



**HAL**  
open science

# Paysage conformationnel de petits modèles de peptides en phase gazeuse : Spectroscopie UV/ IR et approche théorique

Wutharath Chin

► **To cite this version:**

Wutharath Chin. Paysage conformationnel de petits modèles de peptides en phase gazeuse : Spectroscopie UV/ IR et approche théorique. Autre. Université Paris Sud - Paris XI, 2005. Français. NNT : . tel-00158069

**HAL Id: tel-00158069**

**<https://theses.hal.science/tel-00158069>**

Submitted on 27 Jun 2007

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ORSAY

N° D'ORDRE : 8005

**UNIVERSITÉ PARIS 11**  
**UFR SCIENTIFIQUE D'ORSAY**

**THÈSE**

Spécialité : **Chimie physique**

présentée

pour obtenir le grade de

**DOCTEUR EN SCIENCES**

**de l'UNIVERSITÉ PARIS 11, ORSAY**

par **Wutharath CHIN**

Sujet de la thèse :

**Paysage conformationnel de petits modèles de peptides en phase gazeuse :**

**Spectroscopie UV/ IR et approche théorique**

Soutenue le 3 octobre 2005 devant le jury composé de :

**M. Jacques Delaire, Président**

**M. Charles Desfrancois, Rapporteur**

**M. Gilles Ohanessian, Rapporteur**

**M. Christophe Juvet**

**M. Michel Mons**



# Remerciements

Ce travail de thèse a été effectué dans le service Photons, Atomes et Molécules (SPAM) du Commissariat à l'Énergie Atomique de Saclay. Je souhaite remercier Didier Normand de m'avoir accueillie au sein du service.

Je remercie Messieurs Charles Desfrancois et Gilles Ohanessian d'avoir accepté d'être les rapporteurs de cette thèse et pour l'intérêt qu'ils ont porté à mon travail. J'adresse également mes remerciements à Messieurs Jacques Delaire et Christophe Juvet qui ont eu l'amabilité de juger ce travail et de participer au jury.

Je tiens à exprimer toute ma gratitude à Michel Mons qui a dirigé ma thèse dans la continuité de mon stage de DEA et de maîtrise. Je lui suis très reconnaissante d'avoir été attentif à mes travaux. Sa rigueur scientifique, sa pédagogie, sa disponibilité et son humour guideront mes pas dans mon travail de recherche.

Je remercie chaleureusement Madame Iliana Dimicoli de m'avoir accueillie au sein du groupe Biomolécules excitées et d'avoir prêté une attention soutenue à mes travaux. Je remercie François Piuzzi, *maître ès reverse-engineering*, qui m'a montré que « c'est avec du vieux qu'on fait du neuf » (Jacques Brel). Je lui suis reconnaissante pour tous ses conseils et pour ses chansons, dont j'ai pu apprécier la diversité. Je tiens à remercier Benjamin Tardivel pour sa gentillesse, sa disponibilité et ses conseils techniques. Mes pensées vont également vers Clélia Canuel dont la compagnie, la bonne humeur et les discussions scientifiques alternant avec les moments de détente m'ont été bénéfiques. Je remercie tous les membres du groupe pour ces trois années et demie passées dans une ambiance aussi agréable et dynamique.

Je souhaite remercier Jean-Pierre Dognon pour avoir encadré la partie théorique de mes travaux. Je tiens à saluer toute l'équipe de Chimie Théorique « du 2<sup>eme</sup> » pour leur gentillesse



et leur disponibilité.

Je remercie le groupe Molecular Dynamics du Fundamental Optik Material (FOM) Institute for Plasma Physics à Rijnhuizen (Pays-Bas) avec lequel j'ai collaboré. Je remercie à cette occasion les membres du FOM pour leur assistance technique.

Je remercie très chaleureusement mes parents pour leur soutien et leurs encouragements. Je dédie ce travail à Wuthina dont la présence et la bonne humeur me permettent de toujours avancer. Enfin, je souhaite remercier Rodolphe pour son soutien inconditionnel, sa patience et sa présence à mes côtés.

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>9</b>
1.1	Généralités . . . . .	9
1.1.1	Acides aminés . . . . .	12
1.1.2	Diagramme de Ramachandran . . . . .	14
1.1.3	Structures secondaires . . . . .	15
1.2	Caractérisation des peptides et des protéines . . . . .	21
1.2.1	Diffraction X et phase liquide . . . . .	22
1.2.2	Résultats en matrice . . . . .	23
1.2.3	Résultats en phase gazeuse . . . . .	24
1.3	Problématique . . . . .	30
	Références bibliographiques du chapitre 1 . . . . .	34
<b>2</b>	<b>Méthodologies</b>	<b>41</b>
2.1	Dispositif expérimental . . . . .	41
2.1.1	Préparation des molécules . . . . .	42
2.1.2	Spectromètre de masse . . . . .	45
2.1.3	Systèmes lasers . . . . .	47
2.1.4	Synchronisation . . . . .	51
2.2	Méthodes spectroscopiques . . . . .	53
2.2.1	Ionisation résonnante à 2 photons (IR2P) . . . . .	54
2.2.2	Spectroscopie de dépopulation UV-UV (Hole-Burning) . . . . .	55
2.2.3	Spectroscopie de dépopulation IR-UV . . . . .	55

2.3	Méthodes de calculs . . . . .	57
2.3.1	Choix de la méthode . . . . .	58
	Références bibliographiques du chapitre 2 . . . . .	65
<b>3</b>	<b>Préférences conformationnelles locales</b>	<b>67</b>
3.1	Ac-Phe-NH <sub>2</sub> : recherche des formes typiques de la biologie . . . . .	68
3.1.1	Étude expérimentale . . . . .	69
3.1.2	Calculs et attribution . . . . .	77
3.1.3	Établissement d'un modèle phénoménologique . . . . .	81
3.1.4	Effet du cycle aromatique . . . . .	81
3.1.5	Conclusion . . . . .	83
3.2	Dipeptides Ac-Phe-Xxx-NH <sub>2</sub> : effet de l'ajout d'un résidu . . . . .	85
3.2.1	Spectroscopie électronique . . . . .	85
3.2.2	Spectroscopie IR . . . . .	87
3.2.3	Résultats théoriques et attributions . . . . .	91
3.3	Dipeptides Ac-Xxx-Phe-NH <sub>2</sub> : effet de l'ordre de la séquence . . . . .	101
3.3.1	Spectroscopie électronique . . . . .	101
3.3.2	Spectres IR . . . . .	103
3.3.3	Calculs : recherche des conformations pour Ac-Gly-Phe-NH <sub>2</sub> . . . . .	106
3.3.4	Conformères A et B de Ac-Pro-Phe-NH <sub>2</sub> . . . . .	108
3.3.5	Conclusion : effet de l'ordre de la séquence . . . . .	111
3.4	Conclusion du chapitre 3 . . . . .	113
	Références bibliographiques du Chapitre 3 . . . . .	115
<b>4</b>	<b>Coudes <math>\beta</math></b>	<b>119</b>
4.1	Spectroscopie électronique . . . . .	121
4.2	Signature IR . . . . .	121
4.2.1	Amide A . . . . .	122
4.2.2	Utilisation du modèle de couplage vibrationnel . . . . .	124
4.2.3	Amides I et II . . . . .	125

4.2.4	Comparaison avec les calculs . . . . .	126
4.3	Effet de la séquence : phénylalanine vs. tyrosine et NH <sub>2</sub> vs. NHMe . . . . .	132
4.4	Stabilité des coudes $\beta$ . . . . .	137
4.5	Résumé et conclusions sur les coudes $\beta$ . . . . .	142
	Références bibliographiques du Chapitre 4 . . . . .	145
<b>5</b>	<b>À la recherche des hélices</b> . . . . .	<b>147</b>
5.1	Ac-Phe-Gly-Gly-NH <sub>2</sub> . . . . .	148
5.1.1	Diversité du paysage conformationnel . . . . .	149
5.1.2	Conformère B . . . . .	151
5.1.3	Conclusion . . . . .	157
5.2	Effet de la séquence : tripeptides -Ala-Ala-Ala- . . . . .	159
5.2.1	Étude théorique du tripeptide Ac-Ala-Ala-Ala-NH <sub>2</sub> . . . . .	160
5.3	Étude expérimentale des tripeptides . . . . .	163
5.3.1	Conservation des formes locales . . . . .	164
5.3.2	Vers l'hélice $3_{10}$ . . . . .	168
5.3.3	Émergence de l'hélice $3_{10}$ . . . . .	170
5.4	Discussion : effet du résidu aromatique . . . . .	173
5.4.1	Phe-Ala-Ala . . . . .	174
5.4.2	Ala-Ala-Phe . . . . .	177
5.4.3	Ala-Phe-Ala . . . . .	180
5.4.4	Signature de l'hélice $3_{10}$ . . . . .	181
5.5	Conclusion du Chapitre 5 . . . . .	182
	Références bibliographiques du Chapitre 5 . . . . .	185
	<b>Conclusion générale</b> . . . . .	<b>187</b>
	<b>A Modèle de couplage vibrationnel</b> . . . . .	<b>193</b>
	<b>B Article</b> . . . . .	<b>197</b>

<b>C</b>	<b>Tripeptides</b>	<b>213</b>
C.1	Énergies d'une sélection de structures calculées pour le tripeptide Ac-Phe-Gly-Gly-NH <sub>2</sub> . . . . .	213
C.2	Énergies et paramètres géométriques d'une sélection de structures calculées pour les tripeptides de type -Ala-Ala-Ala- . . . . .	213

# Chapitre 1

## Introduction

Ce travail s'est focalisé sur l'observation en phase gazeuse des structures secondaires typiques de la biologie. L'idée directrice fut de voir dans un premier temps si ces structures sont intrinsèques aux petits polypeptides, c'est-à-dire si partant de molécules dépliées à haute température, ces molécules flexibles et complexes adoptent spontanément ces structures secondaires en se refroidissant dans le jet, en l'absence de tout environnement. Dans un second temps, le réseau de liaisons hydrogène de chaque conformère formé dans le jet a été caractérisé avec toute la précision permise par la résolution spectrale apportée par les conditions de détente supersonique, illustrant ainsi la nature des différentes interactions mises en jeu dans le repliement. Après l'étude des structures secondaires de petite taille comme les coudes, nous avons ensuite abordé la question de la formation d'hélices dans les chaînes de plus grande taille.

### 1.1 Généralités

Un acide aminé est un composé de formule générale  $\text{NH}_2\text{-CH-R-COOH}$  comprenant un carbone noté  $\text{C}\alpha$ , en général asymétrique, auquel sont rattachés un groupement amine  $\text{NH}_2$ , un groupement acide carboxylique  $\text{COOH}$  et une chaîne latérale notée  $\text{R}$  suivant le schéma de la Figure 1.1. La condensation des acides aminés par élimination de molécules d'eau conduit à la formation des polypeptides, des polymères linéaires dont les unités monomériques correspon-

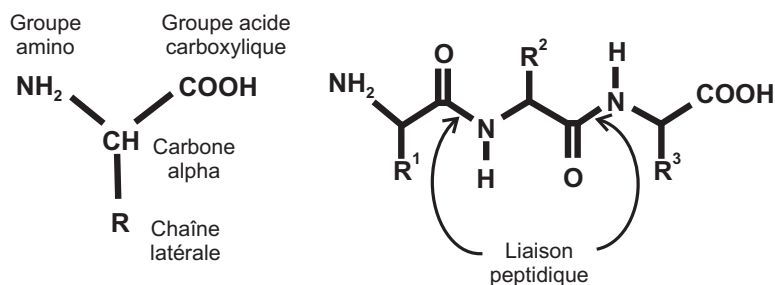


FIG. 1.1 – a) Schéma d'un acide aminé. Le groupe basique NH<sub>2</sub> est placé à gauche et le groupe acide COOH à droite. Le carbone alpha central est asymétrique, excepté pour la glycine où R = H. b) Schéma d'une chaîne de polypeptides contenant trois acides aminés ou résidus. Par convention, la chaîne s'écrit avec la partie N-terminale à gauche et l'extrémité C-terminale à droite.

dant aux acides aminés sont appelés résidus. Ils sont alors reliés par une liaison amide rigide CO-NH appelée liaison peptidique, la succession d'acides aminés et de liaisons peptidiques constituant la chaîne peptidique (Figure 1.1). Lorsque celle-ci comprend moins de 50 résidus on parle souvent d'oligopeptide ou de peptide, au-delà on parle généralement de protéine.

Dans les années 1930, Astbury montra que les cheveux possédaient des profils de diffraction par rayons X très caractéristiques, et que ceux-ci changeaient selon que le cheveu était étiré ou non. L'analyse des profils de diffraction pour la forme étirée le conduisit à proposer un modèle dans lequel la chaîne de polypeptides était complètement étendue et établissait des liaisons hydrogène entre les chaînes adjacentes antiparallèles. De fait, il en déduisit que l'autre forme devait adopter une configuration plus ou moins repliée.

Dans les années 1950, en s'intéressant au concept de liaisons chimiques, rayon atomique, longueur de liaisons, etc., Pauling et collaborateurs mirent en évidence le caractère rigide de la liaison peptidique CO-NH. Ils proposèrent alors deux modèles de structures compatibles avec l'arrangement plan de la liaison peptidique : une structure en hélice dans laquelle les liaisons présentent un changement d'orientation de leur plan d'environ 100° les unes par rapport aux autres et s'arrangent de façon périodique, et une structure étendue dans laquelle les liaisons peptidiques successives présentent des orientations relatives opposées. Ils nommèrent ces deux arrangements hélice  $\alpha$  [1] et feuillet plissé  $\beta$  ou «  $\beta$ -sheet » [2] (Figure 1.2). Ils conclurent que leur formation provenait du caractère à la fois flexible et rigide de la chaîne peptidique. La rigidité est assurée par l'arrangement plan de la liaison peptidique tandis que les rotations

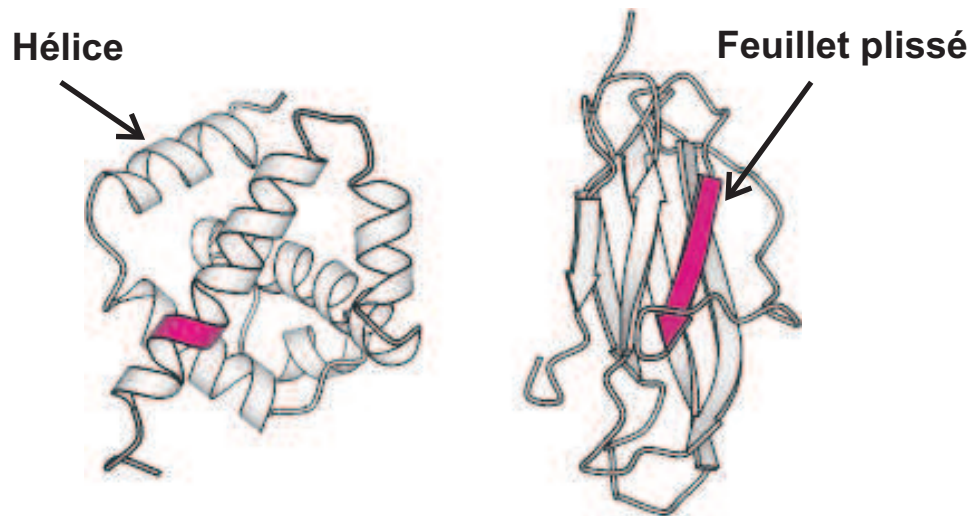


FIG. 1.2 – Exemple de structures de protéines comportant majoritairement des portions en hélices ou en feuillets  $\beta$ .

possibles autour des liaisons C-C, C-N du squelette et des liaisons  $C\alpha-C\beta$  de la chaîne latérale permettent à la chaîne principale d'adopter plusieurs conformations.

Parallèlement, en 1958, Kendrew et collaborateurs identifièrent la première structure tridimensionnelle d'une protéine, la myoglobine [3]. Il a été rapidement démontré que la majorité des fibres pouvaient être regroupées en trois grandes classes de structures selon leur profil de diffraction par rayons X : les structures de type  $\alpha$ -kératine en hélice qui comprennent les protéines fibreuses contenues dans les cheveux, les ongles, les sabots des animaux ; les structures de type  $\beta$ -kératine étendues dans lesquelles sont classées les fibres de cheveux étirés, la soie, etc ; et des structures repliées mais non périodiques.

À la même période, Linderstrøm-Lang au Danemark proposa plusieurs niveaux d'organisation pour la structure des protéines [4], chaque niveau étant caractérisé par des forces spécifiques ; les niveaux « supérieurs » étant composés d'éléments des niveaux « inférieurs ». Cette organisation s'articule de la façon suivante (Figure 1.3) :

- la **structure primaire** : il s'agit de la séquence d'acides aminés. Elle détermine les propriétés de structure et de réactivité chimique de la chaîne de polypeptides d'une protéine donnée.
- la **structure secondaire** : il s'agit en premier lieu d'un élément de terminologie qui résulte de l'observation dans les protéines de motifs remarquables et récurrents. Les



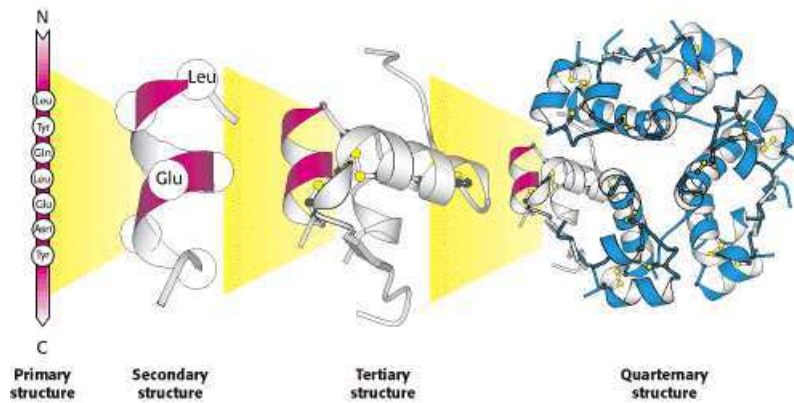


FIG. 1.3 – Schéma des structures primaire, secondaire, tertiaire et quaternaire de la protéine d'insuline.

structures secondaires constituent le premier niveau d'organisation tridimensionnelle. Parmi ces motifs ont été identifiés les hélices  $\alpha$  et  $3_{10}$ , les feuillets  $\beta$  - ces structures sont quasi périodiques - des coudes (ou boucles), et des formes moins régulières appelées polyproline. On parle parfois d'organisation locale de la chaîne car de telles structures sont répétées sur de petits segments le long de la chaîne de polypeptides. Sa conformation est alors contrôlée par l'établissement de liaisons hydrogène, de liaisons de van der Waals, d'interactions électrostatiques entre les différents résidus qui composent la séquence.

- la **structure tertiaire** est due à l'agencement relatif de structures secondaires. Six hélices  $\alpha$  présentant des contacts sont par exemple retrouvées dans la structure de la myoglobine (Figure 1.4).
- la **structure quaternaire** : c'est l'assemblage non covalent de différentes unités identiques et indépendantes. Ainsi, les canaux des protéines transmembranaires sont en fait des hexamères de protéines plus petites.

### 1.1.1 Acides aminés

Les acides aminés en solution aqueuse se présentent sous une forme zwitterionique portant des charges formelles mais globalement neutre, si l'on excepte d'éventuelles charges portées par les chaînes latérales : la fonction amine fixe un proton formant un groupe ammonium  $\text{NH}_3^+$  et la fonction acide prend sa forme basique carboxylate  $\text{COO}^-$ . En phase gazeuse, la forme

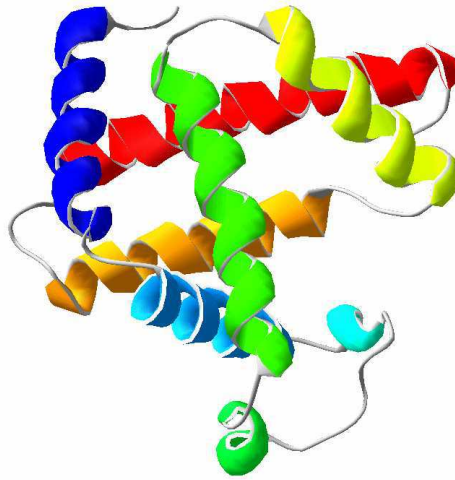
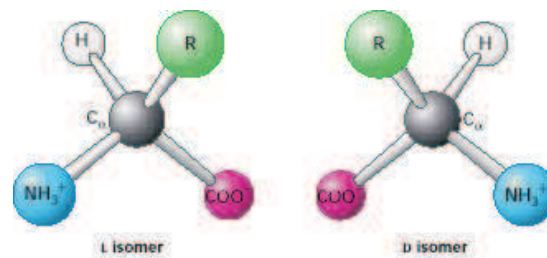


FIG. 1.4 – Structure de la myoglobine.

FIG. 1.5 – Formes chirales L et D d'un acide aminé  $\text{NH}_3^+\text{-CHR-COO}^-$ . Les deux formes sont images l'une de l'autre dans un miroir.

la plus stable est neutre, la forme zwitterionique n'étant généralement pas observée à moins d'être stabilisée par une autre espèce.

On distingue vingt acides aminés naturels pouvant comporter des chaînes latérales hydrophobes, polaires, acides, ou basiques, chargées ou non. La fonctionnalité de la protéine est liée à la nature de ces chaînes latérales. Les groupements polaires sont en général trouvés à la surface des protéines, et donc accessibles aux molécules d'eau, alors que les résidus non polaires, et donc hydrophobes, constituent le coeur de la protéine. On note les carbones de la chaîne latérale  $C\beta$ ,  $C\gamma$ , etc, en partant du carbone  $C\alpha$ . Les acides aminés, hormis la glycine (Gly ou G) dont le carbone n'est pas asymétrique, existent sous deux formes énantiomères présentant deux chiralités traditionnellement notées *L* et *D* (Figure 1.5). La chiralité *L* est la seule forme observée dans le monde vivant.

