Amersham Bioscience, 110 p.
- Baldi, P., Brunak, S., Frasconi, P., Soda, G. et Pollastri, G. (1999) Exploiting the past and the future in protein secondary structure prediction. *Bioinformatics* **15**: 937-946.
- Banville, D., Ahmad, S., Stocco, R. et Shen, S. H. (1994) A novel protein-tyrosine phosphatase with homology to both the cytoskeletal proteins of the band 4.1 family and junction-associated guanylate kinases. *J. Biol. Chem.* **269**: 22320-22327.
- Bodenhausen, G. et Ruben, D. J. (1980) Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett.* **69**: 185-189.
- Bystrov, V. F. (1976) Spin-spin coupling and the conformational states of peptide systems. *Prog. Nucl. Magn. Reson.* **10**: 41-82.
- Cavanagh, J., Fairbrother, W. J., Palmer III, A. G. et Skelton, N. J. (1996). Protein NMR Spectroscopy, Principles and Practice. San Diego, Academic Press, 587 p.
- Cho, K.-O., Hunt, C. A. et Kennedy, M. B. (1992) The rat brain postsynaptic density fraction contains a homolog of the Drosophila discs-large tumor suppressor protein Neuron. *J. Neurosci.* **9**: 929-942.
- Clackson, T. et Wells, J. A. (1995) A hot spot of binding energy in a hormone-receptor interface. *Science* **267**: 383-386.
- Cuppen, E., Gerrits, H., Pepers, B., Wieringa, B. et Hendriks, W. (1998) PDZ motifs in PTP-BL and RIL bind to internal protein segments in the LIM domain protein RIL. *Mol. Biol. Cell* **9**: 671-683.
- Cuppen, E., Nagata, S., Wieringa, B. et Hendriks, W. (1997) No evidence for involvement of mouse protein-tyrosine phosphatase Bas-like Fas associated phosphatase-1 in Fas-mediated apoptosis. *J. Biol. Chem.* **272**: 30215-30220.
- Cuppen, E., van Ham, M., Wansink, D. G., de Leeuw, A., Wieringa, B. et Hendriks, W. (2000) The zyxin-related protein TRIP6 interacts with PDZ motifs in the adaptor protein RIL and the protein tyrosine phosphatase PTP-BL. *Eur. J. Cell Biol* **79**: 283-293.

- Cuppen, E., Van Ham, U., Pepers, B., Wieoringa, B. et Hendriks, W. (1999) Identification and molecular characterization of BP75, a novel bromodomain-containing protein. *FEBS Lett.* **459**: 291-298.
- Delaglio, F., Grzesiek, G., Vuister, G. W., Zhu, G., Pfeifer, J. et Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**: 277-293.
- DeLano, W. L. (1998). The pymol molecular graphics system. San Carlos, DeLano Scientific.
- Doyle, D. A., Lee, A., Lewis, J., Kim, E., Sheng, M. et Mackinnon, R. (1996) Crystal structures of a complexed and peptide-free membrane protein-binding domain: molecular basis of peptide recognition by PDZ. *Cell* **85**: 1067-1076.
- Erdmann, K., Kuhlmann, J., Lessmann, V., Herrmann, L., Eulenburg, V., Müller, O. et Heumann, R. (2000) The adenomatous polyposis coli-protein (APC) interacts with the protein tyrosine phosphatase PTP-BL via an alternatively spliced PDZ domain. *Oncogene* **19**: 3894-3901.
- Flynn, P., Mellor, H., Casamassima, A. et Parker, P. J. (2000) Rho GTPase Control of Protein Kinase C-related Protein Kinase Activation by 3-Phosphoinositide-dependent Protein Kinase. *J. Biol. Chem.* **275**: 11064-11070.
- Gross, C., Heumann, R. et Erdmann, K. S. (2001) The protein kinase C-related kinase interacts with the protein tyrosine phosphatase PTP-BL via a novel PDZ binding motif. *FEBS Lett.* **496**: 101-104.
- Grzesiek, S. et Bax, A. (1992) Improved 3D triple-resonance NMR techniques applied to a 31-kDa Protein. *J. Magn. Reson.* **96**: 432-440.
- Grzesiek, S. et Bax, A. (1993) The importance of not saturating H<sub>2</sub>O in protein NMR. Application to sensitivity enhancement and NOE measurements. *J. Biomol. NMR* **3**: 185-204.
- Harding, S. E. et Chowdhry, B. Z. (2001). Protein-Ligand Interactions: Structure and Spectroscopy. New York, Oxford university press, 436 p.
- Hatada, M. H., Lu, X., Laird, E. R., Green, J., Morgenstern, J. P. et Lou, M. (1995) Molecular basis for interaction of the protein tyrosine kinase ZAP-70 with the T-cell receptor. *Nature* **377**: 32-38.
- He, J. F., Peng, G. W., Min, J., Yu, D.-W., Liang, W.-J., Zhang, S.-Y., Xu, R.-H., Zheng, H.-Y., Wu, X.-W., Xu, J., Wang, Z.-H., Fang, L., Zhang, X., Li, H., Yan, X.-G., Lu, J. H., Hu, Z.-H., Huang, J. C., Wan, Z.-Y., Hou, J.-L., Lin, J.-Y., Song, H.-D., Wang, S.-Y., Zhou, X.-J., Zhang, G.-W., Gu, B.-W., Zheng, H. J., Zhang, X.-L., He, M., Zheng, K.,