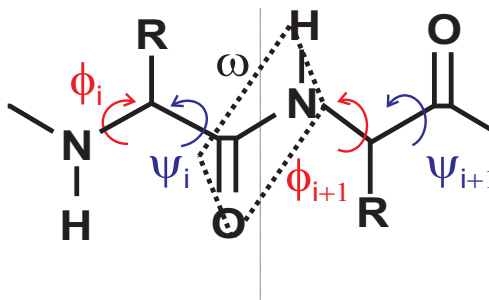


FIG. 1.6 – Angles dièdres  $\phi$  et  $\psi$  définissant la conformation de la chaîne peptidique. L'angle  $\omega$  définit la conformation de la liaison peptidique CO-NH plane. Il est en général proche de  $180^\circ$ , ce qui correspond à une configuration *trans*.

### 1.1.2 Diagramme de Ramachandran

C'est le caractère flexible de la chaîne peptidique, dû aux différentes liaisons simples  $\sigma$ , qui lui permet d'adopter des formes plus ou moins repliées menant à la formation des structures secondaires. Les torsions autour des liaisons N-C $\alpha$  et C $\alpha$ -C sont décrites par les angles dièdres  $\phi$  et  $\psi$  et sont donc très aisées (Figure 1.6). Ces deux angles définissent l'orientation relative de deux liaisons peptidiques ; ils permettent de décrire les diverses conformations que peut adopter localement la chaîne autour d'un résidu. En variant les valeurs de  $\phi$  et  $\psi$ , Ramachandran établit une carte donnant les valeurs d'angles dièdres compatibles en termes de gêne stérique prenant le nom de diagramme de Ramachandran [5]. Les atomes étant considérés comme des sphères dures dont la dimension correspond à leur rayon de van der Waals, les combinaisons de  $(\phi, \psi)$  mènent alors à une zone permise et une zone interdite. Les zones interdites résultent d'encombrements stériques entre la chaîne latérale R et la chaîne principale. Les régions permises correspondent aux valeurs de  $\phi$  et  $\psi$  observées dans les protéines pour les conformations hélice  $\alpha$  et feuillet  $\beta$  (Figure 1.7). Les couples  $(\phi, \psi)$  à la frontière entre les formes  $\alpha$  et  $\beta$  et les combinaisons interdites correspondent à des formes moins stables ; c'est par exemple le cas de l'hélice  $3_{10}$ . L'acide aminé glycine, en raison de l'absence de chaîne latérale, est moins contraint et présente donc une zone accessible beaucoup plus grande. Ont ainsi été identifiés neuf types de conformations notés  $\alpha_{L/D}$ ,  $\beta_L$ ,  $\gamma_{L/D}$ ,  $\delta_{L/D}$  et  $\epsilon_{L/D}$  pour un acide aminé isolé. Les notations *L* et *D* font référence à la propension des acides aminés de chiralité *L* et *D* à adopter ces configurations.

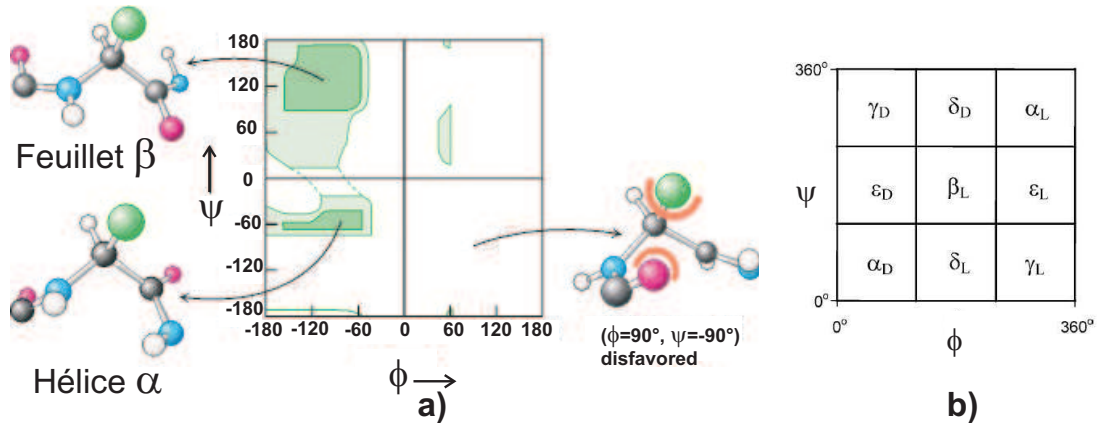


FIG. 1.7 – Diagramme de Ramachandran. a) Les zones foncées (vertes) sont associées à des combinaisons de  $\phi$  et  $\psi$  permises et correspondent aux conformations en hélice  $\alpha$  ( $-150^\circ < \phi < -60^\circ, \psi \approx -60^\circ$ ) et en feuillet  $\beta$  ( $-150^\circ < \phi < -60^\circ, 180^\circ < \psi < 90^\circ$ ). Les zones blanches correspondent à des régions interdites présentant des contacts défavorables, excepté pour la glycine qui n’a pas de chaîne latérale et ne présente pas de problème d’encombrement stérique. b) Conformations possibles pour un acide aminé.

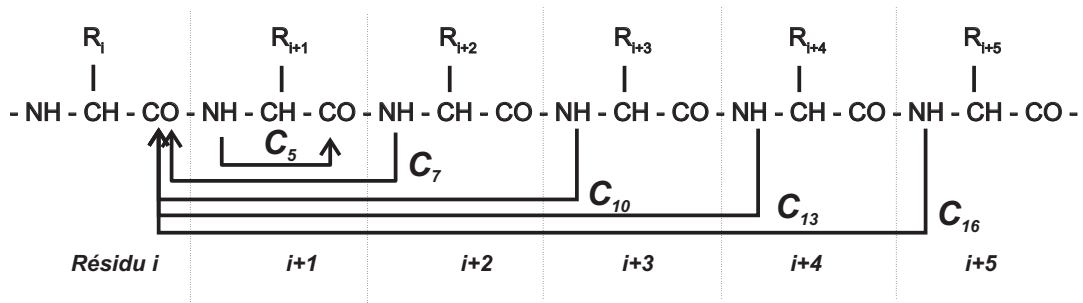


FIG. 1.8 – Schéma des liaisons hydrogène possibles au sein de la chaîne peptidique.

### 1.1.3 Structures secondaires

Les configurations en hélice  $\alpha$  et en feuillet plissé  $\beta$  sont les formes quasi périodiques les plus courantes dans les protéines mais elles sont relativement rigides et donnent aux protéines un aspect fibreux alors qu’elles adoptent en général un aspect globulaire. Les « coudes » sont des structures non périodiques qui assurent le changement de direction de la chaîne peptidique dans ces structures globulaires. Les structures qui ne peuvent être regroupées dans ces trois grandes classes sont rangées dans une catégorie dite pelote globulaire ou « random coil ». Un bref aperçu des diverses structures secondaires est donné ci-après.

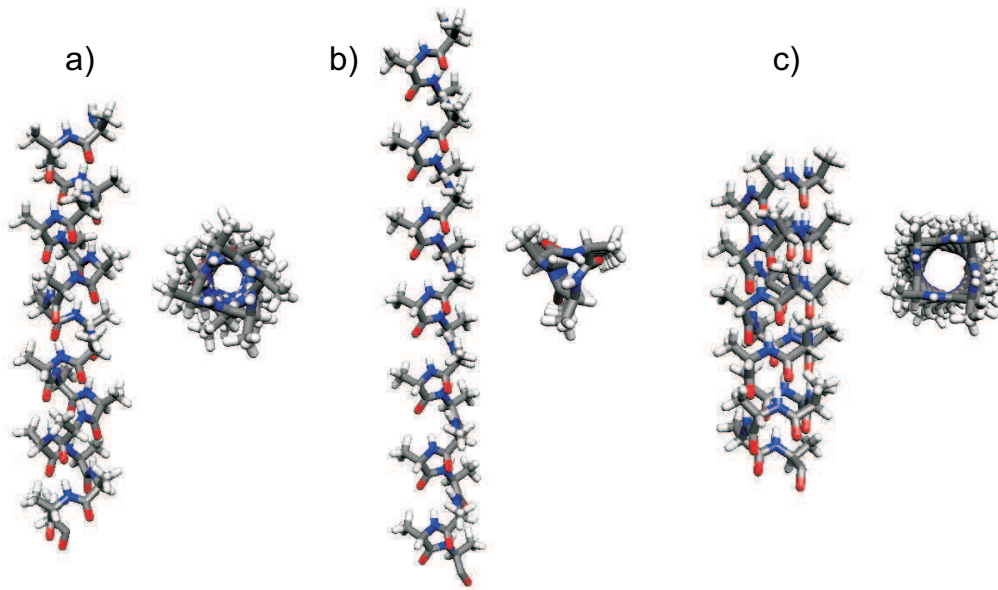


FIG. 1.9 – Conformations en hélice a)  $\alpha$ , b)  $3_{10}$  et c)  $\pi$  de l'oligopeptide  $\text{Ala}_{22}$ .

### Les hélices

Les hélices pseudopériodiques sont donc définies par un couple  $(\phi, \psi)$  caractéristique, et surtout par le nombre de résidus par tour d'hélice, noté  $n$ , et par le pas, noté  $p$ . Les hélices sont entre autres stabilisées par des liaisons hydrogène spécifiques. L'arrangement de l'hélice est tel que les groupes CO et NH des liaisons peptidiques pointent dans des directions proches de l'axe de l'hélice, conduisant à un alignement des dipôles des liaisons peptidiques avec l'axe, c'est-à-dire que tous les groupements CO pointent en direction de l'extrémité C-terminale et les groupements NH en direction de la partie N-terminale. On distingue trois types d'hélices dont les caractéristiques sont résumées dans le Tableau 1.1 (les valeurs de  $\phi$  et  $\psi$  correspondent à des hélices idéales).

L'**hélice**  $\alpha$  est la forme la plus abondante dans les protéines : 32-38% des résidus sont retrouvés dans des hélices  $\alpha$ . Celle-ci compte 3.6 résidus par tour [6]. Sa configuration permet au groupement CO du résidu  $(i)$  d'être en face du groupement NH du résidu  $(i+4)$ . Cet alignement ainsi que la distance entre les deux groupes sont idéaux pour former une liaison hydrogène. Les liaisons hydrogène forment le motif répétitif  $\text{CO}(i)\leftarrow\text{NH}(i+4)$  dans lequel chaque liaison ferme un cycle à 13 atomes. On note ce type de liaison hydrogène  $\text{C}_{13}$  ; on

fait aussi référence à l'hélice  $\alpha$  en tant que hélice  $3.6_{13}$ , conformément à la notation générale  $n_m$  des hélices où  $n$  est le pas en résidu et  $m$  le type de liaison hydrogène (Figures 1.8 et 1.9a). Les chaînes latérales sont à l'extérieur de l'hélice et sont orientées en direction de l'extrémité N-terminale.

L'**hélice  $3_{10}$**  est beaucoup moins fréquente que la structure précédente. Elle est retrouvée principalement au début ou à la fin d'un tour de l'hélice  $\alpha$  et parfois de façon plus isolée, et est limitée à quelques résidus (4-5). Le nom d'hélice  $3_{10}$  résulte de la liaison hydrogène  $\text{CO}(i)\leftarrow\text{NH}(i+3)$  fermant un cycle à 10 atomes (ou  $C_{10}$ ), il y a donc 3 résidus par tour (Figures 1.8 et 1.9b). Dans les protéines, seuls 3.4% des résidus appartiennent à une hélice  $3_{10}$  [6]; ceci s'explique par plusieurs raisons : les valeurs de  $\phi$  et  $\psi$  ( $-74^\circ$ ,  $-4^\circ$ ) correspondent à une région qui est à la limite des zones permises dans le diagramme de Ramachandran de la Figure 1.7; les dipôles des liaisons peptidiques ne sont pas parfaitement alignés dans l'axe de l'hélice mais s'écartent d'environ  $30^\circ$ ; l'arrangement des chaînes latérales est moins favorable que pour l'hélice  $\alpha$ , et enfin l'hélice  $3_{10}$  est moins dense.

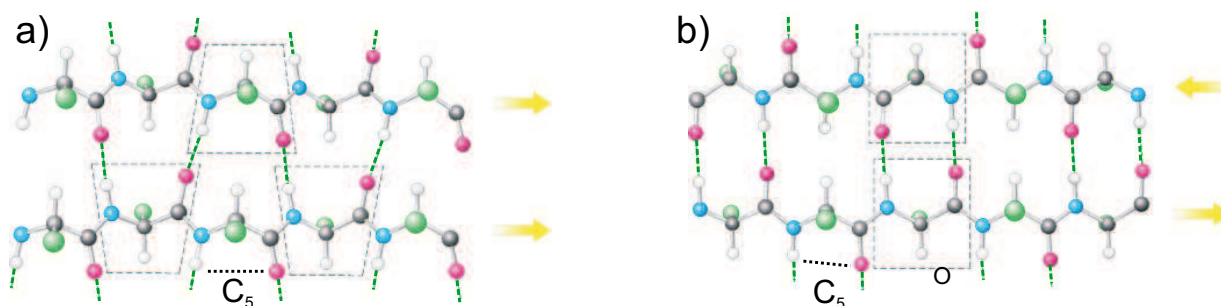
L'**hélice  $\pi$**  est caractérisée par une liaison hydrogène  $\text{CO}(i)\rightarrow\text{NH}(i+5)$ . Elle est extrêmement rare, parfois localisée aux extrémités de l'hélice  $\alpha$ . Sa rareté est due aux valeurs de  $\phi$  et  $\psi$  ( $-57^\circ$ ,  $-70^\circ$ , pour l'hélice  $\pi$  idéale) qui indiquent une région à l'extrême limite d'une zone permise dans le diagramme de Ramachandran.

Hélice	Fréquence	$\phi$	$\psi$	n	liaison H CO $\rightarrow$ NH	nombre d'atomes impliqués dans le cycle
$\alpha$	abondant	-58	-47	3.6	i, i+4	13
$3_{10}$	peu fréquent	-74	-4	3	i, i+3	10
$\pi$	rare	-57	-70	4.4	i, i+5	16

TAB. 1.1 – Caractéristiques des hélices  $\alpha$ ,  $3_{10}$  et  $\pi$  « idéales »

### Les feuillets et les brins $\beta$

On retrouve 20-28% des résidus dans des conformations  $\beta$  [6]. Ce sont des structures étendues, dont les angles  $\phi$  et  $\psi$  correspondent à une zone permise dans le diagramme de

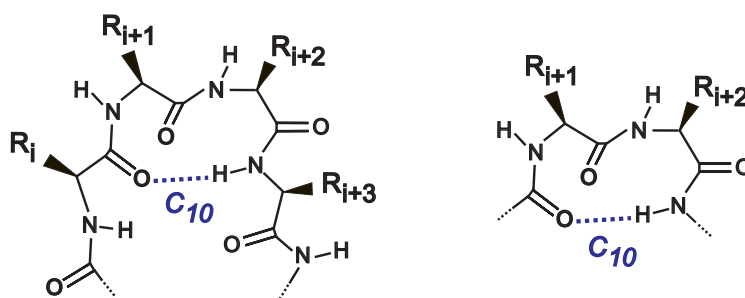
FIG. 1.10 – Schéma de feuillets  $\beta$  a) parallèle et b) antiparallèle.

Ramachandran (correspondant à la conformation  $\beta_L$  de la Figure 1.7b), les valeurs typiques sont  $(-140^\circ, 130^\circ)$  (Figure 1.7). Elles présentent une interaction notée  $C_5$  entre les groupes CO et NH d'un même résidu, relativement faible en raison de la non linéarité de la liaison  $H \cdots O$  (l'angle NHO est proche de  $105^\circ$ ) (Figure 1.8).

Un polypeptide dans la conformation étendue  $\beta$ , où tous les résidus adoptent une conformation  $\beta_L$ , est appelé **brin**  $\beta$  ou «  $\beta$ -strand » : l'agencement de deux brins par la formation de liaisons hydrogène entre les deux chaînes principales constitue un feuillet  $\beta$ . Ces liaisons inter-résidus sont responsables de la stabilité des feuillets des protéines. Ceux-ci peuvent être soit parallèles, c'est-à-dire que les brins ont la même direction, soit antiparallèles, les brins ont des directions opposées (Figure 1.10). Les liaisons peptidiques de résidus adjacents ainsi que les chaînes latérales pointent dans des directions opposées. Les conformations  $\beta$  sont aussi retrouvées dans les structures dites en épingle à cheveux, ou «  $\beta$ -hairpin », dont la chaîne principale est repliée en formant un coude. Ce coude dit  $\beta$  permet la formation de liaisons hydrogène au sein de la chaîne - qui est alors équivalente à deux brins antiparallèles - analogues à celles observées dans les feuillets  $\beta$ . Afin de ne pas confondre les conformations étendues  $\beta$  et les coudes, nous appellerons les formes étendues  $\beta_L$  dans la suite du manuscrit. Nous reviendrons à cette terminologie dans le chapitre 3.

### Les coudes

Les coudes ou « turns » représentent une troisième classe de structures secondaires. Leur fonction principale est d'assurer le changement de direction de la chaîne de polypeptides, comme nous l'avons vu avec le  $\beta$ -hairpin. Ils sont principalement localisés à la surface des

FIG. 1.11 – Schéma d'un coude  $\beta$ .

protéines. Les sites de reconnaissance des anticorps, de phosphorylation, etc, sont souvent retrouvés à proximité des coudes. Il existe deux grandes classes de coudes, les **coudes**  $\beta$  et les **coudes**  $\gamma$  moins fréquents. Les premiers ont été identifiés par Ventakachalam en 1968 [7]. Environ 30% des résidus sont impliqués dans des coudes  $\beta$ .

Les coudes  $\beta$  sont caractérisés par une liaison hydrogène  $\text{CO}(i) \leftarrow \text{NH}(i+3)$  impliquant quatre résidus dans un cycle à dix atomes ; la conformation est communément appelée  $C_{10}$  (Figures 1.8 et 1.11). Suivant la valeur des angles dièdres  $\phi$  et  $\psi$  des résidus centraux ( $i+1$ ) et ( $i+2$ ), les coudes  $\beta$  sont classés dans six types principaux : I, II, I', II', VIa, VIb [8, 9, 10]. Les coudes  $\beta$  qui ne répondent pas aux paramètres géométriques de ces types appartiennent au type IV (Tableau 1.2). Les coudes de type I et II sont les formes les plus abondantes et représentent plus de 95% des coudes  $\beta$  observés dans les protéines [11]. Ils diffèrent par un retournement de la liaison amide centrale. Le squelette peptidique du type I' (respectivement du type II') est l'image du type I (respectivement du type II) dans un miroir (Figure 1.12). Dans les coudes  $\beta$  de type VI, la liaison peptidique centrale entre les résidus ( $i+1$ ) et ( $i+2$ ) adopte une configuration *cis*, alors que la configuration est *trans* dans la majorité des protéines. En fait, le type VI est observé exclusivement avec le résidu Pro en position ( $i+2$ ) dans le coude. L'hélice  $3_{10}$  qui présente une liaison hydrogène fermant un cycle à dix atomes définit également un type de coude  $\beta$  noté type III. Celui-ci étant une structure pseudopériodique, les couples  $(\phi, \psi)$  des résidus ( $i+1$ ) et ( $i+2$ ) sont identiques ; ils ont pour valeur  $(-60^\circ, 30^\circ)$ . Ces valeurs sont comparables à celles du type I ; elles peuvent différer de  $10\text{-}20^\circ$ , aussi parle-t-on également de type I(III) pour qualifier le coude  $\beta$  de l'hélice  $3_{10}$ .



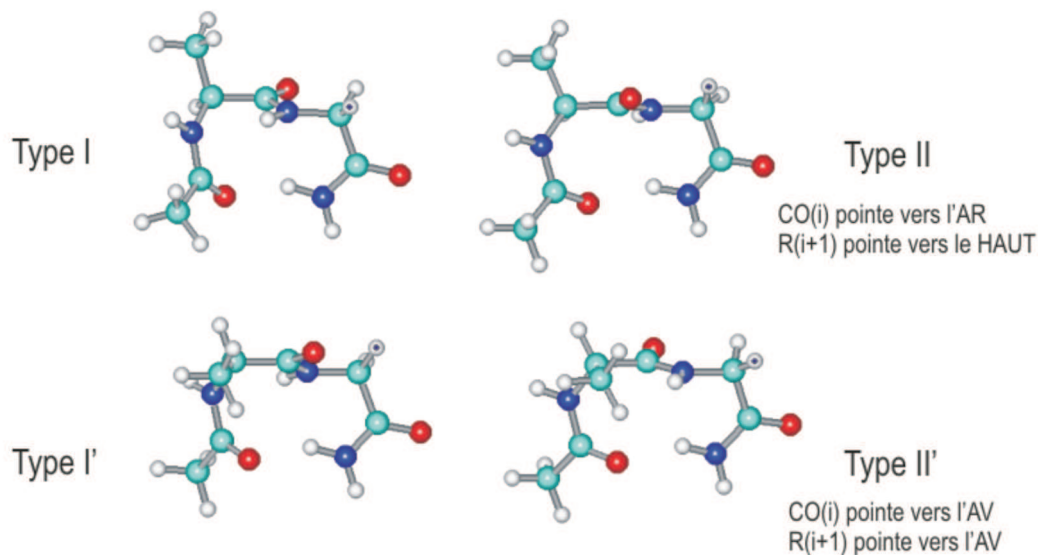


FIG. 1.12 – Exemples de coudes  $\beta$  de type I, I', II et II' pour le système Ac-Ala-Gly-NH<sub>2</sub>. Le groupe acétyle mime un résidu en position (i) et le groupe NH<sub>2</sub> terminal un résidu en position (i+3). Les groupes carbonyles (i) et (i+1) du type I pointent vers l'arrière; le type I' est image du type I dans un miroir, les groupes carbonyles pointent donc vers l'avant. Dans le type II, le groupe CO(i) pointe vers l'arrière, le groupe CO(i+1) vers l'avant, et inversement dans le type II'.