Wang, B., Fu, G., Wang, X.-N., Chen, S.-J., Chen, Z., Hao, P., Tang, H., Ren, S.-X., Zhong, Y., Guo, Z.-M., Liu, Q., Miao, Y.-G., Kong, X.-Y., He, W.-Z., Li, Y.-X., Wu, C.-I., Zhao, G.-P., Rossa, W. K., Stephen, S. C., Tong, Y.-K., Paul, K. S., Tam, J. S. et Dennis, L. Y. M. (2004) Molecular evolution of the SARS coronavirus during the course of the SARS epidemic in china. *Science* **303**: 1666-1669.

Herrmann, L., Dittmar, T. et Erdmann, S. K. (2003) The protein tyrosine phosphatase PTP-BL associates with the midbody and is involved in the regulation of cytokinesis. *Molecular biology of the cell* **14**: 230-240.

Higgins, D., Thompson, J., Gibson, T., Thompson, J. D., Higgins, D. G. et Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**: 4673-4680.

Hitchens, T. K., Lukin, J. A., Zhan, Y., McCallum, S. A. et Rule, G. S. (2003) MONTE: An automated Monte Carlo based approach to nuclear magnetic resonance assignment of proteins. *J. Biomol. NMR* **25**: 1-9.

Ikura, M., Kay, L. E. et Bax, A. (1990) A novel approach for sequential assignment of  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin. *Biochemistry* **29**: 4659-4667.

Jones, D. T. (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* **292**: 195-202.

Jung, Y. S. et Zweckstetter, M. (2004) MARS - robust automatic backbone assignment of proteins. *J. Biomol. NMR* **30**: 11-23.

Kay, L. E., Xu, G. Y. et Yamazaki, T. (1994) Enhanced sensitivity triple-resonance spectroscopy with minimal  $\text{H}_2\text{O}$  saturation. *J. Magn. Reson.* **A109**: 129-133.

Lambert, C., Léonard, N., De Bolle, X. et Depiereux, E. (2002) ESyPred3D: Prediction of proteins 3D structures. *Bioinformatics* **18**: 1250-1256.

Long, J.-F., Tochio, H., Wang, P., Fan, J.-S., Sala, C., Niethammer, M., Sheng, M. et Zhang, M. (2003) Supramodular structure and synergistic target binding of the N-terminal tandem PDZ domains of PSD-95. *J. Mol. Biol.* **327**: 203-214.

Maekawa, K., Imagawa, N., Nagamatsu, M. et Harada, S. (1994) Molecular cloning of a novel protein-tyrosine phosphatase containing a membrane-binding domain and GLGF repeats. *FEBS Lett.* **337**: 179-184.

Maekawa, K., Imagawa, N., Naito, A., Harada, S., Yoshie, O. et Takagi, S. (1999) Association of protein-tyrosine phosphatase PTP-BAS with the transcription-factor-

inhibitory protein I $\kappa$ B $\alpha$  through interaction between the PDZ1 domain and ankyrin repeats. *Biochem. J.* **337**(2): 179-184.

Maudsley, A. A. et Ernst, R. R. (1977) Indirect detection of magnetic resonance by heteronuclear two-dimensional spectroscopy. *Chem. Phys. Lett.* **50**: 368-372.

McGuffin, L. J., Bryson, K. et Jones, D. T. (2000). The PSIPRED protein structure prediction server, UCL. **2004**.

McIntosh, L. P. et Dahlquist, F. W. (1990) Biosynthetic incorporation of  $^{15}\text{N}$  and  $^{13}\text{C}$  for assignment and interpretation of nuclear magnetic resonance spectra of proteins. *Q. Rev. Biophys.* **23**: 1-38.

Montelione, G. T., Emerson, S. D. et Lyons, B. A. (1992) A general approach for determining scalar coupling constants in polypeptides and proteins. *Biopolymers* **32**: 327-334.

Morais, C. J., Petosa, C., Sutcliffe, M. J., Raza, S., Byron, O., Poy, F., Martafia, S. M., Chisti, A. H. et Liddington, R. C. (1996) Crystal structure of a PDZ domain. *Nature* **382**: 649-652.

Morris, G. A. (1980) Indirect measurement of proton relaxation rates by 'INEPT' polarization transfer to carbon-13: proton spin-lattice relaxation in cholesteryl acetate solutions. *J. Magn. Reson.* **41**: 185-188.

Morris, G. A. (1980) Sensitivity enhancement in  $^{15}\text{N}$  NMR: polarization transfer using the INEPT sequence. *J. Am. Chem. Soc.* **102**: 428-429.

Morris, G. A. et Freeman, R. (1979) Enhancement of nuclear magnetic resonance signals by polarization transfer. *J. Am. Chem. Soc.* **101**: 760-761.

Muchmore, D. C., McIntosh, L. P., Russel, B. C., Anderson, D. E. et Dahlquist, W. F. (1989) Expression and nitrogen-15 labelling of proteins for proton and nitrogen-15 nuclear magnetic resonance. *Meth. Enzymol.* **177**: 44-73.

Muhandiram, D. R. et Kay, L. E. (1994) Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity. *J. Magn. Reson. B* **103**: 203-216.

Murthy, K. K., Clark, K., Fortin, Y., Shen, S. H. et Banville, D. (1999) ZRP-1, a zyxin-related protein, interacts with the Second PDZ Domain of the Cytosolic Protein Tyrosine Phosphatase hPTP1E. *J. Biol. Chem.* **274**: 20679-20687.

Palmer, A., Zimmer, M., Erdmann, K. S., Deutsch, U. et Klein, R. (2002) EphrinB phosphorylation and reverse signaling: regulation by Src kinases and PTP-BL phosphatase. *Mol. Cell* **9**: 725-737.

- Pardi, A., Billeter, M. et Wüthrich, K. (1984) Calibration of the angular dependence of the amide proton-C-alpha proton coupling constants,  $^3\text{JHN}\alpha$ , in a globular protein. *J. Mol. Biol.* **180**: 741-751.
- Petsko, G. A. et Ringe, D. (2004). Protein Structure and Function. London, New Science Press Ltd., 195 p.
- Piotto, M., Saudek, V. et Sklenar, V. (1992) Gradient-tailored excitation for single-quantum NMR-spectroscopy of aqueous solutions. *J. Biomol. NMR* **2**: 661 - 666.
- Saras, J., Claesson-Welsh, L., Heldin, C. H. et Gonez, L. J. (1994) Cloning and characterization of PTPL1, a protein tyrosine phosphatase with similarities to cytoskeletal-associated proteins. *J. Biol. Chem.* **269**: 24082-24089.
- Saras, J., Franzen, P., Aspenstrom, P., Hellman, U., Gonez, L. J. et Heldin, C. H. (1997) A novel GTPase-activating protein for Rho interacts with a PDZ domain of the protein-tyrosine phosphatase PTPL1. *J. Biol. Chem.* **272**: 24333-24338.
- Sato, T., Irie, S., Kitada, S. et Reed, J. C. (1995) FAP-1: a protein tyrosine phosphatase that associates with Fas. *Science* **268**: 411-415.
- Schleucher, J., Sattler, M. et Griesinger, C. (1993) Coherence selection via gradients without signal attenuation: application to 3D HNCOC. *Angew. Chem. Int. Ed.* **32**: 1489-1491.
- Sklenar, V., Piotto, M., Leppik, R. et Saudek, V. (1993) Gradient-tailored water suppression for H-1-N-15 HSQC experiments optimized to retain full sensitivity. *J. Magn. Reson.* **102**(Series A): 241-245.
- Songyang, Z., Fanning, A. S., Fu, C., Xu, J., Marfatia, S. M., Chishti, A. H., A., C., Chan, A. C., Anderson, J. M. et Cantley, L. C. (1997) Recognition of unique carboxyl-terminal motifs by distinct PDZ domains. *Science* **275**: 73-77.
- Tochio, H., Zhang, Q., Mandal, P., Li, M. et Zhang, M. (1999) Solution structure of the extended neuronal nitric oxide synthase PDZ domain complexed with an associated peptide. *Nature Struct. Biol.* **6**: 417-421.
- van Ham, M., Croes, H., Schepens, J., Fransen, J., Wieringa, B. et Hendriks, W. (2003) Cloning and characterization of mCRIP2, a mouse LIM only that interact with PDZ domain IV of PTP-BL. *Genomics* **8**: 631-644.
- Vincent, S. et Settleman, J. (1997) The PRK2 kinase is a potential effector target of both RHO and Rac GTPases and regulates actin cytoskeletal organization. *Mol. Cell. Biol.* **17**: 2247-2256.