Type de coude $\beta$	Résidu (i+1)		Résidu (i+2)	
	$\phi$	$\psi$	$\phi$	$\psi$
type I	-60	-30	-90	0
type I'	60	30	90	0
type II	-60	120	80	0
type II'	60	-120	-80	0
type III	-60	30	-60	30
type VIa				
type VIa1	-60	120	-90	0
type VIa2	-120	120	-60	0
type VIb	-120	120	-60	0

TAB. 1.2 – Angles dièdres des résidus centraux (i+1) et (i+2) de différents types de coudes  $\beta$ , les angles sont donnés en degrés.

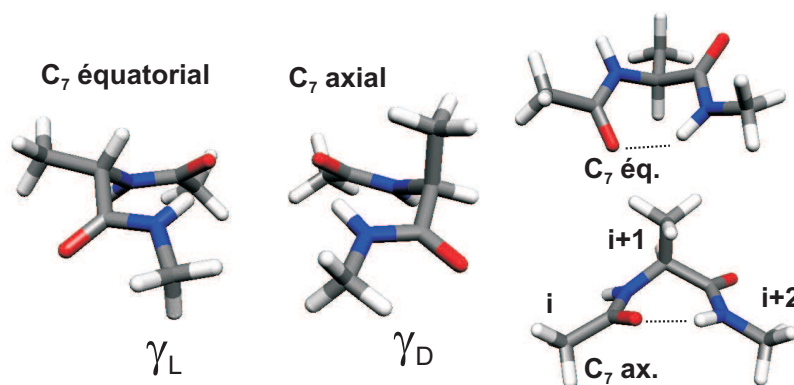


FIG. 1.13 – Conformations coude  $\gamma$  classique  $\gamma_D$  et inverse  $\gamma_L$  de Ac-Ala-NHMe. Le groupement acétyle représente le résidu  $i$ , l’alanine est en position  $i+1$  et le groupement méthylamide représente le résidu  $i+2$ . Les notations axiales et équatoriales correspondent aux positions axiale et équatoriale de la chaîne latérale par rapport au cycle à sept chaînons.

Les coudes  $\gamma$  sont des structures qui permettent également le changement de direction de la chaîne. Ils sont caractérisés par une liaison hydrogène  $\text{CO}(i) \leftarrow \text{NH}(i+2)$  impliquant deux résidus dans un cycle à sept atomes ; la conformation est communément appelée  $C_7$  (Figure 1.8). Ces structures sont plus compactes que les précédentes. On distingue deux types de coudes  $\gamma$  dont les squelettes sont images dans un miroir : le coude  $\gamma_D$  classique et le coude  $\gamma_L$  inverse (Figure 1.13). La position respective de la chaîne latérale par rapport au squelette, plus précisément la position axiale ou équatoriale du carbone  $C\beta$  de la chaîne latérale, définit une configuration  $C_7$  axiale ou équatoriale. Le coude  $\gamma_D$  est beaucoup moins fréquent que le coude inverse  $\gamma_L$  dans les protéines ; on le retrouve par exemple à la fin des boucles des  $\beta$ -hairpin [12]. Les structures idéales  $\gamma_D$  et  $\gamma_L$  sont respectivement définies par les couples  $(\phi, \psi)$   $(80^\circ, -60^\circ)$  et  $(-80^\circ, 60^\circ)$  dans le diagramme de Ramachandran.

## 1.2 Caractérisation des peptides et des protéines

Depuis les travaux de Frederik Sanger en 1953 qui détermina pour la première fois, par hydrolyses successives, la séquence d’une protéine, l’insuline, et prouva qu’une protéine présentait une séquence d’acides aminés bien définie. Depuis, les structures primaires de plus de 100 000 protéines ont été identifiées par les techniques de spectrométrie de masse, de chromatographie, etc. Ces techniques qui permettent de remonter à la séquence d’acides aminés de

manière désormais automatisée ne donnent cependant pas accès aux structures secondaires.

### 1.2.1 Diffraction X et phase liquide

Les études par diffraction des rayons X ont montré leur efficacité dans la détermination de la structure tridimensionnelle des protéines [13]. Plus de 12 000 protéines solubles ont ainsi été étudiées mais les protéines membranaires restent encore très mal connues en raison de leur caractère plutôt hydrophobe, les rendant alors difficile à cristalliser. Il faut en outre souligner que cette méthode ne donne pas accès à la structure secondaire des protéines car les atomes d'hydrogène ne sont pas détectés.

En phase liquide, les méthodes de dichroïsme circulaire (CD), de spectroscopie vibrationnelle IR et Raman et de résonance magnétique nucléaire (RMN) permettent en revanche d'accéder à la structure secondaire des protéines (ou des chaînes de polypeptides). Ainsi, les mesures de RMN, ainsi que par diffraction des rayons X, ont permis d'élaborer la bibliothèque Protein Data Bank (PDB) qui regroupe les structures de plus de 31000 protéines. La RMN présente pourtant l'inconvénient d'avoir un temps caractéristique d'analyse relativement long, de l'ordre de la ms. Il en résulte que les structures observées peuvent correspondre à des structures moyennes car les différentes espèces en présence ont le temps de s'interconvertir. À l'inverse, les méthodes de dichroïsme circulaire et de spectroscopie IR présentent des temps caractéristiques d'analyse beaucoup plus courts, de l'ordre de  $10^{-13}$ s. Néanmoins la première méthode ne donne pas directement accès à la structure secondaire mais donne un pourcentage de chaque type de structure présent dans la protéine et les spectres de CD sont alors déconvolués en quatre composantes : l'hélice  $\alpha$ , le feuillet  $\beta$ , les coudes et la structure « random coil » (Figure 1.14).

Les groupements amides CO-NH étant très sensibles à la présence de liaisons hydrogène, leurs fréquences de vibration, en fait essentiellement celles des modes d'élongation CO appelés amide I, sont mesurées en spectroscopie IR en solution afin de déterminer la conformation de la chaîne peptidique. Le domaine de fréquences correspondant à ces modes est typiquement de  $1648-1662\text{ cm}^{-1}$  pour l'hélice  $\alpha$ , et de  $1620-1640\text{ cm}^{-1}$  et  $1670-1690\text{ cm}^{-1}$  pour le feuillet  $\beta$ .

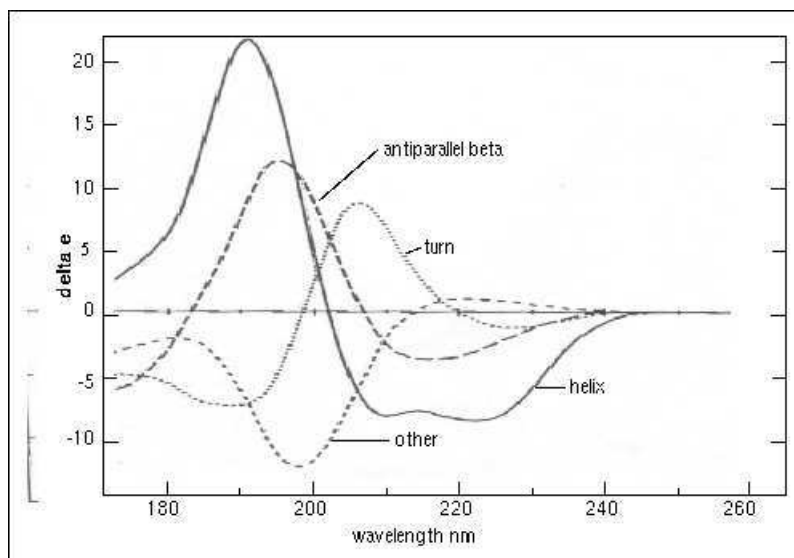


FIG. 1.14 – Spectre de dichroïsme circulaire montrant les domaines d’absorption des structures secondaires hélices  $\alpha$ , feuillets  $\beta$  antiparallèles et des coudes.

Toutefois, les domaines d’absorption de certaines structures étant parfois très proches, la faible résolution obtenue en phase liquide (typiquement de quelques centaines de  $\text{cm}^{-1}$ ) rend difficile la détection des formes minoritaires en raison de recouvrements spectraux ; c’est le cas des coudes  $\beta$  dont la signature est masquée par celle des feuillets  $\beta$  dont le coefficient d’absorption est plus élevé. Dans les études en solution, les biochimistes cherchant à caractériser des types de structures précises ont alors souvent recours à des peptides cycliques afin de contraindre le squelette à adopter la structure voulue. Les systèmes modèles choisis, de petits peptides synthétisés, contiennent alors un nombre entier de liaisons peptidiques [14, 15, 16, 17, 11]. Enfin, outre la faible résolution spectrale, se pose aussi le problème de la déconvolution des spectres, comme dans le cas de la technique de CD, leur interprétation s’appuyant alors fortement sur une contrepartie théorique.

### 1.2.2 Résultats en matrice

Les mesures de spectroscopie IR réalisées en matrice d’argon par le groupe de S. Stepanian ont permis d’accéder à la structure des acides aminés naturels glycine, valine et proline [18, 19, 20, 21]. En accord avec des calculs de chimie quantique menés par le même groupe, les formes observées présentent deux types de liaisons hydrogène intramoléculaires entre les fonctions

acide et basique de l'acide aminé, les liaisons  $\text{H}\cdots\text{O}=\text{C}$  et  $\text{N}-\text{H}\cdots\text{O}$ . Cependant l'inconvénient des études en matrice provient des interactions intermoléculaires avec le milieu, entraînant un décalage des spectres en fonction de la matrice utilisée. On peut par exemple citer les travaux de Bentwood et collaborateurs [22] sur le dimère de l'eau réalisés en matrice d'argon. Les auteurs trouvent des fréquences d'élongation OH plus basses de  $20\text{ cm}^{-1}$  par rapport aux résultats trouvés en phase gazeuse par Miller et collaborateurs [23] et également par rapport aux résultats obtenus par la technique d'anion multipolaire employée par Desfrancois et collaborateurs [24].

### 1.2.3 Résultats en phase gazeuse

#### Calculs

Les progrès apportés par les outils informatiques rendent désormais possibles les études de systèmes de plus en plus complexes. On a recours à deux approches dans l'étude théorique de systèmes biologiques en phase gazeuse : les calculs de chimie quantique, soit *ab initio*, soit utilisant la théorie de la fonctionnelle de la densité [25], et les calculs de dynamique moléculaire faisant appel à des champs de force [26]. En s'affranchissant des effets d'environnement, les études en phase gazeuse permettent d'accéder aux propriétés **intrinsèques** des systèmes : stabilité, réactivité, dynamique, etc.

Depuis le début des années 1990, de nombreux calculs *ab initio* sur de petits peptides ont cherché à caractériser les propriétés de repliement de la chaîne peptidique, mettant en évidence la très grande complexité du paysage conformationnel même pour des systèmes de taille modeste [27, 28, 29, 30, 31, 32, 33, 34, 35], et offrant également une description précise des structures secondaires, notamment en termes de signature spectrale, de stabilité, etc. [36, 37, 38, 39, 40, 41, 42].

#### Expériences

D'un point de vue expérimental, les techniques de la phase gazeuse traditionnellement employées dans la caractérisation des agrégats sont désormais couramment appliquées à l'étude des molécules biologiques [43, 44, 45, 46, 47]. Les résultats obtenus sont directement compa-

rables aux données théoriques et constituent donc un moyen de valider les différentes méthodes de calcul ; réciproquement, il est souvent indispensable de s'appuyer sur les calculs pour l'interprétation des résultats expérimentaux. La résolution spectrale apportée par les techniques de la phase gazeuse permet d'obtenir des données très précises sur les diverses interactions intramoléculaires qui stabilisent les systèmes en l'absence d'environnement. L'étude de complexes permet par la suite de caractériser les changements induits par la présence d'autres molécules et notamment des molécules d'eau en ayant accès aux interactions intermoléculaires et aux sites d'hydratation, ce qui n'est pas possible en phase condensée. En outre, le couplage avec des méthodes de spectrométrie de masse permet désormais de suivre pas à pas l'effet de chaque molécule de solvant et l'évolution d'un système en fonction de sa taille. L'utilisation d'une détente supersonique, dans laquelle la température associée aux degrés de liberté internes de la molécule est très basse, permet une simplification spectrale et est de plus propice à la formation de systèmes faiblement liés tels que les complexes de van der Waals.

Les études spectroscopiques de systèmes biologiques en phase gazeuse couvrent un domaine relativement large passant par l'étude de petits systèmes tels que les neurotransmetteurs [46, 48], les sucres [49, 50, 51], les bases de l'ADN [52, 53], les acides aminés [54, 55, 56], les complexes de ces acides aminés - complexes avec des cations métalliques ou hydrates [46] - et pouvant s'étendre à l'investigation de molécules plus grandes telles que des oligopeptides [57, 58, 59] et des protéines de taille moyenne [60, 61]. Ces études peuvent être conduites selon différents critères. Les techniques sont par exemple spécifiques selon que le système est neutre ou ionique notamment en ce qui concerne la production des molécules. Les molécules neutres sont traditionnellement mises en phase vapeur par des méthodes de chauffage, par désorption laser [62], ou plus récemment dans des gouttelettes d'hélium [63]. Les techniques MALDI (matrix-assisted laser desorption-ionisation) [64, 65] et ESI (electrospray ionisation) [66] sont les plus courantes pour la production d'ions. On peut également distinguer des approches différentes suivant la taille du système. Ainsi, les méthodes non spectroscopiques telles que la mobilité d'ions [67, 43] couplée aux méthodes de capture dissociative d'électrons (ECD pour electron capture dissociation) [61] s'appliquent davantage à des systèmes de grande taille. À l'inverse, d'autres techniques se sont pour l'instant limitées à l'étude des systèmes de

petite taille. La spectroscopie micro-onde permet de remonter à l'orientation du dipôle permanent dans le référentiel de la molécule et aux constantes rotationnelles d'un système. Par comparaison à des structures théoriques, elle permet l'attribution des conformères observés. Elle a été appliquée à l'étude de purines et de pyrimidines ainsi que leurs hydrates [68, 69], aux acides aminés [70, 71, 72, 73], ainsi qu'à des dérivés d'acides aminés [74]. La technique de double résonance IR-UV en ionisation et en fluorescence a permis d'accéder, entre autres, à la structure des acides aminés phénylalanine, tryptophane et leurs hydrates [45, 46, 75, 76, 77]. C'est la technique que nous avons employée dans nos travaux, nous y reviendrons dans la section Dispositif expérimental du Chapitre 2.

Les techniques citées permettaient d'accéder aux minimas de la surface d'énergie potentielle. Afin d'aller un pas plus loin, notamment afin d'estimer les barrières séparant les différents minimas, Zwier et collaborateurs ont très récemment pu mesurer et modifier la population de différents conformères d'un dérivé de l'acide aminé tryptophane en appliquant la technique d'isomérisation ou « Hole-filling spectroscopy » [78]. La technique repose sur l'excitation vibrationnelle au début de la détente d'un conformère donné, le spectre vibrationnel de la molécule est mesuré au préalable. L'énergie apportée au système est suffisante pour que celui-ci explore la surface d'énergie potentielle et relaxe soit vers la conformation initiale, soit vers les états qui correspondent à d'autres conformères de la molécule. L'évolution de la population de chaque conformère peut alors être suivie en mesurant le signal de fluorescence. En appliquant la méthode d'émission stimulée à cette technique, la même équipe a également pu mesurer les hauteurs de barrière des différents conformères d'un dérivé du tryptophane, la tryptamine [79] (Figure 1.15).

Nous allons à présent nous intéresser à certaines des techniques mentionnées précédemment, plus précisément à celles qui permettent de déterminer la structure des acides aminés, petits peptides et polypeptides en phase gazeuse. Nous traiterons dans des parties séparées les techniques appliquées à des systèmes ioniques et à des systèmes neutres.

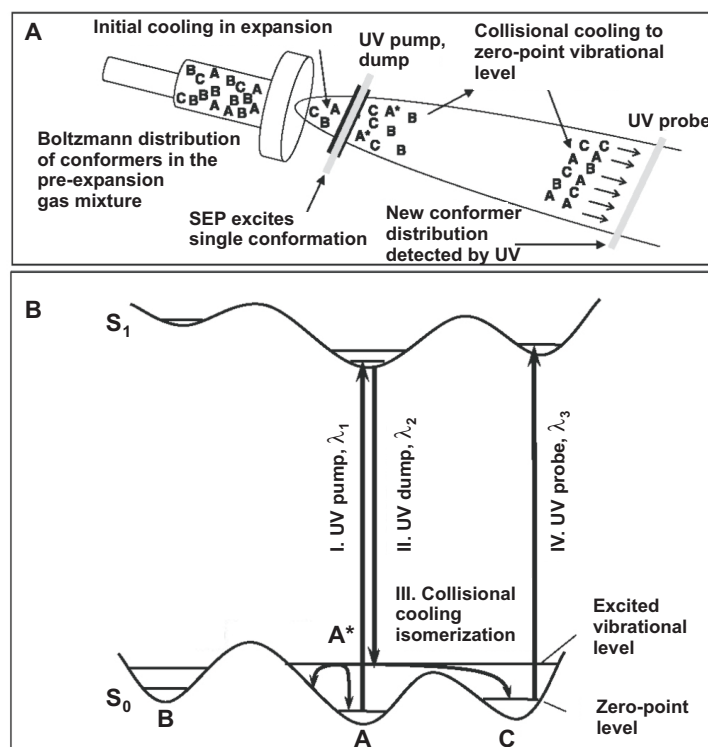


FIG. 1.15 – Schéma de principe de la technique d'isomérisation par émission stimulée tiré de la référence [79]. (B I) Les molécules A sont excitées par un premier faisceau ( $\lambda_1$ ) en début de détente. (B II) Une fraction des conformères A excités ( $A^*$ ) est amenée à un niveau vibrationnel excité de l'état fondamental par un second laser ( $\lambda_2$ ). (B III) Par refroidissement collisionnel dans la détente, la molécule relaxe soit vers son état fondamental, soit vers les autres conformères (ici le conformère C), si l'énergie de ce niveau vibrationnel est supérieure à la hauteur de barrière avec C. (B IV) Un troisième laser UV sonde la population restante du conformère A ou celle du conformère C. Lorsque  $\lambda_1$  est fixé sur une transition de A,  $\lambda_2$  est également fixe pour amener la molécule vers un niveau vibrationnel donné et  $\lambda_3$  est variable, on mesure l'isomérisation du conformère A vers les différents conformères. Si  $\lambda_1$  est fixé sur A et  $\lambda_2$  est variable, on mesure le changement de population de C en fixant  $\lambda_3$  sur sa transition. On mesure ainsi la hauteur de barrière entre A et C.

## Molécules neutres

**Spectroscopie électronique** Le groupe de Levy en 1985 a été le premier à appliquer la méthode d'ionisation résonnante à deux photons en jet supersonique pour mesurer les spectres d'excitation des acides aminés aromatiques isolés phénylalanine (Phe), tyrosine (Tyr) et tryptophane (Trp) [58, 80, 56]. Les spectres obtenus étaient très riches et présentaient de nombreuses structures dues à des bandes vibroniques et à la présence de plusieurs conformères. Le nombre de conformères a par la suite été déterminé grâce aux méthodes de saturation [54], de fluorescence dispersée [80] et de dépopulation UV-UV. Outre le nombre de conformères



présents dans le jet, cette dernière permet aussi de séparer leurs contributions spectrales, à condition que les bandes d'absorption ne se recouvrent pas [81]. Bien que très puissantes, ces techniques ne permettent cependant pas d'accéder à la structure des espèces observées. En appliquant une « variante » de la spectroscopie de dépopulation UV-UV à la spectroscopie IR de certains groupements fonctionnels tels que les vibreurs CH, NH et OH ( $2600-3800\text{ cm}^{-1}$ ), Page et collaborateurs ont pu accéder à la structure vibrationnelle de l'état fondamental  $S_0$  de la molécule de benzène et de son dimère de façon très sélective [82]. En effet, en fixant le faisceau de sonde UV sur une transition donnée du benzène et en faisant varier le laser de pompe IR, ils ont obtenu un signal équivalent au spectre d'absorption IR du benzène. L'utilisation d'une telle méthode, dite aussi de double résonance, permet de mesurer le spectre IR de chaque espèce. Le premier groupe à avoir appliqué cette technique à des acides aminés, Phe et Trp, est celui de John P. Simons à Oxford. En mesurant les fréquences des modes d'élongation NH et OH des groupes  $\text{NH}_2$  et  $\text{COOH}$ , le groupe a ainsi déterminé la structure des acides aminés [83, 84]. Nous avons utilisé ces deux dernières méthodes dans nos travaux ; elles seront exposées en détail dans le chapitre qui suit.