Vuister, G. W., Grzesiek, G., Delaglio, F., Wang, A. C., Tschudin, R., Zhu, G. et Bax, A. (1994) Measurement of homo- and heteronuclear J couplings from quantitative J correlation. *Meth. Enzymol.* **239**: 79-105.

Wagner, G., Braum, W., Havel, T. H., Schaumann, T., Go, N. et Wüthrich, K. (1987) Protein structures in solution by nuclear magnetic resonance and distance geometry: the polypeptide fold of the basic pancreatic trypsin inhibitor determined using two different algorithms. *J. Mol. Biol.* **196**: 611-639.

Walma, T., Aelen, J., Nabuurs, S. B., M., O., van den Berk, L., Hendriks, W. et Vuister, G. W. (2004) A closed binding pocket and global destabilization modify the binding properties of an alternatively spliced form of the second PDZ domain of PTP-BL. *Structure* **12**: 11-20.

Walma, T., Spronk, C. A. E., Tessari, M., Aelen, J., Schepens, J., Hendriks, W. et Vuister, G. W. (2002) Structure, dynamics and binding characteristics of the second PDZ domain of PTP-BL. *J. Mol. Biol.* **316**: 1101-1110.

Willott, E., Balda, M. S., Fanning, A. S., Jameson, B., Van Itallie, C. et Anderson, J. M. (1993) The tight junction protein ZO-1 is homologous to the Drosophila discs-large tumor suppresser protein of septate junctions. *Proc. Natl. Acad. Sci. U. S. A.* **90**: 7834-7838.

Woods, D. F. et Bryant, P. J. (1991) The discslarge tumor suppressor gene of Drosophila encodes a guanylate kinase homolog localized at septate junctions. *Cell* **66**: 451-464.



## **Annexe**

### **Déplacements chimiques observé**

3.CA	63.734	18.CA	54.863	32.HN	8.360
3.CB	29.561	18.CB	31.141	32.CA	44.574
4.N	109.594	18.C	174.663	33.C	174.721
4.HN	8.493	19.N	126.151	34.N	117.751
4.CA	46.606	19.HN	8.868	34.HN	8.864
5.CA	57.260	19.CA	53.179	34.CA	58.794
5.C	176.667	19.CB	45.227	35.N	117.749
6.N	121.357	19.C	173.897	35.HN	8.864
6.HN	8.571	20.N	126.323	35.CA	58.944
6.CA	58.865	20.HN	8.469	36.N	100.790
6.C	175.885	20.CA	49.861	36.HN	6.841
7.N	120.993	20.CB	20.023	36.CA	65.789
7.HN	8.333	20.C	177.376	37.CA	62.054
7.CA	56.504	21.N	120.634	37.CB	32.885
9.CA	62.787	21.HN	8.471	37.C	177.642
9.CB	32.492	21.CA	57.734	38.N	118.532
9.C	175.863	21.CB	34.703	38.HN	7.564
10.N	123.336	21.C	177.191	38.CA	58.395
10.HN	8.755	22.N	113.848	38.CB	29.559
10.CA	53.690	22.HN	8.071	39.N	124.791
11.CA	64.571	22.CA	61.274	39.HN	8.492
11.CB	32.370	23.CA	55.359	39.CA	54.128
11.C	177.629	23.CB	40.277	40.CA	59.718
12.N	113.250	23.C	176.487	40.CB	-2.119
12.HN	8.907	24.N	102.502	40.C	175.232
12.CA	45.070	24.HN	8.727	41.N	109.972
12.C	174.079	24.CA	45.825	41.HN	8.471
13.N	121.812	25.N	113.977	41.CA	43.900
13.HN	8.519	25.HN	8.017	42.N	104.974
13.CA	54.540	25.CA	61.271	42.HN	7.764
13.CB	41.956	25.CB	70.183	42.CA	44.493
13.C	176.640	25.C	174.139	43.N	119.280
14.N	113.129	26.N	122.004	43.HN	8.453
14.HN	8.423	26.HN	8.197	43.CA	56.969
14.CA	60.691	26.CA	57.595	43.CB	36.333
14.CB	71.238	27.C	176.130	43.C	174.590
14.C	174.881	28.N	115.163	44.N	123.206
15.N	119.950	28.HN	7.595	44.HN	8.954
15.HN	9.308	28.CA	60.213	44.CA	55.919
15.CA	56.145	28.C	174.029	44.CB	42.635
15.C	171.313	29.N	119.179	44.C	175.984
16.N	118.967	29.HN	8.486	45.N	120.449
16.HN	8.547	29.CA	56.065	45.HN	9.268
16.CA	54.038	30.CA	58.710	45.CA	64.452
16.C	176.027	30.CB	35.622	45.CB	31.711
17.N	119.797	30.C	174.541	45.C	175.822
17.HN	8.577	31.N	115.751	46.N	134.698
17.CA	61.117	31.HN	9.266	46.HN	9.565
17.CB	36.945	31.CA	59.562	46.CA	56.960
17.C	173.120	31.CB	70.559	46.CB	34.628
18.N	125.940	31.C	171.968	46.C	174.708
18.HN	7.828	32.N	111.883	47.N	117.289

47.HN	7.508	59.HN	7.946	69.C	175.469
47.CA	51.465	59.CA	58.044	70.N	128.778
47.CB	22.546	59.CB	31.051	70.HN	10.525
47.C	177.142	59.C	175.852	70.CA	54.134
48.N	120.607	60.N	119.899	70.CB	37.073
48.HN	8.582	60.HN	8.372	70.C	175.183
48.CA	60.579	60.CA	61.615	71.N	103.137
48.C	175.986	60.CB	39.459	71.HN	9.124
49.N	129.674	60.C	173.639	71.CA	45.128
49.HN	7.783	61.N	123.244	71.C	173.911
49.CA	58.715	61.HN	9.042	72.N	123.422
50.CA	64.301	61.CA	54.783	72.HN	8.044
50.CB	32.452	61.CB	31.824	72.CA	62.192
50.C	177.270	61.C	174.158	72.C	175.657
51.N	117.667	62.N	120.820	73.N	120.846
51.HN	9.361	62.HN	8.670	73.HN	8.673
51.CA	57.703	62.CA	58.493	73.CA	56.690
51.CB	29.544	62.CB	32.452	73.C	177.535
51.C	176.316	62.C	177.047	74.N	120.349
52.N	106.133	63.N	114.626	74.HN	8.288
52.HN	7.824	63.HN	9.113	74.CA	51.491
52.CA	44.704	63.CA	44.796	75.CA	55.896
52.C	174.437	63.C	174.377	75.CB	31.659
53.N	121.069	64.N	121.510	75.C	176.053
53.HN	8.846	64.HN	8.021	76.N	117.528
53.CA	55.231	64.CA	55.927	76.HN	9.682
53.CB	18.899	64.CB	41.084	76.CA	45.379
53.C	178.241	64.C	174.718	76.C	173.549
54.N	118.988	65.N	123.225	77.N	122.052
54.HN	7.786	65.HN	9.048	77.HN	7.879
54.CA	54.963	65.CA	54.701	77.CA	52.570
54.CB	18.612	65.CB	31.704	77.CB	20.100
54.C	180.946	65.C	175.587	77.C	179.929
55.N	124.566	66.N	126.082	78.N	111.644
55.HN	9.300	66.HN	8.943	78.HN	8.244
55.CA	59.989	66.CA	62.306	78.CA	60.874
55.C	174.152	66.CB	31.444	78.CB	73.309
56.N	113.421	66.C	175.352	78.C	174.856
56.HN	7.989	67.N	127.221	79.N	121.752
56.CA	61.367	67.HN	9.187	79.HN	9.164
56.CB	63.288	67.CA	56.081	79.CA	62.019
56.C	175.674	67.CB	31.730	79.CB	30.853
57.N	120.108	67.C	180.915	79.C	176.773
57.HN	7.704	68.N	118.934	80.N	116.694
57.CA	57.242	68.HN	7.817	80.HN	8.501
57.CB	44.591	68.CA	52.222	80.CA	59.704
57.C	177.741	68.CB	22.930	80.CB	33.301
58.N	102.665	68.C	175.424	80.C	179.097
58.HN	7.449	69.N	120.346	81.N	117.552
58.CA	46.389	69.HN	8.281	81.HN	7.682
58.C	174.893	69.CA	60.476	81.CA	58.334
59.N	117.818	69.CB	34.437	81.CB	30.047