**Anion multipolaire** La technique d'anion multipolaire repose sur le transfert d'électron entre un atome de Rydberg possédant un électron faiblement lié et un complexe polaire - on parle de transfert d'électrons de Rydberg ou RET (Rydberg Electron Transfer) - couplé à la technique de spectroscopie IR [24]. L'interaction entre un électron de Rydberg et le moment dipolaire d'une molécule neutre mène à la formation d'un anion dipolaire, c'est-à-dire pour lequel l'électron est placé dans une orbitale très diffuse, Parce qu'il est éloigné de la molécule, sa présence ne modifie pas la géométrie du système neutre et on a donc accès à la structure du neutre, à condition que le moment dipolaire total  $\mu$  soit supérieur à  $2.2\text{ D}$  [85]. Cette technique ne nécessite pas la présence d'un chromophore UV. L'énergie fournie par le laser IR étant supérieure à l'énergie de liaison du complexe, il y a dissociation de celui-ci lorsque le laser IR est résonnant avec une vibration intramoléculaire du neutre. On mesure alors la diminution du signal d'ion à la masse du parent résultant de cette dissociation en fonction de la longueur d'onde du laser IR. L'énergie de liaison dépend du moment dipolaire total de la

molécule neutre et donc de sa conformation.

### Molécules ioniques

**Dissociation multiphotonique IR** La technique de dissociation multiphotonique IR (IRMPD pour IR Multiphoton Dissociation) fait appel à un processus de fragmentation qui résulte de l'absorption successive de photons IR suivie de la relaxation du système par IVR et lui permettant d'atteindre sa limite de dissociation. Couplée à une analyse en masse, cette technique permet de mesurer le signal d'apparition du fragment en fonction de l'énergie du photon IR. Elle est rendue possible grâce à la forte puissance crête du laser IR. Une telle puissance ne peut être délivrée que par des lasers à électrons libres, CLIO à Orsay et FELIX aux Pays-Bas.

Dans leurs travaux sur des complexes d'acides aminés Gly-Na<sup>+</sup> et Pro-Na<sup>+</sup>, G. Ohanessian et collaborateurs ont ainsi mis en évidence que l'acide aminé Pro prenait une forme zwitterionique stabilisée par la formation d'un pont salin entre le cation métallique et les deux oxygènes de l'acide carboxylique. En revanche, ils ont montré que l'acide aminé Gly adoptait une structure neutre à transfert de charge au sein du complexe [86]. Ces résultats sont conformes aux calculs de chimie quantique menés par le groupe. En étendant leurs travaux à des oligopeptides sodiés Gly<sub>n</sub>-Na<sup>+</sup> et Ala<sub>n</sub>-Na<sup>+</sup> (n = 1-4), l'attribution en termes de structures s'est révélée plus complexe, en l'occurrence à cause de la présence de plusieurs conformères [87]. Ils ont néanmoins pu attribuer les formes observées à des structures maximisant le nombre de coordination avec le cation Na<sup>+</sup>.

Bien que la méthode soit très prometteuse, les auteurs ont conclu qu'elle ne donnait pas accès à la structure précise d'un système mais permettait de distinguer des classes de structures. Ceci peut être lié à la température des ions formés qui conduit à des spectres complexes comportant les contributions des différentes formes peuplées, contrairement aux conditions de système isolé des jets supersoniques qui apportent une grande simplification spectrale. Par conséquent, la largeur spectrale du laser à électrons libres CLIO qui est de 5-10 cm<sup>-1</sup> dans le domaine étudié - modes d'élongation CO et OH et de déformation NH et CH (1200-1800 cm<sup>-1</sup>) - ne permet pas toujours de séparer les contributions spectrales de chacune de ces

formes lorsqu'il y a des recouvrements spectraux.

**Mobilité ionique** La technique de mobilité ionique mesure la vitesse à laquelle un ion se meut en présence d'un champ électrique faible dans un tube à dérive (« drift tube »), le principe est similaire à la chromatographie en phase gazeuse. Le temps mis pour traverser le tube donne accès à la section efficace de collision [88]. Les conformères très repliés, typiquement les conformations en hélice, possèdent une section efficace relativement faible et correspondront donc à un temps de migration faible, tandis que les structures non repliées, et donc de plus grande section efficace, mèneront à des temps de traversée plus élevés [43].

En réalisant des mesures à différentes températures, Jarrold et collaborateurs ont pu déterminer une température de transition entre la forme repliée du cytochrome c et sa forme dénaturée [89]. L'étude de chaînes de polyalanines substituées par un résidu lysine (Lys) qui possède une chaîne latérale protonée, a permis au même groupe de montrer que la présence et la position d'une charge positive affectaient la stabilité de l'hélice  $\alpha$  : celle-ci est stable lorsque la charge est placée en position C-terminale dans la séquence de polyalanines (à partir de  $n=10$ ) mais n'est plus un minimum lorsque le résidu Lys est en position N-terminale ou en milieu de chaîne ; la structure observée prend alors un aspect globulaire [90, 91]. Dans le cas de chaînes de polyglycines, seule la forme globulaire est formée [92]. La technique de mobilité ionique s'applique donc efficacement à des peptides de taille moyenne mais elle s'est avérée moins concluante pour de petits peptides ( $n < 10$ ). De plus, cette technique ne donne pas accès à la structure précise d'une espèce mais permet uniquement de différencier des classes de structures (repliées vs. non repliées).

### 1.3 Problématique

Nous avons cherché à observer en phase gazeuse les structures secondaires typiques de la biologie. L'objectif de cette thèse était en premier lieu de savoir si leur formation est une propriété intrinsèque de la chaîne peptidique, puis de déterminer les forces qui gouvernent leur formation. Nous avons alors orienté notre étude sur la caractérisation des liaisons hydrogène intramoléculaires, qui jouent un rôle fondamental dans la cohésion des structures secondaires.

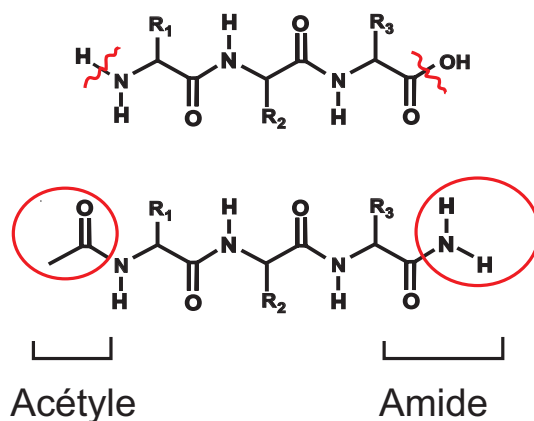


FIG. 1.16 – Peptide  $R_1$ - $R_2$ - $R_3$  pour lequel on a modifié les groupements terminaux  $NH_2$  et  $COOH$  pour donner le peptide protégé  $Ac$ - $R_1$ - $R_2$ - $R_3$ - $NH_2$ .

Nous avons appliqué les techniques de spectroscopie de double résonance IR-UV en phase gazeuse, associées à des calculs de chimie quantique, pour l'étude de petits peptides mimant des segments de protéines. Nous avons donc choisi d'étudier des peptides protégés dont les extrémités acide et basique ont été chimiquement modifiées afin qu'elles n'interfèrent pas dans le repliement de la chaîne. Les groupes  $NH_2$  et  $COOH$  des acides aminés naturels peuvent en effet interagir en formant des liaisons spécifiques qui n'existent pas au sein de la chaîne peptidique ; ils ont été substitués par un groupe acétyle ( $CH_3$ -CO) en position N-terminale et par un groupement amide  $NH_2$  en position C-terminale. Les protections assurent un nombre entier de liaisons peptidiques, les peptides protégés sont alors de la forme  $Ac$ -...- $NH_2$  (Figure 1.16). Selon que l'on fait référence au nombre de résidus ou de liaisons peptidiques, on parle de  $n$ -peptides ou de  $n+1$ -peptides : par exemple, un peptide protégé (chimiquement) qui comporte deux résidus, et donc trois liaisons peptidiques, est un *dipeptide* ou bien un modèle de *tripeptide*. Dans la suite du manuscrit, nous emploierons le premier terme *dipeptide* ou *dipeptide protégé* pour faire référence à un tel système.

L'emploi de la spectroscopie UV nous permettra d'accéder à la signature électronique des systèmes observés et pourra nous indiquer, par l'intermédiaire de la technique de dépopulation UV-UV, le nombre de conformères présents dans le jet. Leur caractérisation se poursuivra en appliquant les techniques de spectroscopie de double résonance IR-UV qui, couplées à une

sélectivité en masse, nous donneront accès à la structure des espèces formées de façon sélective, en nous renseignant notamment sur le réseau de liaisons hydrogène intramoléculaires au sein de la chaîne peptidique. Les groupements amides CO-NH des liaisons peptidiques nous serviront de sonde structurale en nous indiquant la présence éventuelle de liaisons hydrogène. Nous nous appuierons notamment sur les bandes amide A, qui correspondent aux vibrations d'élongation des groupements NH, et les bandes amide I et II qui sont associées à des modes de vibration couplés.

L'étude de ces systèmes par spectroscopie UV implique la présence d'un chromophore UV; notre choix s'est alors porté sur le résidu aromatique phénylalanine (Phe). Nous avons procédé par étapes en commençant par nous intéresser au comportement de l'acide aminé protégé Ac-Phe-NH<sub>2</sub> qui constituera notre système de référence. Nous avons ensuite étendu notre investigation à des peptides plus longs contenant deux et trois résidus, en conservant le résidu Phe dans la séquence. Nous verrons que la présence du résidu aromatique pourra affecter significativement la conformation de la chaîne peptidique en interagissant avec celle-ci. Nous reviendrons tout au long de la discussion sur les éventuels effets de la chaîne latérale.

Le chapitre qui suit sera consacré à la description du dispositif expérimental, des outils spectroscopiques et théoriques utilisés. Nous expliquerons ce qui a guidé notre choix pour la méthode théorique.

Le troisième chapitre exposera les résultats obtenus sur le système Ac-Phe-NH<sub>2</sub>. Ces résultats nous serviront de point de départ pour l'investigation de deux séries de dipeptides Ac-Phe-Xxx-NH<sub>2</sub> et Ac-Xxx-Phe-NH<sub>2</sub> qui seront également présentés dans ce chapitre. Nous montrerons notamment que la chaîne peptidique présente une organisation à une échelle locale s'articulant autour de chaque résidu dans la séquence.

Nous verrons ensuite qu'en raison de la flexibilité de la chaîne peptidique, les deux séries de dipeptides mèneront à d'autres structures qui s'avèreront différentes de celles observées dans le chapitre 3. Nous montrerons qu'il s'agit d'une des structures secondaires de la biologie et que sa formation est indépendante de la nature de la séquence.

Enfin, dans le chapitre 5, nous essaierons de voir comment la longueur de la chaîne peut affecter sa conformation et en particulier si les propriétés mises en évidence pour des séquences plus courtes sont toujours observées. Cette dernière partie nous permettra finalement de confirmer l'ubiquité des structures secondaires et en particulier de discuter de la stabilité des formes hélicoïdales.

**Références bibliographiques du chapitre 1**

- [1] L. Pauling, R. B. Corey and H. R. Branson, Proc. Natl. Acad. Sci. U.S. **37**, 205 (1951).
- [2] L. Pauling and R. B. Corey, Proc. Natl. Acad. Sci. U.S. **37**, 729 (1951).
- [3] J. C. Kendrew et al., Nature **181**, 662 (1958).
- [4] K. U. Linderstrøm-Lang and J. A. Shellman, *The Enzymes, Protein structure and enzyme activity* (Academic Press, New York, 1959), vol. 1, 2nd edn.
- [5] G. N. Ramachandran, C. Ramakrishnan and V. Sasisekharan, J. Mol. Biol. **7**, 95 (1963).
- [6] J. M. Berg, J. L. Tymoczko and L. Stryer, *Biochemistry 5th Edition* (W. H. Freeman and Company, New York, 2002).
- [7] C. M. Venkatachalam, Biopolymers **6**, 1425 (1968).
- [8] J. S. Richardson, Adv. Prot. Chem. **34**, 167 (1981).
- [9] K. Chou, Anal. Biochem. **286**, 1 (2000), et ref.citées.
- [10] K. Guruprasad and S. Rajkumar, J. Biosc. **25**, 143 (2000).
- [11] E. Vass, M. Hollósi, F. Besson and R. Buchet, Chem. Rev. **103**, 1917 (2003), et réf. citées.
- [12] E. Milner-White, B. R. Moss, R. Ismail, K. Belhjad-Mostefa and R. Poet, J. Mol. Biol. **204**, 777 (1988).
- [13] A. Litjas, Science **285**, 2077 (1999).
- [14] J. A. Smith and L. G. Pease, CRC Crit. Rev. Biochem. **9**, 315 (1980).
- [15] G. Rose, L. M. Gierasch and J. A. Smith, Adv. Prot. Chem. **37**, 1 (1985).
- [16] G. Boussard and M. J. Marraud, J. Am. Chem. Soc. **107**, 1825 (1985), et réf. citées.
- [17] G.-B. L. and C. J. Rito and S. H. Gellman, J. Am. Chem. Soc. **114**, 440 (1992), et réf. citées.
- [18] S. G. Stepanian, I. D. Reva, E. D. Radchenko, M. T. S. Rosado, M. L. T. S. Duarte, R. Fausto and L. Adamowicz, J. Phys. Chem. A **102**, 1041 (1998).
- [19] S. G. Stepanian, I. D. Reva, E. D. Radchenko and L. Adamowicz, J. Phys. Chem. A **102**, 4623 (1998).

- [20] S. G. Stepanian, I. D. Reva, E. D. Radchenko and L. Adamowicz, *J. Phys. Chem. A* **103**, 4404 (1999).
- [21] S. G. Stepanian, I. D. Reva, E. D. Radchenko and L. Adamowicz, *J. Phys. Chem. A* **105**, 10664 (2001).
- [22] R. M. Bentwood, A. J. Barnes and W. J. Orville-Thomas, *J. Mol. Spec.* **84**, 391 (1980).
- [23] Z. S. Huang and R. E. Miller, *J. Chem. Phys.* **91**, 6613 (1989).
- [24] B. Lucas, F. Lecomte, B. Reimann, H. D. Barth, G. Grégoire, Y. Bouteiller, J. P. Schermann and C. Desfrancois, *Phys. Chem. Chem. Phys.* **5**, 2600 (2004).
- [25] P. Hohenberg and W. Kohn, *Phys. Rev.* **864**, B136 (1964).
- [26] P. K. Weiner and P. A. Kollman, *J. Comput. Chem.* **2**, 287 (1991).
- [27] A.-M. Spase, L. Mallah-Levy, S. D. Daniels and B. W. Erickson, *J. Am. Chem. Soc.* **109**, 3526 (1987).
- [28] J. A. Bombasaro, M. A. Zamora, H. A. Baldoni and R. D. Enriz, *J. Am. Chem. Soc.* **98**, 5648 (2005).
- [29] A. Perczel and I. G. Csizmadia, *Int. Rev. Phys. Chem.* **14**, 127 (1995).
- [30] M. D. Beachy, D. Chasman, R. B. Murphy, T. A. Halgern and R. A. Friesner, *J. Am. Chem. Soc.* **119**, 5908 (1997), et réf. citées.
- [31] K. Möhle and H.-J. Hofmann, *J. Mol. Model.* **4**, 5 (1998), et réf. citées.
- [32] C.-H. Yu, L. Schafer and M. Ramek, *J. Phys. Chem. A* **103**, 8337 (1999).
- [33] G. A. Chasse, A. M. Rodriguez, M. L. Mak, E. Deretey, A. Perczel, C. P. Sosa, R. D. Enriz and I. G. Csizmadia, *J. Molec. Struct. (THEOCHEM)* **537**, 319 (2001), et réf. citées.
- [34] R. Vargas, J. Garza, B. P. Hay and D. Dixon, *J. Phys. Chem. A* **106**, 3213 (2002).
- [35] K. J. Jalkanen, M. Elstner and S. Suhai, *J. Mol. Struct. (Theochem)* **675**, 61 (20004), et réf. citées.
- [36] K. Möhle, M. Gussman, A. Rost, R. Cimiraglia and H.-J. Hofmann, *J. Phys. Chem. A* **101**, 8571 (1997).



- [37] M. Elstner, K. J. Jalkanen, M. Knapp-Mohammady, T. Frauenheim and S. Suhai, *Chem. Phys.* **256**, 15 (2000), et réf. citées.
- [38] R. Improta, V. Barone, K. N. Kudin and G. E. Scuseria, *J. Am. Chem. Soc.* **123**, 3311 (2001).
- [39] P. Bour, J. Kubelka and T. A. Keiderling, *Biopolymers* **65**, 45 (2002), et réf. citées.
- [40] R. Wiczorek and J. J. Dannenberg, *J. Am. Chem. Soc.* **125**, 14065 (2003), et réf. citées.
- [41] N. Kobko and J. J. Dannenberg, *J. Phys. Chem. A* **107**, 6688 (2003).
- [42] A. Perczel, I. Jákli and I. G. Csizmadia, *Chem. Eur. J.* **9**, 5332 (2004).
- [43] M. F. Jarrold, *Annu. Chem. Rev.* **32**, 360 (1999), et ref. citées.
- [44] C. Desfrancois, S. Carles and J. P. Schermann, *Chem. Rev.* **100**, 3943 (2000).
- [45] T. S. Zwier, *J. Phys. Chem. A* **105**, 8827 (2001), et ref. citées.
- [46] E. G. Robertson and J. P. Simons, *Phys. Chem. Chem. Phys.* **3**, 1 (2001).
- [47] R. Weinkauf, J. P. Schermann, M. S. de Vries and K. Kleinermanns, *Eur. Phys. J. D* **20**, 309 (2002), et ref. citées.
- [48] G. M. Florio, R. A. Christie, K. D. Jordan and T. S. Zwier, *J. Am. Chem. Soc.* **124**, 10236 (2002), et réf. citées.
- [49] F. O. Talbot and J. P. Simons, *Phys. Chem. Chem. Phys.* **4**, 3562 (2002).
- [50] R. A. Jockusch, R. T. Kroemer, F. O. Talbot and J. P. Simons, *J. Phys. Chem. A* **107**, 10725 (2003).
- [51] I. Hünig, A. J. Painter, R. A. Jockusch, P. Çarçalbal E. M. Marzluff, L. C. Snoek, D. P. Gamblin, B. G. Davis and J. P. Simons, *Phys. Chem. Chem. Phys.* **7**, 2474 (2005).
- [52] E. Nir, P. Imhof, K. Kleinermanns and M. S. de Vries, *J. Am. Chem. Soc.* **122**, 8091 (2000).
- [53] F. Piuze, M. Mons, I. Dimicoli, B. Tardivel and Q. Zhao, *Chem. Phys.* **270**, 205 (2001).
- [54] T. R. Rizzo, Y. D. Park, L. Peteanu and D. H. Levy, *J. Chem. Phys.* **84**, 2534 (1986).
- [55] T. R. Rizzo, Y. D. Park and D. H. Levy, *J. Chem. Phys.* **85**, 6945 (1986).

- [56] S. J. Martinez III, J. C. Alfano and D. H. Levy, *J. Mol. Spectrosc.* **156**, 421 (1992), et réf. citées.
- [57] R. Cable, M. J. Tubergen and D. H. Levy, *J. Am. Chem. Soc.* **109**, 6198 (1987).
- [58] R. Cable, M. J. Tubergen and D. H. Levy, *J. Am. Chem. Soc.* **110**, 7349 (1988).
- [59] R. Cohen, B. Brauer, E. Nir, L. Grace and M. S. de Vries, *J. Phys. Chem. A* **104**, 6351 (2000).
- [60] K. B. Shelimov, D. E. Clemmer, R. R. Hudgins and M. F. Jarrold, *J. Am. Chem. Soc.* **119**, 2240 (1997).
- [61] P. E. Barran, N. C. Polfer, D. J. Campopiano, D. J. Clarke, J. R. W. Govan, A. Maxwell, J. R. Dorin, R. P. Millar and M. T. Bowers, *Int. J. Mass Spectrom.* **240**, 273 (2005), et réf. citées.
- [62] J. Grotemeyer, U. Boesl, K. Walter and E. W. Schlag, *J. Am. Chem. Soc.* **108**, 4233 (1986).
- [63] A. Lindinger, J. P. Toennies and A. Vilesov, *J. Chem. Phys.* **110**, 1429 (1999).
- [64] M. Karas, D. Bachmann, U. Bahr and F. Hillenkamp, *Int. J. of Mass spectrometry and ion process* **78**, 53 (1987).
- [65] M. Karas and F. Hillenkamp, *Anal. Chem.* **60**, 2299 (1988).
- [66] J. B. Fenn, M. Mann, C. K. Meng, S. F. Mong and C. M. Whitehouse, *Science* **246**, 64 (1989).
- [67] G. von Helden, T. Wyttenbach and M. T. Bowers, *Science* **267**, 1483 (1995).
- [68] W. Caminati, G. Maccaferri, P. G. Favero and L. B. Favero, *Chem. Phys. Lett.* **265**, 561 (1997).
- [69] S. Melandri, M. E. Sanz, W. Caminati, G. Maccaferri, P. G. Favero and Z. Kiesel, *J. Am. Chem. Soc.* **120**, 11504 (1998).
- [70] P. D. Godfrey and R. D. Brown, *J. Am. Chem. Soc.* **117**, 2019 (1995).
- [71] P. D. Godfrey and R. D. Brown, *J. Am. Chem. Soc.* **120**, 10724 (1998).