81.C	179.452	94.HN	8.744	107.C	172.812
82.N	123.890	94.CA	56.243	108.N	131.461
82.HN	8.520	94.CB	40.765	108.HN	8.726
82.CA	55.518	94.C	174.390	108.CA	50.05
82.CB	19.701	95.N	126.954	108.C	176.298
82.C	178.705	95.HN	8.982	109.N	121.802
83.N	117.467	95.CA	53.491	109.HN	8.402
83.HN	8.294	95.CB	45.848	109.CA	56.835
83.CA	67.135	95.C	173.490	110.CA	63.566
83.CB	26.936	96.N	123.470	110.CB	32.461
83.C	179.171	96.HN	7.960	110.C	176.661
84.N	120.373	96.CA	49.123	111.N	123.210
84.HN	8.114	96.C	175.106	111.HN	8.399
84.CA	59.423	97.N	127.186	111.CA	52.284
84.CB	29.420	97.HN	8.883	111.CB	20.299
84.C	178.536	97.CA	53.947	111.C	178.297
85.N	117.440	97.CB	46.311	112.N	107.841
85.HN	8.226	97.C	174.043	112.HN	7.987
85.CA	67.158	98.N	118.985	112.CA	44.513
85.CB	38.148	98.HN	8.933	113.CA	60.727
85.C	173.650	98.CA	54.047	113.CB	33.925
86.N	119.800	98.CB	34.717	113.C	175.980
86.HN	7.914	98.C	176.419	114.N	129.226
86.CA	58.299	99.N	128.451	114.HN	10.111
87.N	121.148	99.HN	9.173	114.CA	54.152
87.HN	8.223	99.CA	56.039	114.CB	28.871
87.CA	58.785	100.CA	45.235	114.C	175.414
87.C	180.970	100.C	172.902	115.N	116.465
88.N	119.684	101.N	117.390	115.HN	7.869
88.HN	7.785	101.HN	8.161	115.CA	59.888
88.CA	52.663	101.CA	55.233	118.CA	63.208
88.CB	37.108	101.CB	30.737	118.CB	33.019
88.C	174.942	101.C	176.583	119.N	119.656
89.N	110.010	102.N	123.510	119.HN	8.332
89.HN	7.647	102.HN	8.297	119.CA	54.255
89.CA	59.831	102.CA	60.258	119.CB	35.463
90.N	109.644	103.CA	63.284	119.C	176.590
90.HN	8.471	103.CB	32.443	120.N	120.349
90.CA	43.996	103.C	177.139	120.HN	8.288
91.CA	58.412	104.N	114.547	120.CA	56.450
91.C	175.194	104.HN	8.251	120.CB	29.812
92.N	115.939	104.CA	61.600	120.C	176.094
92.HN	7.733	104.CB	104.786	121.N	121.075
92.CA	61.323	104.C	173.687	121.HN	8.324
92.CB	53.308	105.N	115.881	121.CA	54.804
92.C	175.822	105.HN	8.254	121.CB	41.541
93.N	128.556	106.CA	53.474	121.C	176.466
93.HN	9.020	106.C	174.486	122.N	123.965
93.CA	61.201	107.N	117.126	122.HN	8.130
93.CB	35.240	107.HN	9.114	122.CA	53.170
93.C	173.219	107.CA	56.220	122.CB	19.535
94.N	126.641	107.CB	31.141	122.C	178.198

123.N	117.969	138.CA	53.771	150.CB	33.558
123.HN	8.225	138.CB	43.658	150.C	176.323
123.CA	56.073	138.C	175.518	151.N	124.572
124.CB	42.632	139.N	116.443	151.HN	9.025
124.C	175.253	139.HN	8.678	151.CA	61.430
125.N	111.353	139.CA	53.759	151.CB	36.608
125.HN	7.168	139.CB	45.437	151.C	173.342
125.CA	55.410	140.N	118.701	152.N	122.960
125.CB	29.738	140.HN	8.020	152.HN	8.504
125.C	175.143	140.CA	58.621	152.CA	55.015
126.N	126.780	140.CB	64.147	152.CB	33.391
126.HN	8.337	140.C	173.682	152.C	175.486
126.CA	50.591	141.N	123.320	153.N	123.227
127.CA	64.478	141.HN	8.165	153.HN	8.838
127.C	174.597	141.CA	60.502	153.CA	52.673
128.N	131.222	141.C	174.608	153.C	174.671
128.HN	9.165	142.N	112.269	154.N	121.202
128.CA	57.096	142.HN	6.254	154.HN	8.615
128.CB	29.357	142.CA	55.413	154.C	176.298
128.C	174.308	142.CB	38.617	155.N	121.746
129.N	117.047	142.C	175.218	155.HN	8.384
129.HN	7.129	143.N	126.010	155.CA	56.738
129.CA	55.993	143.HN	7.109	155.CB	32.304
129.C	176.418	143.CA	63.587	156.N	116.090
130.N	121.399	144.N	124.620	156.HN	7.960
130.HN	8.008	144.HN	8.373	156.CA	53.116
130.CA	62.233	144.CA	62.101	156.C	174.643
130.CB	33.239	145.N	125.332	157.N	118.683
130.C	175.344	145.HN	8.245	157.HN	8.345
131.N	127.971	145.C	174.192	157.CA	60.709
131.HN	8.360	146.N	113.977	157.CB	62.130
131.CA	52.341	146.HN	8.017	157.C	177.781
131.C	177.490	146.CA	54.715	158.N	116.116
132.N	117.027	146.CB	42.927	158.HN	8.615
132.HN	8.279	146.C	177.029	158.CA	59.105
132.CA	56.168	147.N	117.680	158.CB	64.147
132.CB	29.923	147.HN	7.602	158.C	175.330
133.N	118.591	147.CA	51.709	159.N	110.436
133.HN	8.277	147.CB	39.140	159.HN	8.033
133.CA	56.477	147.C	174.860	159.CA	45.059
134.N	124.123	148.N	111.220	160.CA	55.203
134.HN	8.589	148.HN	7.314	160.CB	44.242
134.CA	61.830	148.CA	60.666	160.C	177.785
136.N	123.570	148.C	174.675	161.N	104.841
136.HN	8.403	149.N	120.817	161.HN	8.751
136.CA	55.081	149.HN	8.775	161.CA	46.203
137.N	118.653	149.CA	56.063	161.C	175.116
137.HN	7.593	149.CB	39.486	162.N	116.703
137.CA	60.142	149.C	171.603	162.HN	7.421
137.C	175.962	150.N	121.822	162.CA	55.633
138.N	125.332	150.HN	8.718	162.CB	41.771
138.HN	8.245	150.CA	54.299	162.C	173.141

163.N	115.751	178.HN	8.255	191.C	180.445
163.HN	8.740	178.CA	58.945	192.N	126.628
163.CA	55.628	178.C	173.839	192.HN	10.236
163.CB	33.261	179.N	121.572	192.CA	55.221
163.C	177.222	179.HN	8.617	192.CB	20.299
164.N	120.231	179.CA	60.793	192.C	180.223
164.HN	8.537	179.CB	42.058	193.N	116.848
164.CA	55.456	179.C	172.815	193.HN	8.375
164.CB	43.112	180.N	130.104	193.CA	58.394
164.C	173.149	180.HN	9.555	193.CB	30.272
165.N	116.253	180.CA	61.087	193.C	175.500
165.HN	9.472	180.C	176.716	194.N	112.965
165.CA	56.365	181.N	119.521	194.HN	7.863
165.C	178.105	181.HN	7.923	194.CA	61.174
166.N	114.617	181.CA	58.311	194.CB	64.643
166.HN	8.403	181.C	173.515	194.C	175.877
166.CA	56.247	182.N	123.136	195.N	109.041
166.CB	29.451	182.HN	8.008	195.HN	7.314
166.C	176.232	182.CA	60.125	195.CA	46.464
167.N	121.215	183.N	125.162	195.C	174.776
167.HN	8.078	183.HN	8.594	196.N	117.762
167.CA	56.209	183.CA	59.095	196.HN	7.761
168.CA	54.442	183.CB	28.342	196.CA	56.100
168.CB	35.268	183.C	178.186	196.C	174.817
169.N	123.164	184.N	121.058	197.N	117.169
169.HN	8.869	184.HN	7.817	197.HN	7.731
170.CA	57.895	184.CA	57.009	197.CA	60.172
171.N	121.148	184.CB	64.018	197.C	173.149
171.HN	8.223	184.C	175.980	198.N	128.693
172.CA	65.063	185.N	127.453	198.HN	9.084
172.CB	33.748	185.HN	9.288	198.CA	51.639
172.C	175.860	185.CA	53.251	198.CB	38.704
173.N	115.018	185.CB	44.866	199.N	119.599
173.HN	7.203	185.C	176.073	199.HN	8.455
173.CA	55.101	186.N	122.158	199.CA	65.376
173.CB	31.461	186.HN	7.528	199.CB	63.799
173.C	175.154	186.CA	53.461	199.C	176.101
174.N	125.291	187.CA	64.069	200.N	117.217
174.HN	9.016	187.CB	31.339	200.HN	9.430
174.CA	55.165	187.C	177.949	200.CA	44.534
174.CB	30.171	188.N	112.095	200.C	173.932
175.N	122.457	188.HN	8.836	201.N	121.946
175.HN	8.257	188.CA	45.170	201.HN	7.802
175.CA	58.990	189.N	120.687	201.CA	56.051
176.CA	53.202	189.HN	8.041	201.C	175.154
176.CB	39.123	189.CA	56.200	202.N	121.707
176.C	175.034	190.CA	62.481	202.HN	9.034
177.N	107.901	190.C	174.486	202.CA	60.481
177.HN	8.277	191.N	125.909	202.C	175.854
177.CA	45.374	191.HN	9.159	203.N	125.789
177.C	175.456	191.CA	56.211	203.HN	8.794
178.N	115.886	191.CB	19.390	203.CA	61.277