- [72] A. Lesarri, E. J. Cocinero, J. C. López and J. L. Alonso, *Angew. Chem. Int. Ed.* **43**, 605 (2004).
- [73] A. Lesarri, E. J. Cocinero, J. C. López and J. L. Alonso, *J. Am. Chem. Soc.* **127**, 2572 (2005).
- [74] W. Caminati, *Phys. Chem. Chem. Phys.* **6**, 2806 (2004).
- [75] L. C. Snoek, R. T. Kroemer and J. P. Simons, *Phys. Chem. Chem. Phys.* **4**, 2130 (2002).
- [76] J. P. Simons, *C. R. Chimie* **6**, 17 (2003).
- [77] P. Çarçabal, L. C. Snoek, R. T. Kroemer, J. P. Simons, J. M. Bakker, I. Compagnon, G. Meijer and G. von Helden, *Phys. Chem. Chem. Phys.* **6**, 4546 (2004).
- [78] B. C. Dian, A. Longarte and T. S. Zwier, *Science* **296**, 2369 (2002), et ref. citées.
- [79] B. C. Dian, J. R. Clarkson and T. S. Zwier, *Science* **303**, 11969 (2004).
- [80] J. R. Cable, M. J. Tubergen and D. H. Levy, *J. Am. Chem. Soc.* **111**, 9032 (1989).
- [81] R. Lipert and S. Colson, *J. Chem. Phys.* **93**, 3894 (1989).
- [82] R. H. Page, Y. R. Shen and Y. L. Lee, *J. Chem. Phys.* **88**, 4621 (1988).
- [83] L. C. Snoek, E. G. R. R. T. Kroemer and J. P. Simons, *Chem. Phys. Lett.* **321**, 49 (2000).
- [84] L. C. Snoek, R. T. Kroemer, M. R. Hockridge and J. P. Simons, *Phys. Chem. Chem. Phys.* **3**, 1819 (2001).
- [85] C. Desfrancois, H. Abdoul-Carime, N. Khelifa, and J. P. Schermann, *Phys. Rev. Lett.* **73**, 2436 (1994).
- [86] C. Kapota, J. Lemaire, P. Maître and G. Ohanessian, *J. Am. Chem. Soc.* **126**, 1836 (2004).
- [87] C. Kapota, *Interactions du cation sodium avec des molécules d'intérêt biologique : acides aminés et oligopeptides* (Thèse de doctorat, 2005).
- [88] E. A. Mason and E. W. McDaniel, *Transport Properties of ions in Gases* (Wiley, New York, 1988).
- [89] R. R. Hudgins, J. Woenckhaus and M. F. Jarrold, *Int. J. Mass Spectrom. Ion Processes* **165/166**, 497 (1997), et ref. citées.

- [90] M. Kohtani, J. E. Schneider, T. C. Jones and M. F. Jarrold, *J. Am. Chem. Soc.* **126**, 16981 (2004), et réf. citées.
- [91] M. Kohtani, T. C. Jones, J. E. Schneider and M. F. Jarrold, *J. Am. Chem. Soc.* **126**, 7420 (2004), et réf. citées.
- [92] R. R. Hudgins and M. F. Jarrold, *J. Phys. Chem. B* **104**, 2154 (2000).



## Chapitre 2

# Méthodologies

Nous présenterons dans ce chapitre les outils dont nous avons disposé pour l'étude des peptides protégés. Nous avons fait appel à deux approches, expérimentale et théorique, afin d'aboutir à des interprétations les plus affinées possibles.

La stratégie expérimentale repose sur les techniques de spectroscopie de double résonance qui, associées à l'emploi d'un jet supersonique, permettent d'obtenir des informations très précises sur des systèmes de taille croissante. Néanmoins, les systèmes que nous cherchons à caractériser sont, nous le verrons par la suite, des systèmes très complexes. Aussi, malgré la précision des informations collectées, nous verrons qu'il est parfois nécessaire de recourir à des outils théoriques pour caractériser les espèces observées. La première partie de ce chapitre sera consacrée à la description du montage expérimental, dans la seconde partie seront exposées les techniques spectroscopiques que nous avons appliquées dans nos travaux et nous développerons dans une dernière partie les méthodes de calcul employées dans cette étude.

### 2.1 Dispositif expérimental

L'étude expérimentale consiste à produire les molécules, les exciter optiquement par des sources lasers et détecter les ions qui en résultent par un spectromètre de masse à temps de vol. Les différentes parties du montage seront présentées dans les paragraphes qui suivent.

### 2.1.1 Préparation des molécules

La source de vaporisation, conçue au laboratoire [1] repose sur la technique de désorption laser couplée à une détente supersonique.

#### Détente supersonique

L'utilisation d'une détente supersonique permet de créer des conditions de « système isolé » froid conduisant à une simplification des spectres [2]. La température très basse dans le jet permet en outre la formation de complexes faiblement liés tels que les complexes de van der Waals ou les agrégats moléculaires. Le principe est fondé sur la différence de pression entre deux chambres séparées par un orifice (ou tuyère) de très faible diamètre. Dans notre cas, la vanne pulsée joue le rôle de tuyère. Le gaz porteur, à une pression initiale  $P_0$  de quelques bars dans une bouteille à haute pression, subit une détente adiabatique à travers la tuyère dans une enceinte maintenue sous pression réduite  $P_1$  de l'ordre de  $10^{-4}$  mbar (chambre d'expansion). À l'entrée de l'enceinte, les collisions entre le gaz porteur et les molécules entraînent une conversion de l'énergie interne rovibrationnelle en énergie de translation. Il s'ensuit que seuls les niveaux rotationnels et vibrationnels les plus bas sont peuplés. On obtient ainsi un jet monocinétique et directionnel. Le quantum de vibration étant plus grand que le quantum de rotation, le refroidissement vibrationnel est généralement moins efficace que le refroidissement rotationnel. Les températures rotationnelles et vibrationnelles obtenues dans le jet sont typiquement inférieures à 10 K et 50 K respectivement. Le refroidissement des molécules dépend de  $P_0$  qui doit être élevée afin d'observer une densité d'espèces suffisante pour les expériences en phase gazeuse et de la nature du gaz porteur. Nous avons ainsi choisi d'utiliser l'argon comme gaz porteur (la vitesse de jet est de 560 m/s) afin d'assurer un refroidissement collisionnel efficace, bien qu'il favorise la formation d'agrégats en raison de sa plus grande polarisabilité. Néanmoins, cette caractéristique de l'argon à former des complexes pourrait par la suite être utilisée dans l'étude de complexes pour nos systèmes.

La détente est engendrée à l'aide d'une vanne pulsée General Valve (orifice de 300 ou 500  $\mu\text{m}$ , cadence de 10 Hz, durée d'ouverture de l'ordre de 100  $\mu\text{s}$ ). Une détente impulsionnelle permet d'obtenir un vide résiduel suffisant pour la détente, en maintenant un système de

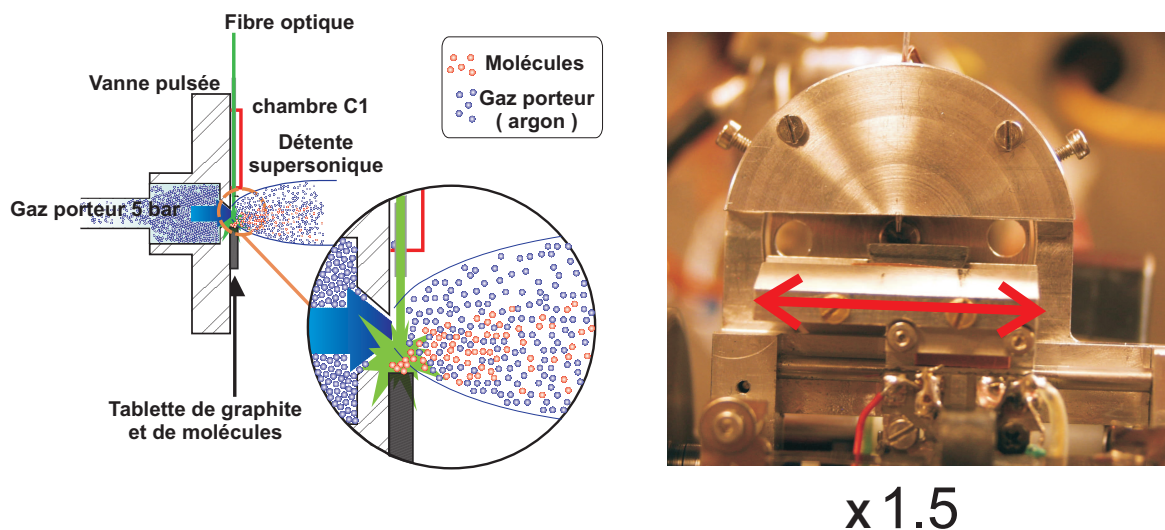


FIG. 2.1 – Schéma de la désorption laser.

pompage relativement modeste. Elle est par ailleurs compatible avec les sources lasers qui sont également impulsionnelles.

L'utilisation d'un spectromètre de masse à temps de vol nécessite la présence d'un écorceur entre l'enceinte du jet et la chambre d'ionisation afin de réduire la pression dans la chambre du spectromètre. L'écorceur permet d'extraire la partie centrale du jet et ainsi de sélectionner les molécules les plus froides.

### Désorption laser

Il n'est pas aisé de garder les molécules biologiques intactes lors de leur mise en phase gazeuse. Deux types de méthodes sont employées pour la mise en phase gazeuse de molécules neutres, le chauffage thermique et la désorption laser. La première consiste à placer les molécules dans un four chauffé à l'aide de cordons chauffants (Thermocoax) que l'on règle à l'aide d'une mesure en température par thermocouple. Dans un montage de jet moléculaire, ce four est placé en amont de la vanne qui doit également être chauffée pour éviter les phénomènes de condensation. Les températures de chauffage dépendent du système étudié. Outre la limitation due à la plage de fonctionnement en température de la vanne, la méthode par chauffage n'est pas toujours adaptée à certaines molécules biologiques qui se dégradent sous l'effet de la chaleur : c'est par exemple le cas d'une des bases de l'ADN, la guanine [3, 4].



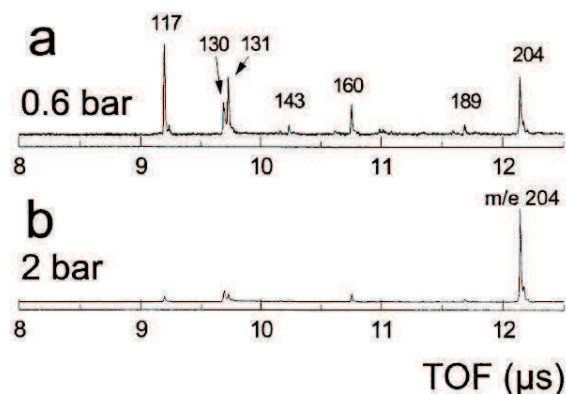


Fig. 2. REMPI mass spectra of desorbed tryptophan ( $m/e$  204) obtained with desorption laser energy of 2.3 mJ at excitation wavelength of 270 nm under different argon pressure: (a) 0.6 bar and (b) 2 bar.

FIG. 2.2 – Spectre de masse de l'acide aminé tryptophane (Trp) obtenu par excitation laser, à partir de pastilles de Trp, pour 2 valeurs de pression  $P_0$  [1]. Les différentes masses correspondent à des produits de dégradation lorsque le refroidissement est insuffisant ; c'est par exemple le cas avec une pression  $P_0$  de 0.6 bar.

La vaporisation par désorption laser a pour avantage de causer un faible taux de dégradation lorsque que le refroidissement est suffisant, comme l'ont montré les travaux de Piuzzi *et al.* sur l'acide aminé tryptophane [1]. Dans le dispositif, les molécules sont vaporisées en aval de la tuyère. Le refroidissement se fait en *post-expansion* et conduit à une zone de refroidissement plus petite que dans une détente en *co-expansion* pour les méthodes traditionnelles de chauffage, où les molécules sont refroidies dès la sortie de la vanne. Le principe de la vaporisation par désorption repose sur l'absorption de la lumière laser par une matrice absorbante : l'échauffement rapide et important a lieu sur une très faible profondeur, conduisant à une vaporisation « instantanée » de la matrice et des molécules qu'elle contient. L'échantillon se présente sous la forme d'une pastille de 6 ou 13 mm de diamètre selon la quantité de produits dont nous disposons. Elle est obtenue en pressant dans une presse hydraulique (pression de 5 tonnes typiquement) un mélange de poudre de graphite et de peptides réduits en poudre dans des proportions molaires de 1 : 4 (1 molécule pour 4 atomes de carbone). Le choix de la matrice s'est porté sur le graphite car il ne réagit pas avec les systèmes étudiés ; il est apolaire et est facile d'utilisation.

Le faisceau d'un laser Nd : YAG (diamètre de 2 mm) doublé en fréquence (532 nm vert) est

introduit dans une fibre optique (diamètre de 400  $\mu\text{m}$ , 5 m de longueur, 70 % de transmission) à l'aide d'un coupleur optique. La distance entre la fibre et l'échantillon est inférieure à 2 mm, la distance entre l'échantillon et la vanne est d'environ 1 mm (Figure ??). Des énergies de 1-3 mJ par impulsion laser (correspondant à une densité d'énergie de 0.2-0.7 J/cm<sup>2</sup>) sont utilisées pour désorber les espèces. L'échantillon présente une surface plane et rectangulaire de dimension typique 2mm \* 6mm ou 2mm \* 13mm. Il est monté sur un chariot permettant sa translation dans une direction parallèle à sa longueur afin que des tirs laser consécutifs ne frappent pas la même surface. En pratique, la pastille est limée afin d'avoir une surface assez grande pour effectuer un « balayage » avec le laser de désorption. La partie fibre optique-échantillon-vanne est placée sur un dispositif de translation XYZ réglable, de sorte qu'il est possible d'optimiser la quantité de molécules collectées en ajustant la position du jet par rapport à l'écorceur et à l'axe du spectromètre. La deuxième harmonique (532 nm) du laser Nd :YAG de désorption vient frapper la matrice de carbone. En s'échauffant, le carbone transfère son énergie aux molécules. Les molécules éjectées sont alors entraînées par le jet supersonique à travers l'écorceur. L'utilisation du rayonnement à 532 nm (rayonnement vert visible) permet de régler aisément la position de la fibre par rapport à l'orifice de la vanne et à la pastille. Outre la minimisation de la dégradation des molécules lors de leur vaporisation, la désorption laser offre également l'avantage de pouvoir travailler avec de faibles quantités de produits, contrairement à la méthode par chauffage. Celles-ci sont typiquement de l'ordre de quelques dizaines de mg (30 à 150) quand les techniques par chauffage réclament des grammes.

### 2.1.2 Spectromètre de masse

L'utilisation d'un spectromètre de masse permet d'obtenir une détection sélective en masse, ce qui est nécessaire lorsque l'on souhaite s'assurer de la nature des molécules étudiées quand le produit de synthèse n'est pas purifié, ou par la suite mener des études en agrégats. Le spectromètre de masse à temps de vol est équipé d'un miroir électrostatique et peut être utilisé en mode linéaire ou en mode réflectron. Il est composé de trois zones principales, une zone d'extraction, une zone d'accélération et une zone sans champ, dite de vol libre.

Le jet supersonique, dont l'axe est parallèle à celui du spectromètre, entraîne les molécules

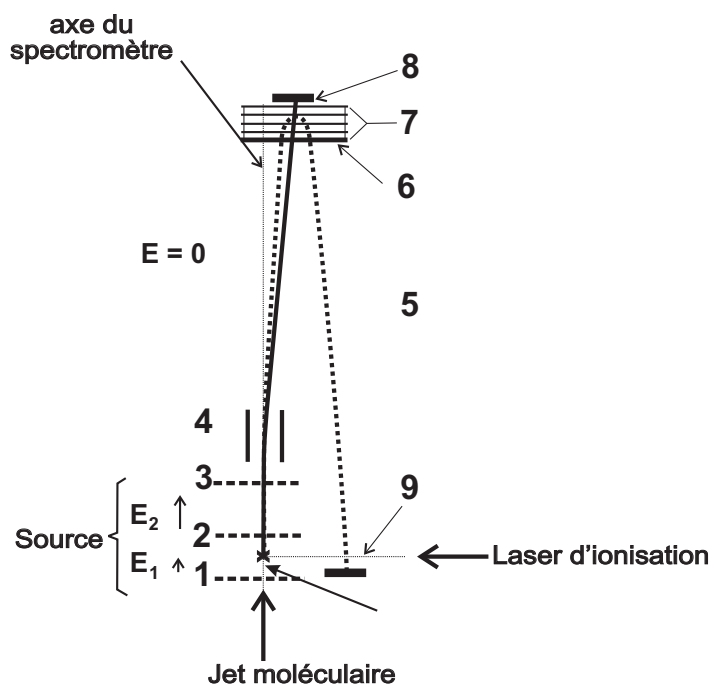


FIG. 2.3 – Schéma du spectromètre de masse à temps de vol. **1** : Grille "repousseur" d'ions ( $V = 1283$  V). **2** : Grille extractrice ( $V = 1200$  V). **3** : Grille de sortie de source, à la masse. **4** : Plaques de déflexion. **5** : Zone de vol libre. **6** : Zone de freinage du miroir. **7** : Zone de réflexion des ions. **8** : Galettes de micro-canaux pour la détection des ions en mode linéaire. **9** : Galettes de micro-canaux pour la détection en mode réflectron. Les traits pleins indiquent la trajectoire des ions en mode linéaire, les pointillés en mode réflectron.

vers la zone d'extraction : elles y sont excitées puis ionisées selon un processus d'ionisation résonnante à deux photons par un faisceau laser entrant perpendiculairement à l'axe du spectromètre. Un champ uniforme  $E_1$  amène les ions formés vers la zone d'accélération. Ceux-ci sont alors soumis à un autre champ uniforme  $E_2$  avant de traverser la zone de temps de vol (Figure 2.3). Le temps que met un ion à parcourir la distance entre la zone d'extraction et le détecteur définit le temps de vol de l'ion (exprimé en  $\mu s$ ) suivant la relation :

$$T_i = a\sqrt{\frac{m}{q}} + b \quad (2.1)$$

où  $a$  et  $b$  sont des constantes qui dépendent du spectromètre ( $b$  est une correction due à l'électronique de synchronisation),  $m$  et  $q$  la masse (exprimée en u.m.a.) et la charge de l'ion. Le temps de vol d'un ion est donc proportionnel à sa masse et à sa charge. En raison de leur vitesse initiale non nulle, les ions formés ont une distribution d'énergie cinétique, c'est-

à-dire que les ions de même masse n'ont pas tous la même vitesse initiale. L'effet de cette dispersion des vitesses est corrigé lorsque les réglages de l'appareil répondent aux conditions de focalisation.

En mode linéaire, le spectromètre de masse est de type Wiley et McLaren [5]. Les ions traversent la zone sans champ et sont détectés par des galettes de micro-canaux placées à l'extrémité de cette zone. Les galettes produisent un flux d'électrons proportionnel au courant d'ions induit et au gain du détecteur. En sortie de la source, deux plaques de déflexion permettent d'ajuster la trajectoire des ions transversalement au jet pour les guider vers le détecteur qui est légèrement décalé de l'axe du jet afin de ne pas être pollué par celui-ci (Figure 2.3). La tension appliquée aux plaques doit être d'autant plus faible que la taille des espèces considérées augmente. La résolution du spectromètre dépend des paramètres géométriques et électriques de l'appareil et est optimale lorsque les conditions de focalisation de Wiley-McLaren [5] sont réalisées. La résolution  $m/\Delta m$ , définie comme le rapport entre la masse associée à un pic et l'écart de masse correspondant à la largeur à mi-hauteur du pic, est de l'ordre de 150 pour la masse  $m = 450$  u.m.a.

L'utilisation du spectromètre en mode réflectron permet d'améliorer la résolution de l'appareil en allongeant le temps de vol parcouru et en corrigeant la dispersion d'énergie cinétique initiale des ions. La résolution est alors de l'ordre de 2000 pour la masse  $m = 450$  u.m.a., conformément aux conditions de focalisation de Mamyrin [6, 7]. En sortie de source, les ions traversent la zone sans champ, puis sont freinés et réfléchis dans le miroir électrostatique situé à l'extrémité du spectromètre (Figure 2.3). En raison de la distribution d'énergie cinétique, les ions les plus rapides pénètrent plus loin dans le miroir que les plus lents. Mais après avoir parcouru une seconde fois la zone de vol libre, tous les ions sont finalement focalisés temporellement sur les galettes de microcanaux.

### 2.1.3 Systèmes lasers

Le système laser engendre un rayonnement impulsionnel et accordable en fréquence. Il est constitué d'un laser à colorant (Lambda Physik, FL 2002 et FL 3000) pompé par un laser à excimère XeCl (Lambda Physik EMG 103) pour l'excitation et l'ionisation des molécules dans

l'UV, et d'un oscillateur paramétrique optique (OPO, Euroscan) pour la spectroscopie IR des modes d'élongation NH (domaine spectral 3  $\mu\text{m}$ ). Les mesures de spectres électroniques et vibrationnels dans la région de 3  $\mu\text{m}$  ont été menées dans le laboratoire, à Saclay. La spectroscopie IR des modes d'élongation CO et des modes de déformation NH (domaine spectral 5-8  $\mu\text{m}$ ) a été réalisée avec le laser à électrons libres FELIX à Rijnhuizen (Pays-Bas).

### Lasers UV

Le principe du laser à excimère repose sur l'excitation d'un mélange constitué de gaz rare et d'halogène provoquée par une décharge électrique et entraînant la formation d'un complexe stable à l'état excité ou excimère. L'émission laser a lieu à partir de la désexcitation radiative de l'excimère. Le mélange utilisé est composé de Xe, HCl et d'un gaz tampon (He ou He/Ne). Le complexe XeCl\* a une longueur d'onde fixe d'émission de 308 nm. Le rayonnement UV résultant a une largeur temporelle de l'ordre de 15 ns. L'énergie par impulsion est d'environ 150 mJ.