203.CB	18.044	217.HN	8.322	231.CA	55.703
203.C	176.470	217.CA	52.450	231.C	177.201
204.N	125.931	217.CB	19.948	232.N	124.253
204.HN	9.136	217.C	176.679	232.HN	8.632
204.CA	56.283	218.N	124.715	232.CA	53.274
204.C	179.392	218.HN	7.955	232.CB	44.522
205.N	118.102	218.CA	57.169	232.C	174.550
205.HN	8.109	218.C	179.253	233.N	128.141
205.CA	55.088	219.N	124.257	233.HN	8.986
205.C	176.224	219.HN	8.457	233.CA	53.405
206.N	126.006	219.CA	67.463	233.C	174.531
206.HN	9.013	219.C	177.050	234.N	124.546
206.CA	61.025	220.N	119.139	234.HN	8.416
207.CA	53.926	220.HN	8.021	234.CA	54.260
207.C	175.655	220.CA	65.976	234.CB	41.393
208.N	103.030	220.C	179.082	234.C	173.961
208.HN	9.039	221.N	113.351	235.N	124.341
208.CA	45.234	221.HN	8.106	235.HN	8.039
208.C	174.413	221.CA	61.951	235.CA	52.428
209.N	126.274	221.C	195.833	235.C	177.465
209.HN	8.016	222.N	127.245	236.N	120.867
209.CA	50.671	222.HN	8.000	236.HN	8.294
210.CA	61.719	222.CA	54.736	236.CA	57.838
210.C	176.925	222.CB	18.263	237.C	175.207
211.N	123.894	222.C	179.625	238.N	124.620
211.HN	7.918	223.N	115.825	238.HN	8.373
211.CA	54.797	223.HN	7.802	238.CA	50.819
212.N	119.515	223.CA	56.190	239.CA	63.656
212.HN	9.427	223.C	177.315	239.C	175.319
212.CA	57.494	224.N	118.832	240.N	109.644
212.CB	30.995	224.HN	7.761	240.HN	8.471
212.C	177.267	224.CA	57.139	240.CA	45.086
213.N	105.826	224.C	177.465	240.C	174.698
213.HN	7.681	225.N	109.644	241.N	121.355
213.CA	45.284	225.HN	8.471	241.HN	7.540
213.C	174.588	225.CA	45.181	241.CA	65.522
214.N	120.952	226.CA	57.470	241.CB	29.490
214.HN	7.608	226.C	172.694	241.C	176.134
214.CA	54.924	227.N	126.274	242.N	117.948
214.CB	42.659	227.HN	8.014	242.HN	8.336
214.C	177.561	227.CA	48.765	242.CA	50.881
215.N	119.515	228.CA	61.625	243.CA	62.211
215.HN	9.427	228.C	174.233	244.N	129.723
215.CA	57.538	229.N	114.880	244.HN	8.022
215.CB	45.679	229.HN	8.307	244.CA	58.140
215.C	176.931	229.CA	55.913	245.CA	62.790
216.N	119.670	230.N	121.148	245.CB	33.019
216.HN	8.040	230.HN	8.223	245.C	176.593
216.CA	55.765	230.CA	62.261	246.N	122.892
216.CB	31.483	230.C	175.633	246.HN	8.470
216.C	176.492	231.N	124.461	246.CA	53.783
217.N	123.227	231.HN	8.326	248.CA	63.984

248.CB 32.906  
248.C 178.050  
249.N 110.778  
249.HN 9.336  
249.CA 45.545  
249.C 173.492  
250.N 118.573  
250.HN 7.224  
250.CA 54.086