En sortie du laser à excimère, le faisceau UV vient pomper le laser à colorant qui engendre alors un rayonnement accordable (Figure 2.4). Une première partie du faisceau UV passe par la cavité composée d'un réseau, qui assure l'accordabilité du laser, d'un miroir, ainsi qu'une cuve oscillatrice (et pré-amplificatrice) contenant le colorant. Il passe ensuite dans une cuve d'amplification contenant le même colorant que la cuve oscillatrice. Le faisceau amplifié est alors doublé en fréquence par un cristal non linéaire biréfringent de BBO (ou KDP thermostaté) qui permet d'étendre le domaine spectral vers l'UV. Un système de quatre prismes de Pélin-Broca permet de séparer les composantes UV et visible du rayonnement. Le faisceau UV en sortie du laser à colorant a une durée d'impulsion de l'ordre de 10 ns et une largeur spectrale de 0.2  $\text{cm}^{-1}$ . Les énergies par impulsion sont de l'ordre de 10 mJ pour le faisceau visible et de 700  $\mu\text{J}$  pour le faisceau UV.

Les systèmes étudiés ont pour chromophore un cycle phényle. Nous avons donc utilisé le rayonnement doublé en fréquence de la Coumarine 540A, dont le maximum d'émission se situe à 540 nm, qui permet d'accéder à la bande d'absorption du phényle dans le proche UV (266 nm).

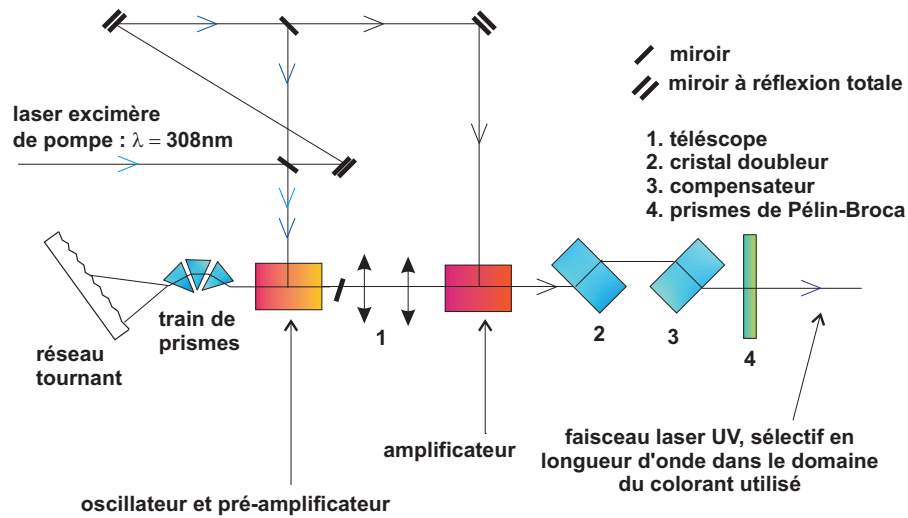


FIG. 2.4 – Schéma montrant la chaîne laser UV.

### Oscillateur paramétrique optique OPO

La chaîne laser OPO est composée d'un laser de pompe Nd : YAG et d'un oscillateur paramétrique optique. La longueur d'onde 1064 nm du laser YAG est utilisée pour pomper l'OPO, la cadence de répétition est de 10 Hz, l'énergie par impulsion utilisée est de 150 mJ et la finesse spectrale est de  $1\text{ cm}^{-1}$ .

L'oscillation paramétrique optique est un phénomène non linéaire qui a lieu au sein d'une cavité optique contenant un cristal non linéaire  $\text{LiNbO}_3$ . Une onde incidente de fréquence  $\omega_p$  et de vecteur d'onde  $\vec{k}_p$  pénètre dans le cristal non linéaire et engendre deux ondes dites signal ( $\vec{k}_s, \omega_s$ ) et complémentaire (ou « idler ») ( $\vec{k}_c, \omega_c$ ). Ces trois ondes satisfont aux conditions de conservation de l'énergie  $\omega_p = \omega_s + \omega_c$ , et d'accord de phase  $\vec{k}_p = \vec{k}_s + \vec{k}_c$ . L'accordabilité est obtenue en variant la longueur d'onde pour laquelle est obtenu l'accord de phase. Cela consiste à ajuster la position du cristal en changeant l'angle entre l'onde incidente et l'axe du cristal. Le cristal  $\text{LiNbO}_3$  utilisé permet de couvrir une région allant de  $2500$  à  $4000\text{ cm}^{-1}$  (ou  $4.0$ - $2.5\ \mu\text{m}$ ). La largeur spectrale de l'onde complémentaire est de l'ordre de  $20\text{ cm}^{-1}$ . On ajoute alors une lame étalon de Fabry-Pérot, de distance interfrange (ou intervalle spectral libre) de  $32\text{ cm}^{-1}$ , dans la cavité afin d'améliorer la résolution du laser IR. On ajuste le maximum de la courbe de gain du cristal avec l'un des maxima du système de franges de la lame étalon, dont l'intervalle spectral libre est plus grand que la largeur de la courbe de gain du cristal.

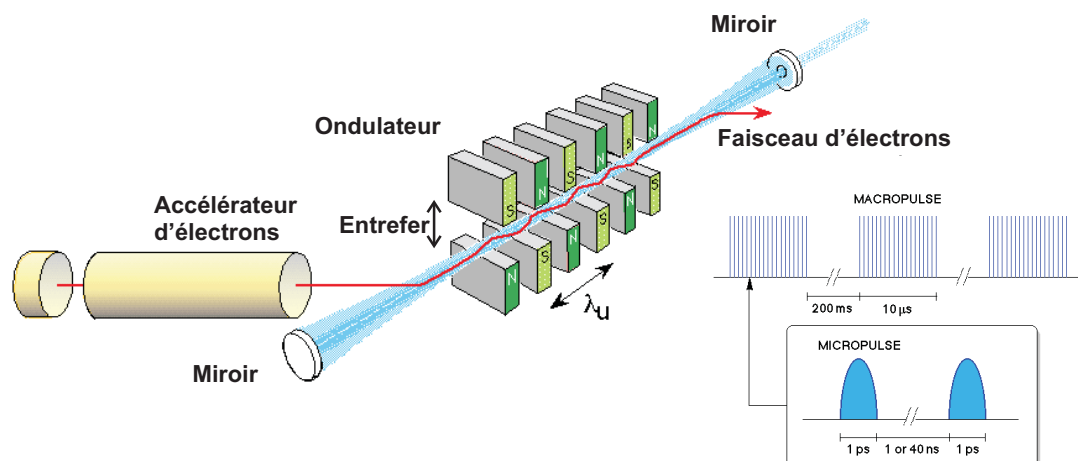


FIG. 2.5 – Schéma du principe et structure temporelle du laser à électrons libres FELIX.

On atteint une largeur spectrale intrinsèque inférieure à  $0.6 \text{ cm}^{-1}$ , néanmoins limitée par celle de l'onde pompe qui est de  $1 \text{ cm}^{-1}$  (laser Nd :YAG non injecté). Les énergies par impulsion obtenues sont, dans le cas présent, de 1-3 mJ dans la région spectrale  $2.5\text{-}3.3 \mu\text{m}$  (NH stretch), hormis dans la région  $2.85\text{-}2.89 \mu\text{m}$  pour laquelle l'intensité IR chute fortement en raison d'une absorption importante dans le cristal de l'OPO.

### Laser à électrons libres FELIX

La spectroscopie dans les régions amide I et amide II ( $1500\text{-}1800 \text{ cm}^{-1}$ ), qui sont, respectivement, principalement composées des modes d'élongation de la liaison CO et des modes de déformation dans le plan CNH (voir section 2.2) de nos systèmes a été réalisée à l'aide du laser à électrons libres FELIX [8], en collaboration avec l'équipe de G. Meijer au Fundamental Optik Material (FOM) Institute for Plasma Physics à Rijnhuizen (Pays-Bas). Le dispositif est constitué d'une source laser UV pour l'excitation et l'ionisation des molécules dans l'UV, et de FELIX pour la spectroscopie IR.

Le laser à électrons libres FELIX (Free Electron Laser Infrared eXperiment) est une source accordable en énergie conçue pour accéder au moyen et lointain IR, en délivrant une fluence suffisamment élevée pour réaliser l'étude spectroscopique de systèmes en milieu dilué comme c'est le cas des études en phase gazeuse. Le principe du laser à électrons libres repose sur le rayonnement émis par des électrons énergétiques en mouvement non rectiligne uniforme. Un

paquet d'électrons relativistes, produits par un accélérateur radiofréquence, est injecté dans un ensemble composé d'aimants permanents (onduleur) situé à l'intérieur de la cavité optique. Le champ magnétique dans l'onduleur est perpendiculaire à la direction des électrons et est modulé spatialement, entraînant une déflexion périodique des électrons et un rayonnement émis par ceux-ci lorsqu'ils traversent l'onduleur. Le rayonnement est amplifié à chaque passage par le faisceau d'électrons dans l'onduleur. Les oscillations qu'ils effectuent au sein de la cavité permettent de les mettre en phase et ainsi d'amplifier l'intensité du rayonnement. La longueur d'onde du rayonnement dépend du champ magnétique à l'intérieur de l'onduleur. Elle est balayable par ajustement de la distance de l'entrefer.

La structure temporelle du laser IR consiste en un train de micro-impulsions formant des macro-impulsions d'une durée de 5  $\mu\text{s}$  espacées de 100 ms (taux de répétition de 10 Hz) dans nos expériences. La durée d'une micro-impulsion est comprise entre 300 fs et quelques ps, les micro-impulsions sont espacées de 1 à 40 ns (taux de répétition de 1 GHz à 25 MHz). La largeur spectrale à mi-hauteur est de l'ordre de 1%, soit d'environ 15  $\text{cm}^{-1}$  dans le domaine étudié (1500-1800  $\text{cm}^{-1}$ ).

#### 2.1.4 Synchronisation

Le fonctionnement en régime impulsionnel des sources lasers exige une bonne synchronisation temporelle de tous les éléments du montage. L'utilisation d'un générateur d'impulsions retardées permet de contrôler le déclenchement des différents éléments (vanne, laser de désorption, laser UV, etc.).

La commande d'ouverture de la vanne est le temps de référence ( $t = 0$ ). L'impulsion gazeuse n'apparaît dans la chambre qu'au bout d'un temps de 1.2 ms qui correspond à l'ouverture réelle de la vanne. On déclenche alors le laser de désorption ( $t_2$ ). Les molécules extraites sont entraînées par le jet supersonique. Lorsque celles-ci atteignent la zone d'interaction avec le laser d'ionisation, on déclenche ce dernier ( $t_3$ ). Le déclenchement de l'OPO est fonction de celui du laser UV d'ionisation. Il s'effectue une centaine de ns avant le laser UV.

À chaque tir du laser de désorption, le faisceau ne frappe pas la même surface de l'échantillon. Toutefois, en raison de l'usure de la tablette d'échantillon, les conditions de désorption peuvent



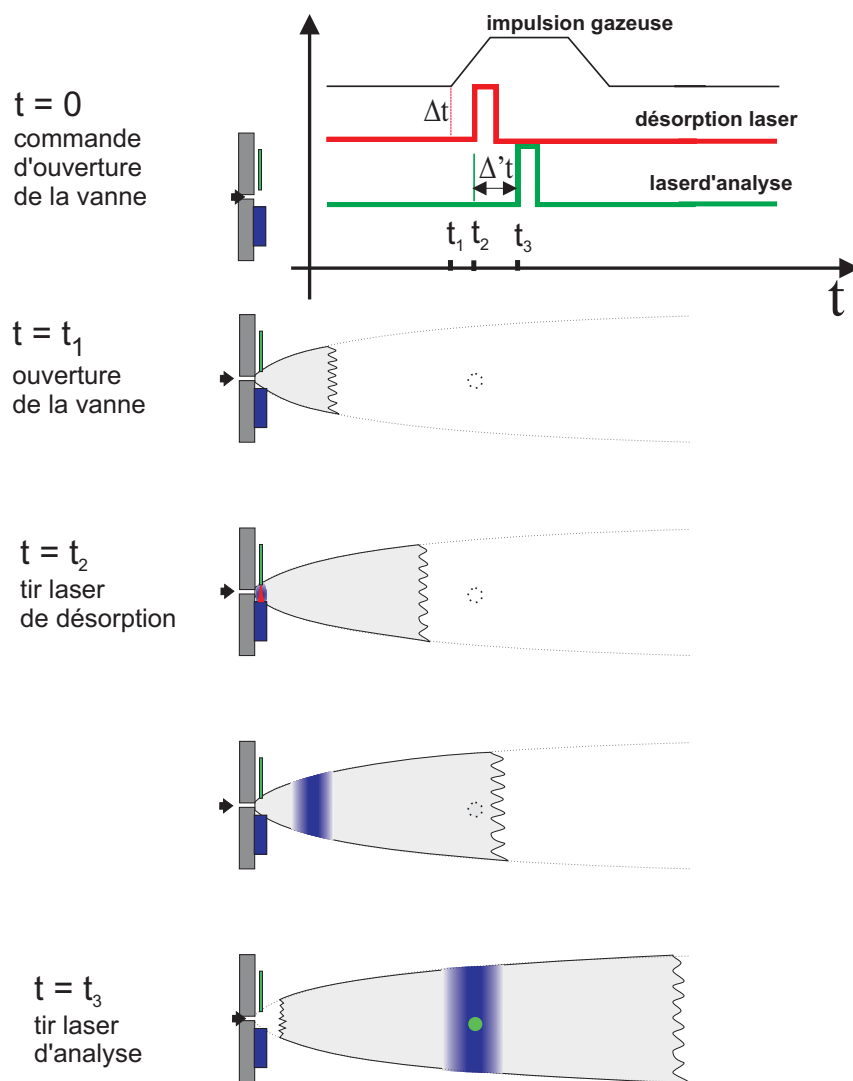


FIG. 2.6 – Schéma du chronogramme de l'expérience.

varier durant l'enregistrement d'un spectre.

Le paramètre  $\Delta t = t_2 - t_1$  entre le déclenchement du laser de désorption et l'ouverture de la vanne permet de contrôler le moment où on ensemence le jet. Ceci permet en l'occurrence de se déplacer dans l'impulsion gazeuse afin de sélectionner une zone du jet pour laquelle les molécules seront froides. Le délai  $\Delta't = t_3 - t_2$  entre le déclenchement du laser d'ionisation et le laser de désorption correspond au temps de trajet des molécules entre le moment où elles sont vaporisées et celui où elles atteignent la zone d'interaction avec le laser d'analyse. Ce délai permet de sélectionner les molécules interrogées par le laser d'analyse.

Les collisions multiples sont nombreuses au voisinage de la tuyère, leur nombre décroît ensuite rapidement avec la distance par rapport à la tuyère. Il est alors possible de sélectionner un système en fonction de sa taille, ou plus précisément de sa stoechiométrie en sondant une zone plus ou moins froide du jet. Ainsi les monomères froids sont-ils présents dès le début de la détente, tandis que des complexes, dimères, trimères, etc, dont la formation nécessite plus de collisions, seront obtenus pour des délais plus longs dans l'impulsion gazeuse. Le paramètre  $\Delta t$  qui contrôle la zone excitée dans l'impulsion gazeuse donne accès à la stoechiométrie du système : un monomère aura donc un délai plus petit qu'un dimère, qu'un trimère, etc. Nous nous sommes donc assurés que les systèmes étudiés étaient bien des monomères.

## 2.2 Méthodes spectroscopiques

Les outils spectroscopiques que nous avons utilisés reposent sur la spectroscopie UV, qui permet d'accéder à la signature électronique d'un système, et sur la spectroscopie IR qui nous donnera accès à la structure des espèces formées.

La signature électronique d'un système traduit la perturbation induite par l'environnement sur le chromophore UV. Dans notre cas, le résidu phénylalanine (Phe), dont la chaîne latérale R est un cycle phényle, joue le rôle de chromophore, et la perturbation est induite par les interactions entre le cycle aromatique et les groupements CO-NH des liaisons peptidiques. Ces perturbations induisent un décalage suffisamment grand entre les transitions électroniques  $S_0-S_1$  pour distinguer différentes conformations de la chaîne. Il s'ensuit que des environnements

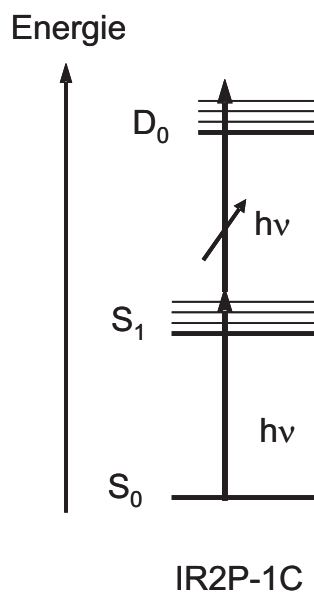


FIG. 2.7 – Schéma du principe d'ionisation multiphotonique à une couleur.

différents autour du chromophore aromatique conduiront à des spectroscopies électroniques relativement différentes en terme de domaine spectral et d'activité Franck-Condon.

### 2.2.1 Ionisation résonnante à 2 photons (IR2P)

L'énergie nécessaire pour ioniser la plupart des molécules aromatiques se situe entre 7 et 9 eV, ce qui nécessite l'absorption d'au moins deux photons laser. Dans le processus d'ionisation résonnante à deux photons (IR2P), le système est porté dans un état excité intermédiaire par un premier photon ( $h\nu_1$ ), c'est-à-dire le premier état électronique excité  $S_1$  dans notre cas, puis est ionisé par l'absorption du deuxième photon ( $h\nu_2$ ) (Figure 2.7). Le processus IR2P permet d'accéder à la spectroscopie d'un système dans le proche UV. Parmi les schémas d'ionisation multiphotonique possibles, les photons peuvent provenir du même laser ou de sources laser différentes. Dans notre cas, les deux photons sont issus de la même impulsion laser, on parle alors de processus d'ionisation à une couleur (IR2P-1C). Le passage par un état intermédiaire résonnant, un état vibronique de  $S_1$  dans notre cas, apporte une sélectivité en espèce puisque seules les espèces excitées par le premier photon sont ionisées et analysées. Les ions sont détectés par le spectromètre de masse à temps de vol. En collectant les ions correspondant à un temps de vol défini en fonction de la longueur d'onde d'excitation ( $h\nu_1$ ), on obtient un

spectre d'excitation en ionisation résolu en masse. On accède à la structure électronique du système en faisant balayer l'énergie d'excitation du premier photon autour de la bande origine  $0_0^0$  du toluène ( $37477 \text{ cm}^{-1}$  [9]) qui est le chromophore présent dans le résidu phénylalanine. Le spectre obtenu est équivalent au spectre d'absorption si la durée de vie de l'état  $S_1$  est suffisamment longue pour absorber le deuxième photon.

### 2.2.2 Spectroscopie de dépopulation UV-UV (Hole-Burning)

La méthode de dépopulation est une technique pompe-sonde, employant deux lasers. Elle est couramment employée pour déterminer le nombre de conformères présents dans un spectre [10]. Elle permet en particulier de séparer la contribution spectrale de chaque espèce en présence, à condition toutefois que les transitions des différentes espèces ne se recouvrent pas.

Le laser de pompe  $h\nu_p$ , très intense, balaie le domaine spectral autour de la transition électronique  $S_1 \leftarrow S_0$  d'une espèce A. Après un délai  $\Delta t$  d'une centaine de ns, le laser de sonde  $h\nu_s$ , issu d'un second laser, fixé sur la transition électronique de l'espèce A permet l'ionisation de celle-ci. Le signal d'ions correspondant, c'est-à-dire le signal de sonde, mesure la population de l'espèce A dans son état fondamental. On enregistre par conséquent un signal constant qui correspond à l'excitation de la transition  $S_0$ - $S_1$  de l'espèce. Lorsque le laser de pompe est en résonance avec une bande de l'espèce sondée A, le niveau fondamental est dépeuplé entraînant une diminution (ou déplétion) du signal d'ions (Figure 2.8a). Ces déplétions signifient que les bandes correspondantes balayées par le laser de pompe appartiennent à A. Si en revanche, le laser de pompe est résonnant avec une transition de l'espèce B, le signal de sonde, toujours fixé sur A, reste constant (Figure 2.8b).

### 2.2.3 Spectroscopie de dépopulation IR-UV

La spectroscopie IR donne accès à la structure vibrationnelle d'un système. En particulier, les groupements amide CO-NH des liaisons peptidiques sont très sensibles à la présence de liaisons hydrogène. Les signatures IR de certains groupements fonctionnels vont donc nous

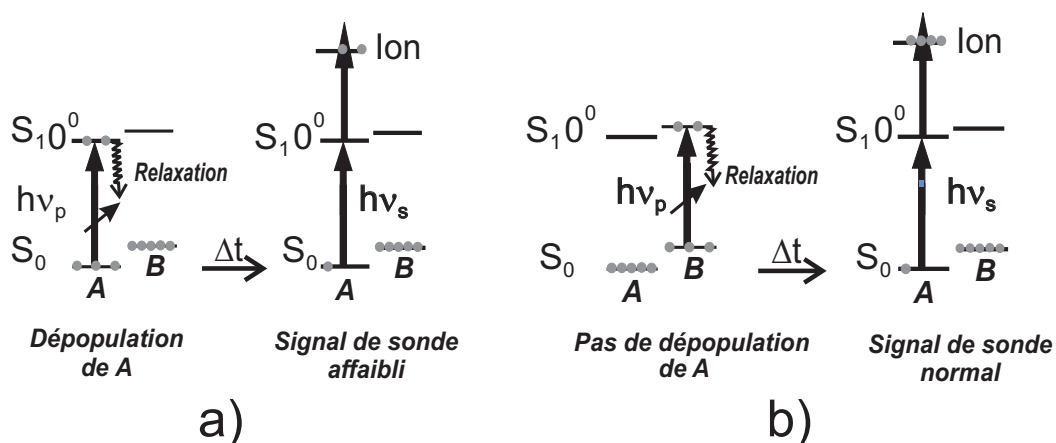


FIG. 2.8 – Schéma de la spectroscopie de dépopulation UV-UV.

permettre de caractériser la conformation de la chaîne peptidique en termes de liaisons hydrogène. On distingue ainsi trois principaux modes de vibration du groupement amide CO-NH :

- le mode amide I ( $1660\text{-}1800\text{ cm}^{-1}$ ) qui correspond principalement à l'élongation (stretching CO) de la liaison CO couplée aux modes d'élongation de la liaison CN et de déformation (pliage, bending NH) dans le plan CNH.
- le mode amide II ( $1480\text{-}1600\text{ cm}^{-1}$ ) qui correspond principalement au mode de déformation dans le plan CNH couplé aux modes d'élongation CO et CN et au mode de déformation CH.
- le mode amide A ( $3200\text{-}3600\text{ cm}^{-1}$ ) qui correspond à l'élongation de la liaison NH.

On constate donc que seuls les modes de vibration amide A sont relativement peu couplés entre eux, et aux autres oscillateurs de la molécule. L'attribution à partir des bandes d'élongation NH donnera donc des informations structurales plus directes.

La fréquence d'élongation NH d'un groupement NH libre apparaît typiquement à une valeur proche de  $3470\text{ cm}^{-1}$  [11]. Ainsi, toute bande décalée vers les basses fréquences (on dira dans la suite du manuscrit « décalées vers le rouge ») sera considérée comme la signature d'un groupement NH perturbé, en l'occurrence par une liaison hydrogène. Par exemple, une liaison hydrogène N-H...O conduit à un décalage typique de  $100\text{-}200\text{ cm}^{-1}$  [11]. De même, les modes amide I voient leur fréquence abaissée lorsque les groupes CO correspondants sont impliqués dans des liaisons hydrogène. Plus la liaison hydrogène est forte, plus le décalage mesuré est

grand et l'on observera une bande élargie et très intense. À l'inverse, les modes amide II seront décalés vers les hautes fréquences.

La technique de spectroscopie de double résonance IR/UV permet de mesurer le spectre IR d'une espèce sélectionnée par son spectre UV. Cette technique est comparable à la technique de dépopulation UV/UV décrite précédemment : le laser de pompe est alors le laser IR et le faisceau de sonde reste le laser UV. La mesure du signal d'ions en fonction de la longueur d'onde du laser IR permet donc d'avoir un spectre de dépopulation en IR. A chaque pas du laser IR, des signaux avec (« IR on ») et sans (« IR off ») laser IR sont enregistrés. La différence « IR on »/« IR off » est ensuite normalisée par le signal « IR off » afin de s'affranchir des fluctuations dues aux conditions de désorption. L'enregistrement du signal « IR off » est assuré par un obturateur « coupant » le faisceau IR et contrôlé par un ordinateur. Dans la suite du manuscrit, les spectres IR ne seront pas présentés sous la forme de déplétion mais sous la forme de bandes d'absorption.

### 2.3 Méthodes de calculs

La partie théorique de cette thèse a été menée en collaboration avec le Groupe de Chimie Théorique du SPAM sous l'encadrement de Jean-Pierre Dognon. Elle repose sur des calculs de fréquences harmoniques et d'énergies relatives. La comparaison entre les spectres IR expérimentaux et théoriques nous donnera des informations concernant la structure des espèces observées en indiquant la présence éventuelle de liaisons hydrogène et leur signature. En outre, par relaxation conformationnelle dans la détente, on s'attend à ce que seuls les minimas de la surface d'énergie potentielle soient peuplés. Les calculs d'énergie nous permettront alors d'identifier les conformations les plus stables. Réciproquement, la confrontation des calculs avec l'expérience sera un moyen d'évaluer leur précision et notamment de comparer différents niveaux de calculs.

Nous avons mené une exploration préliminaire de la surface d'énergie potentielle pour les systèmes Ac-Phe-NH<sub>2</sub>, Ac-Phe-Gly-NH<sub>2</sub> et Ac-Gly-Phe-NH<sub>2</sub> en utilisant le champ de force

Amber [12], à l'aide du logiciel Hyperchem 6.0 [13]. Les structures les plus stables dans un intervalle de 3 kcal/mol ont alors été réoptimisées par des méthodes de chimie quantique avec le logiciel Gaussian 98 [14]. Toutefois, bien que le champ de force Amber soit couramment employé dans l'étude des protéines, nous avons dû générer manuellement certaines structures de départ qui n'étaient pas trouvées par le champ de force telles que les structures secondaires de type coude  $\beta$ . Ces structures ont été construites en prenant les valeurs idéales des biochimistes pour les angles  $\phi$  et  $\psi$ . Aucune contrainte géométrique n'a été appliquée lors de ces calculs. Les calculs de fréquences ont ensuite été réalisés sur ces structures optimisées.

### 2.3.1 Choix de la méthode

Nous avons choisi de recourir à la théorie de la fonctionnelle de la densité (DFT pour density functional theory) pour mener notre étude théorique [15, 16]. Cette méthode, où l'énergie et en fait toutes les propriétés d'un système dans son état fondamental non dégénéré sont des fonctionnelles de la densité électronique de celui-ci, est désormais couramment utilisée pour étudier des interactions non covalentes telles que les interactions de van der Waals, les liaisons hydrogène, etc. [17, 18] intervenant dans la structure de molécules biologiques. Les calculs en DFT présentent plusieurs avantages, notamment un temps de calcul qui reste acceptable. Celui-ci varie en  $N^3$  où  $N$  est la taille de la base d'orbitales atomiques, alors qu'il varie respectivement en  $N^4$  et  $N^5$  pour les calculs effectués au niveau Hartree-Fock et par la théorie des perturbations MP2 (Møller-Plesset [19] au second ordre). En outre, contrairement aux méthodes Hartree-Fock, une partie de la corrélation électronique est prise en compte par la DFT. D'après les équations de Kohn-Sham, l'énergie peut s'exprimer comme la somme d'un terme d'énergie cinétique, d'un terme d'interaction électron-noyau, d'un terme d'interaction coulombienne et d'un terme d'échange-corrélation. Le principal inconvénient des calculs en DFT réside dans le fait que l'expression exacte de ce dernier terme n'est pas connue. Cette énergie d'échange-corrélation est alors calculée au moyen de diverses fonctionnelles approchées qui diffèrent par leurs termes d'échange et de corrélation.

Par ordre de qualité croissante, on trouve d'abord des fonctionnelles qui dépendent uniquement de la densité électronique (LDA pour Local Density Approximation). Viennent ensuite des

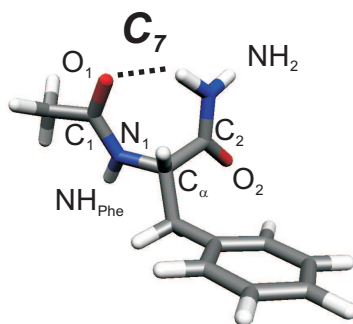


FIG. 2.9 – Conformation  $\gamma_L(a)$  du système Ac-Phe-NH<sub>2</sub> optimisée avec la méthode B3LYP/6-31G(d,p).

fonctionnelles qui dépendent de la densité et du gradient de la densité (GGA pour Generalized Gradient Approximation) ou bien du laplacien de la densité (meta GGA), etc. Il est aussi d'usage d'employer des fonctionnelles hybrides dont le terme d'échange est de type Hartree-Fock, c'est le cas de la fonctionnelle B3LYP [20, 21]. Dans ce cas, le temps de calcul varie alors en  $N^4$ .

Il nous a donc fallu choisir une méthode de calcul qui sera par la suite appliquée à tous nos systèmes. Nous avons alors testé plusieurs fonctionnelles et plusieurs bases. La structure de référence sur laquelle ont été effectués tous les calculs est une conformation repliée  $\gamma_L(a)$ , équivalente à un coude  $\gamma$ , qui présente une liaison hydrogène C<sub>7</sub> fermant un cycle à sept atomes (Figure 2.9). Le terme (a) fait référence à l'orientation du cycle aromatique par rapport au groupement NH de la liaison peptidique, cette orientation correspond à un angle dièdre  $\chi_1$  d'environ 180°. Les fréquences harmoniques ont été calculées pour chaque fonctionnelle avec la base 6-31G(d,p), à partir des géométries optimisées avec la même méthode. Elles ont alors été comparées avec le spectre IR dans la région des modes d'élongation NH du conformère C de l'acide aminé protégé Ac-Phe-NH<sub>2</sub> (Figure 2.10). La fonctionnelle et la base retenues sont celles qui donnent le meilleur accord avec l'expérience.

### Choix de la fonctionnelle

Les fonctionnelles testées sont la fonctionnelle hybride B3LYP [20, 21] et les fonctionnelles GGA BLYP [22, 20], BPW91 [22, 23], BP86 [22, 24], PW91PW91 [23]. Il est courant d'appliquer des facteurs d'échelle aux fréquences calculées afin de tenir compte du caractère



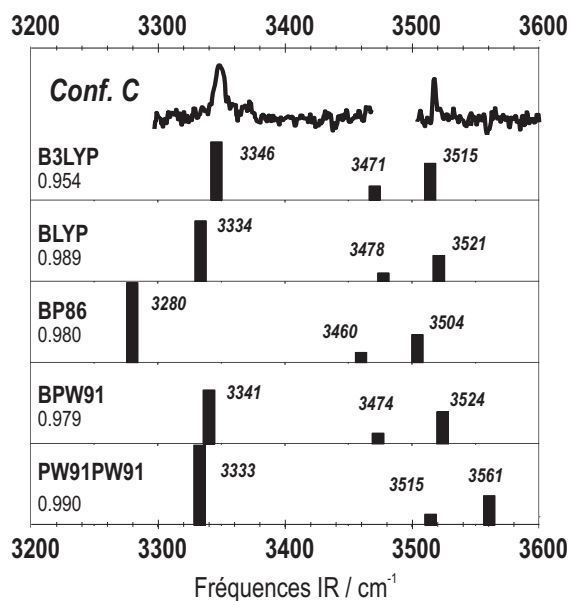


FIG. 2.10 – Comparaison entre le spectre IR expérimental du conformère C du système Ac-Phe-NH<sub>2</sub> et les fréquences harmoniques calculées avec diverses fonctionnelles pour la structure  $\gamma_L(a)$ . Les fréquences sont corrigées avec différents facteurs pour tenir compte de l'anharmonicité et de la taille de la base 6-31G(d,p).

anharmonique du potentiel et du défaut inhérent à la taille de la base utilisée [25]. Ces facteurs sont différents selon la fonctionnelle employée. Nous les avons déterminés en ajustant les fréquences théoriques sur le spectre IR du conformère C. La fréquence la plus basse et la plus haute correspondent aux modes d'élongation symétrique et antisymétrique du groupement NH<sub>2</sub> ( $\nu_{sym}$  et  $\nu_{anti}$ ). Les valeurs de  $\nu_{sym}$ , toutes inférieures à 3400 cm<sup>-1</sup>, indiquent bien que le groupe NH<sub>2</sub> est impliqué dans une liaison hydrogène forte. La troisième fréquence est associée au groupement amide NH<sub>Phe</sub> qui est libre (Figure 2.10). Les fréquences calculées avec les fonctionnelles BPW91, BLYP et B3LYP donnent un très bon accord avec le spectre IR du conformère C. Les autres méthodes reproduisent moins bien les valeurs expérimentales, elles surestiment notamment l'écart entre les composantes symétrique et antisymétrique du groupement NH<sub>2</sub>. La fonctionnelle B3LYP semble en fait donner les meilleurs résultats puisque l'écart théorique est identique à la valeur expérimentale (169 cm<sup>-1</sup>). Enfin, nous avons aussi comparé les paramètres géométriques de ces différentes structures avec une structure  $\gamma_L(a)$  optimisée au niveau MP2/6-31G(d,p) (Tableau 2.1). La structure calculée avec la fonctionnelle B3LYP est celle qui donne les paramètres les plus proches de celle obtenue au niveau MP2,

	BLYP	B3LYP	BP86	BPW91	PW91PW91	MP2
Distances (Å)						
C <sub>1</sub> -N <sub>1</sub>	1.375	1.362	1.370	1.371	1.367	1.363
N <sub>1</sub> -C $\alpha$	1.481	1.468	1.473	1.471	1.468	1.459
C $\alpha$ -C <sub>2</sub>	1.568	1.551	1.561	1.559	1.558	1.535
C <sub>2</sub> -N <sub>2</sub>	1.371	1.360	1.365	1.365	1.361	1.360
C <sub>1</sub> =O <sub>1</sub>	1.247	1.233	1.246	1.246	1.243	1.239
C <sub>2</sub> =O <sub>2</sub>	1.238	1.225	1.237	1.237	1.235	1.233
O $\cdots$ H	2.058	2.068	1.993	2.113	1.985	2.116
Angles (°)						
$\angle$ COH	106.469	106.9	106.3	106.3	106.4	107.2
$\angle$ NHO	142.8	141.3	144.4	140.1	144.2	139.2
Dihèdres (°)						
$\phi$	-81.5	-81.9	-81.5	-83.24	-81.7	-82.5
$\psi$	78.7	79.5	77.0	79.8	76.5	84.5

TAB. 2.1 – Paramètres géométriques du coude  $\gamma_L(a)$  calculé avec différentes fonctionnelles et au niveau MP2 avec la base 6-31G(d,p).

bien que le coude  $\gamma$  soit légèrement plus déformé pour la géométrie MP2, conduisant à une liaison hydrogène C<sub>7</sub> plus faible. Nous avons par conséquent choisi de réaliser tous nos calculs avec la fonctionnelle B3LYP.

### Choix de la base

Nous avons ensuite testé plusieurs bases afin d'évaluer l'influence de la taille de la base. Les bases utilisées sont les bases double-zeta de Pople 6-31G(d), 6-31+G(d), 6-31G(d,p), 6-311G(d,p), 6-31++G(d,p) [26]. Les fréquences ont également été corrigées avec un facteur d'échelle, que nous avons évalué, pour tenir compte de l'anharmonicité. La comparaison entre les calculs et le spectre IR du conformère C montre que les différentes bases donnent des valeurs de fréquences relativement comparables (Figure 2.11). Celles-ci diffèrent peu entre les bases moyennes et les grosses bases, de même que les paramètres géométriques semblent être peu influencés par la taille de la base (Tableau 2.2). Finalement, nous avons retenu la base 6-31+G(d) qui offre un bon compromis entre la précision et le temps de calcul, ce dernier paramètre étant essentiel puisque l'on souhaite étendre les calculs à des systèmes plus grands tels que les dipeptides et les tripeptides protégés.

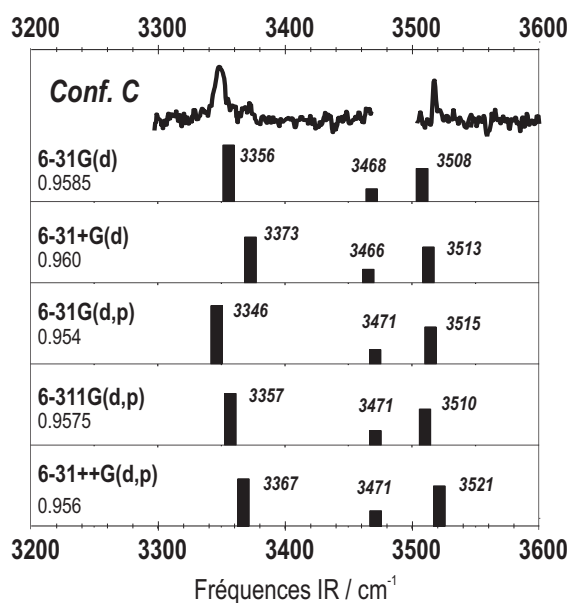


FIG. 2.11 – Comparaison entre le spectre IR expérimental du conformère C du système Ac-Phe-NH<sub>2</sub> et les fréquences harmoniques de la structure  $\gamma_L(a)$  calculées avec la fonctionnelle B3LYP avec différentes bases. Les fréquences sont corrigées avec différents facteurs pour tenir compte de l'anharmonicité et de la taille de la base.

	6-31G(d)	6-31+G(d)	6-31G(d,p)	6-31++G(d,p)	6-311G(d,p)
Distances (Å)					
C <sub>1</sub> -N <sub>1</sub>	1.362	1.364	1.362	1.364	1.361
N <sub>1</sub> -C $\alpha$	1.468	1.468	1.467	1.467	1.467
C $\alpha$ -C <sub>2</sub>	1.551	1.550	1.551	1.551	1.551
C <sub>1</sub> -N <sub>1</sub>	1.360	1.360	1.358	1.359	1.357
C <sub>1</sub> =O <sub>1</sub>	1.233	1.253	1.233	1.253	1.226
C <sub>2</sub> =O <sub>2</sub>	1.225	1.227	1.226	1.228	1.219
O...H	2.068	2.156	2.067	2.155	2.097
Angles (°)					
$\angle$ OCH	106.9	107.3	106.8	107.1	107.0
$\angle$ NHO	141.3	137.8	141.2	137.7	139.7
Dihèdres (°)					
$\phi$	-81.9	-82.9	-82.3	-83.4	-82.7
$\psi$	79.5	81.8	78.7	80.8	79.0

TAB. 2.2 – Paramètres géométriques du coude  $\gamma_L(a)$  calculés pour différentes bases avec la fonctionnelle B3LYP.

Nous verrons dans la suite du manuscrit que nous avons mesuré les spectres IR dans trois domaines différents, le domaine d'absorption amide A qui correspond aux modes d'élongation (stretching) NH, le domaine amide I qui correspond aux modes d'élongation CO couplés aux modes de pliage (bending ou déformation dans le plan CNH) de la liaison NH, et le domaine amide II correspondant aux modes de bending NH couplés aux modes de stretching CO et de bending CH. Ces trois modes de vibration étant de nature différente, nous avons alors appliqué des facteurs d'échelle à chaque type de vibration. Ceux-ci ont été ajustés à partir de structures susceptibles de correspondre aux trois conformères A, B et C observés pour le système Ac-Phe-NH<sub>2</sub>, à savoir les formes repliées  $\gamma_L$  avec les orientations (g+) et (g-) du phényle - les notations correspondent respectivement aux angles  $\chi_1 \sim +60^\circ$  et  $\chi_1 \sim -60^\circ$  - et une forme étendue  $\beta_L$ , équivalente au brin  $\beta$ , avec une orientation (a) du phényle et ne présentant pas de liaison hydrogène (la discussion détaillée des structures est présentée dans le chapitre 3). Le meilleur accord a été trouvé pour les facteurs suivants : 0.960 dans le domaine amide A, 0.980 dans le domaine amide I et 0.970 dans le domaine amide II.

Nous avons ensuite étendu les calculs aux dipeptides et tripeptides. Néanmoins notre objectif n'était pas d'explorer de façon exhaustive le paysage conformationnel de chaque système mais de proposer l'attribution la plus précise possible pour les conformères observés. Deux approches différentes ont déterminé le choix des structures à calculer. La première a consisté à considérer les géométries trouvées comme étant les plus stables pour l'acide aminé isolé, c'est-à-dire les formes qu'adopte le résidu de façon spontanée, et à les juxtaposer.

Dans la seconde approche, nous avons considéré les structures secondaires typiques de la biologie telles que les formes complètement étendues brins  $\beta$ , les formes repliées coudes  $\beta$  ou les hélices  $3_{10}$ .

Bien que les calculs effectués avec la fonctionnelle B3LYP donnent des géométries correctes, notamment pour des systèmes présentant des liaisons hydrogène, ils ne tiennent pas compte des effets de dispersion [27, 28, 29, 17]. Les énergies relatives ont donc toutes été calculées au niveau MP2/6-31+G(d) à partir des géométries optimisées avec la méthode B3LYP/6-

31+G(d). Nous verrons dans la suite du manuscrit que l'ordre énergétique est alors modifié suivant que les énergies sont calculées en DFT ou au niveau MP2.

**Références bibliographiques du chapitre 2**

- [1] F. Piuzzi, I. Dimicoli, M. Mons, B. Tardivel and Q. Zhao, *Chem. Phys. Lett.* **320**, 282 (2000).
- [2] R. Campargue, *Rarefield Gas Dynamics* (Academic Press, New York, 1966), vol. II.
- [3] E. Nir, L. Grace, B. Brauer and M. S. de Vries, *J. Am. Chem. Soc.* **121**, 4896 (1999).
- [4] M. Mons, I. Dimicoli, F. Piuzzi, B. Tardivel and M. Elhanine, *J. Phys. Chem. A* **106**, 5088 (2002).
- [5] W. C. Wiley and I. H. McLaren, *Rev. Sci. Instrum.* **26**, 1150 (1955).
- [6] V. I. Karatov, B. A. Mamyryn and D. V. Schmikk, *Sov. Phys.-JETP* **16**, 1177 (1972).
- [7] B. A. Mamyryn, V. I. Karatov, D. V. Schmikk and V. A. Zagulin, *Sov. Phys.-JETP* **37**, 45 (1973).
- [8] J. M. Bakker, L. M. Aleese, G. Meijer and G. von Helden, *Phys. Rev. Lett.* **91**, 203003 (2003).
- [9] M. Schauer, K. Law and E. Bernstein, *J. Chem. Phys.* **81**, 49 (1981).
- [10] R. Lipert and S. Colson, *J. Chem. Phys.* **93**, 3894 (1989).
- [11] E. G. Robertson, M. R. Hockridge, P. D. Jelfs and J. P. Simons, *Phys. Chem. Chem. Phys.* **3**, 786 (2001).
- [12] W. D. Cornell, P. Cieplack, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.* **117**, 5179 (1995).
- [13] H. . ., Hypercube, Inc., 115 NW 4th street, Gainesville, FL 32601 (USA).
- [14] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. J. A. Montgomery, T. Vreven, K. N. Kudin, J. C. Burant, J. M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala,

- K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, Gaussian 98, Revision A.11.3, Gaussian, Inc., Pittsburgh PA, 2002.
- [15] P. Hohenberg and W. Kohn, Phys. Rev. **864**, B136 (1964).
- [16] W. Kohn and L. J. Sham, Phys. Rev. **1133**, A140 (1965).
- [17] Y. Zhao and D. G. Truhlar, J. Chem. Theory Comput. **1**, 415 (2005), et réf. citées.
- [18] R. A. Friesner, Proc. Natl. Acad. Sci. U.S. **102**, 6648 (2005), et réf. citées.
- [19] C. M. Iler and M. Plesset, Phys. Rev. **46**, 618 (1934).
- [20] C. Lee, W. Yang and R. G. Parr, Phys. Rev. B **37**, 785 (1988).
- [21] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).
- [22] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
- [23] P. Perdew, *Electronic structure of solids* (Akademie Verlag, Berlin, 1991).
- [24] J. P. Perdew, Phys. Rev. B **33**, 8822 (1986).
- [25] M. P. Andersson and P. Uvdal, J. Phys. Chem. A **109**, 2937 (2005), et ref. citées.
- [26] W. J. Hehre, L. Radom, P. v. R. Schleyer and J. A. Pople, *Ab Initio Molecular Orbital Theory* (Wiley, New York, 1999).
- [27] J. Sadlej, R. Moszynski, C. J. Dobrowolski and A. P. Mazurek, J. Phys. Chem. A **103**, 8528 (1999).
- [28] A. D. Rabuck and G. E. Scuseria, Theor. Chem. Acc. **104**, 439 (2000).
- [29] A. K. Rappé and E. R. Bernstein, J. Phys. Chem. A **104**, 6117 (2000).

## Chapitre 3

# Préférences conformationnelles locales

La préférence conformationnelle locale est la conformation qu'adopte l'acide aminé protégé isolé. Au sein de la chaîne peptidique, cette conformation peut aussi correspondre à celle que prend un petit segment de la chaîne présentant un résidu unique. De nombreux calculs de chimie quantique indiquent que cette forme locale, c'est-à-dire la forme la plus stable, est une structure compacte et repliée notée  $\gamma$  [1, 2, 3, 4, 5, 6, 7]. Une structure non repliée est également identifiée par les calculs ; il s'agit d'une conformation complètement étendue nommée  $\beta_L$ . Celle-ci est néanmoins trouvée plus haute en énergie que la structure repliée. La structure étendue  $\beta_L$  est équivalente au brin  $\beta$  dont l'association avec un autre brin mène à la formation de l'une des structures secondaires les plus fréquentes dans les protéines, le feuillet  $\beta$ .

La plus grande stabilité de la conformation  $\gamma$  est due à la présence d'une liaison hydrogène intramoléculaire au sein de la chaîne principale qui assure sa forme repliée. Cette structure est en fait équivalente à un coude  $\gamma$  qui est la plus petite structure pouvant être responsable du changement de direction de la chaîne peptidique [8] ; on peut aussi le retrouver à l'intérieur de feuillets  $\beta$  antiparallèles. Le coude  $\gamma$  est constitué de trois résidus reliés entre eux par deux liaisons peptidiques. La liaison hydrogène intramoléculaire relie les groupes CO(i) et NH(i+2) dans un cycle à sept atomes, on parle alors de liaison hydrogène C<sub>7</sub> et de conformation C<sub>7</sub>. Il existe deux formes  $\gamma$  dont les squelettes peptidiques, c'est-à-dire sans la chaîne latérale, sont images dans un miroir l'un de l'autre : la forme  $\gamma_D$  dite normale ou classique et la forme  $\gamma_L$  dite

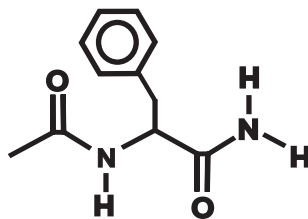
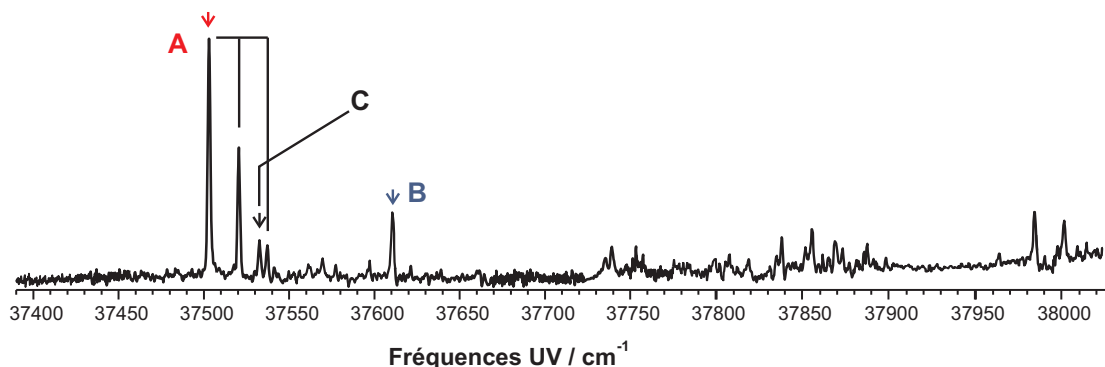


inverse. La structure  $\gamma_L$  est la forme la plus courante dans les protéines et est calculée comme étant la plus stable. La configuration  $\gamma_D$  est beaucoup moins stable en raison d'interactions de contact défavorables entre la chaîne latérale et le squelette peptidique [8, 9].

Ainsi, confirmant les nombreux calculs de structure indiquant que le coude  $\gamma$  est la forme la plus stable, de récents travaux en phase gazeuse par spectroscopie micro-onde sur l'acide aminé Ac-Ala-NHMe et par IR sur l'acide aminé Z-Pro-NHMe (la protection N-terminale consiste en un cycle benzyloxycarbone  $C_6H_5-C=O$ ) ont montré que ces systèmes adoptaient une conformation repliée  $\gamma$  [10, 11]. Pourtant, d'autres études théoriques ont montré que la forme étendue  $\beta_L$  pouvait être le minimum le plus stable [12, 13, 14, 15]. Ces résultats théoriques ont été confortés par les travaux de Zwier et collaborateurs qui ont en effet observé des formes  $\beta_L$  pour les systèmes Ac-Trp-NH<sub>2</sub> et Ac-Trp-NHMe [16]. Afin de comprendre pourquoi ces différentes études aboutissent à des conclusions divergentes, nous nous sommes intéressés à la structure que pouvait adopter un acide aminé protégé. La présence d'un chromophore aromatique est indispensable à notre étude puisque nous avons recours à la spectroscopie UV pour caractériser nos systèmes. Parmi les trois acides aminés aromatiques phénylalanine (Phe), tyrosine (Tyr) et tryptophane (Trp), nous avons choisi d'étudier l'acide aminé Phe. Nous avons donc commencé nos travaux sur le système protégé Ac-Phe-NH<sub>2</sub>. Nous verrons que les résultats obtenus se sont révélés très intéressants car ils nous ont permis d'affiner notre compréhension de la spectroscopie de ce système. Ils nous permettront par la suite d'étendre notre étude à des systèmes plus complexes.

### 3.1 Ac-Phe-NH<sub>2</sub> : recherche des formes typiques de la biologie

Les techniques de spectroscopie de double résonance IR-UV, désormais couramment utilisées afin de déterminer la structure d'un système en phase gazeuse [17], ont été appliquées au système Ac-Phe-NH<sub>2</sub> (Figure 3.1). Couplées à une analyse en masse, elles offrent une sélectivité en masse et en espèces permettant d'obtenir la signature IR d'un système donné, sans être gêné par celle des autres espèces en présence. Outre la spectroscopie IR des modes d'élongation NH et CO, nous verrons que la spectroscopie électronique permet également d'ob-

FIG. 3.1 – Schéma du système Ac-Phe-NH<sub>2</sub>.FIG. 3.2 – Spectre d'excitation en ionisation de Ac-Phe-NH<sub>2</sub>. Les bandes origine des conformères A, B et C sont indiquées par des flèches.

tenir des informations structurales en nous renseignant sur d'éventuelles interactions entre le chromophore aromatique Phe et la chaîne principale, et par conséquent de proposer une attribution préliminaire en termes de conformation de la chaîne. L'attribution finale sera effectuée à la lumière des calculs de fréquences harmoniques et d'énergie relative pour une sélection de structures de la molécule Ac-Phe-NH<sub>2</sub>.

### 3.1.1 Étude expérimentale

#### Spectroscopie UV

Le spectre d'excitation en ionisation, réalisé en sélectionnant la masse du système Ac-Phe-NH<sub>2</sub> ( $m = 206$  u.m.a.), est présenté sur la Figure 3.2. Il présente des bandes fines - qui témoignent d'un refroidissement rotationnel et vibrationnel efficace dans la détente - dans un domaine spectral compris entre 37500 et 38000  $\text{cm}^{-1}$ , proche du domaine d'absorption du toluène dont l'origine est située à 37477  $\text{cm}^{-1}$  [18]. On distingue dans la partie rouge du spectre ( $\nu < 37700$   $\text{cm}^{-1}$ ) plusieurs systèmes de bandes, à savoir des bandes intenses espacées

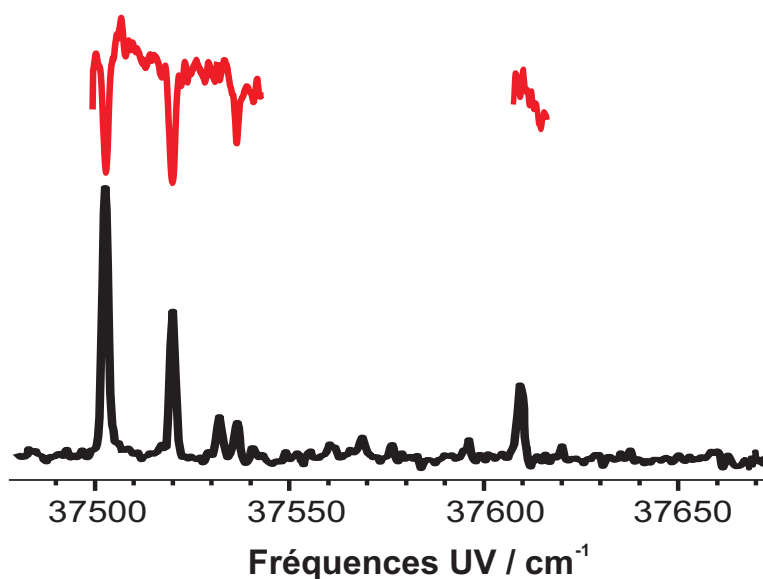


FIG. 3.3 – Spectre de dépopulation UV-UV de Ac-Phe-NH<sub>2</sub>. Le laser de sonde est placé sur la bande origine située à 37503 cm<sup>-1</sup>.

d'une vingtaine de cm<sup>-1</sup> et deux bandes situées à 37610 et 37580 cm<sup>-1</sup> (B et C). Les structures d'intensité plus faible situées dans la partie bleue du spectre ( $\nu > 37700$  cm<sup>-1</sup>) appartiennent à ces systèmes de bandes.

Des mesures de spectroscopie de dépopulation UV-UV ont mis en évidence deux conformères notés A et B dont les origines sont indiquées par des flèches dans le spectre d'excitation de la Figure 3.2. Le laser de sonde a été successivement fixé sur les bandes à 37503 et 37610 cm<sup>-1</sup>, tandis que le laser de pompe balayait le domaine inférieur à 37550 cm<sup>-1</sup> couvrant les bandes les plus rouges du spectre. Le spectre UV de l'espèce A seule a été obtenu, celui de B n'a pas été enregistré en raison de la faible intensité du signal. Le signal d'ions enregistré pour le laser de sonde fixé à 37503 cm<sup>-1</sup> (spectre de dépopulation sur la Figure 3.3) montre trois diminutions du signal à 37503, 37521 et 37539 cm<sup>-1</sup>, prouvant que ces trois bandes appartiennent au conformère A. Celles-ci constituent une progression vibrationnelle de 18 cm<sup>-1</sup>. En revanche, lorsque le faisceau de sonde est fixé sur la bande isolée à 37610 cm<sup>-1</sup> (spectre non montré), aucune dépopulation n'est observée indiquant que cette bande n'appartient pas au conformère A mais à un conformère différent que l'on a noté B.

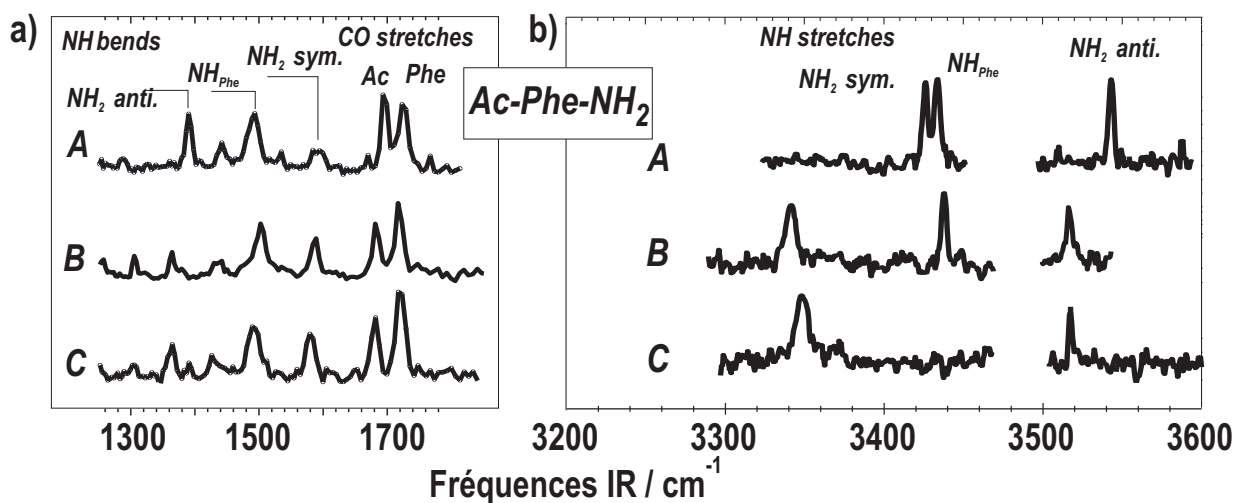


FIG. 3.4 – Spectres de dépopulation IR des conformères A, B et C de Ac-Phe-NH<sub>2</sub> dans la région amide I-II et amide A.

### Spectroscopie IR

Les spectres de double résonance IR-UV, qui seront présentés comme des bandes dans le manuscrit, ont été mesurés dans les domaines d'absorption amide A (modes d'élongation NH,  $\sim 3000\text{ cm}^{-1}$ , réalisés dans le laboratoire à Saclay) et amide I et II (élongations CO,  $\sim 1700\text{ cm}^{-1}$  et modes de déformation NH,  $\sim 1500\text{ cm}^{-1}$ , enregistrés avec le laser à électrons libres FELIX). Ainsi, en fixant successivement le laser de sonde UV sur les bandes situées à 37503, 37521, 37533, 37539 et 37610  $\text{cm}^{-1}$ , trois conformères différents ont été identifiés (Figure 3.4 a et b), dont les espèces A et B mises en évidence par spectroscopie de dépopulation UV-UV. Les origines de A, B et C sont indiquées par des flèches sur le spectre d'excitation de la Figure 3.2.

Les bandes de la progression vibrationnelle montrent toutes la même signature IR dans la région amide A et ont donc été attribuées au conformère A. La progression témoigne d'un changement de géométrie entre l'état fondamental  $S_0$  et le premier état électronique excité  $S_1$ . L'intensité de la bande origine  $0_0^0$  du conformère A indique que c'est l'espèce majoritaire, en supposant que les espèces ont la même section efficace d'absorption. Seules les bandes origines des conformères B et C sont présentes dans le spectre électronique : elles sont respectivement décalées de 107 et 30  $\text{cm}^{-1}$  vers les hautes énergies par rapport à A. Le spectre électronique de A est très différent de celui de B et C, indiquant que l'environnement du cycle aromatique

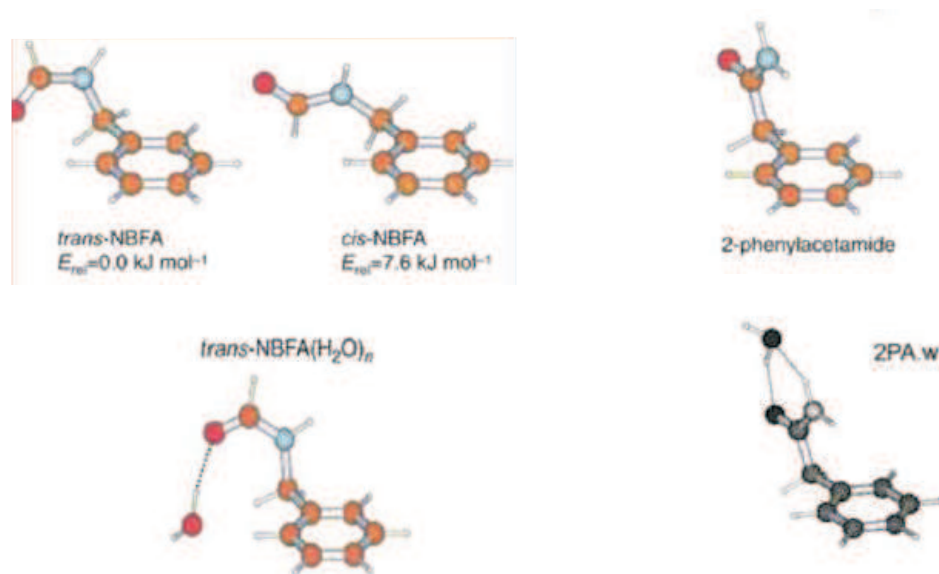


FIG. 3.5 – Systèmes modèles 2-phénylacétamide (2PA) et N-benzylformamide (NBFA) mimant la liaison peptidique. Les structures (B3LYP/6-31+G(d)) sont tirées des Ref. [19] et [20].

dans A diffère de celui de B et C. Par conséquent, la structure générale du conformère A n'est pas la même que celle des conformères B et C.

Nous commencerons par discuter les spectres IR dans la région des modes d'élongation NH (amide A) où les modes normaux correspondent à des élongations pratiquement pures, faiblement couplées aux autres modes de vibration. Puis nous aborderons les spectres mesurés dans les régions amide I et II. Les modes d'élongation CO (amide I) sont couplés à des modes de déformation NH (amide II), eux-mêmes couplés à des modes de torsion dans le plan C-N-H et d'élongation C-N du squelette peptidique et de déformation C-H (voir Section 2.2).

### Amide A

Les travaux de Simons et collaborateurs par spectroscopie de double résonance IR-UV sur le système 2-phénylacétamide ( $C_6H_5-CH_2-CO-NH_2$ , noté 2PA), système modèle mimant le groupement  $CO-NH_2$ , ont montré que les modes d'élongation  $\nu_{sym}$  et  $\nu_{anti}$  du groupe  $NH_2$  libre absorbaient à des fréquences de respectivement 3438 et 3556  $cm^{-1}$  [19] (Figure 3.5). La présence d'une liaison hydrogène entre le groupement  $NH_2$  et une molécule d'eau dans le cas

Système	Fréquences d'élongation		
	NH <sub>2</sub> sym.	NH <sub>2</sub> anti.	NH
2-phénylacétamide	3438	3556	
2-phénylacétamide-eau	3362	3525	
N-benzylformamide			3478
N-benzylformamide-eau			3440 <sup>a</sup>

TAB. 3.1 – Fréquences ( $\text{cm}^{-1}$ ) d'élongation NH des systèmes modèles étudiés par Simons et collaborateurs. Les valeurs expérimentales du 2-phénylacétamide et de son hydrate sont tirées de la Ref. [19], celles du N-benzylformamide et de l'hydrate de la Ref. [20]. <sup>a</sup> Dans ce dernier cas, on a porté la fréquence calculée avec la méthode B3LYP/6-31+G(d), corrigée avec le facteur 0.977 pour tenir compte de l'anharmonicité, car seul le complexe où l'eau est donneur sur le groupe carbonyle est observé expérimentalement.

de l'hydrate 1 : 1 de la molécule 2PA abaisse les fréquences  $\nu_{sym}$  et  $\nu_{anti}$  à des valeurs de 3362 et 3525  $\text{cm}^{-1}$ . Les études sur un autre système modèle, la molécule de N-benzylformamide ( $\text{C}_6\text{H}_5\text{-CH}_2\text{-NH-COH}$ , notée NBFA) qui mime la liaison peptidique CO-NH (Figure 3.5), ont montré que lorsque le groupement amide était libre, sa fréquence d'élongation NH prenait une valeur de 3478  $\text{cm}^{-1}$ . En revanche, lorsque celui-ci est impliqué dans une liaison hydrogène avec une molécule d'eau, cette fréquence est abaissée de 38  $\text{cm}^{-1}$ . Cette dernière valeur correspond en fait à la valeur calculée avec la méthode B3LYP/6-31+G(d) car, expérimentalement, la molécule d'eau est liée à l'atome d'oxygène dans l'hydrate. En résumé, on peut tenter de prendre les valeurs des fréquences amide A de Simons et collaborateurs comme valeurs de référence (Tableau 3.1) :

- les fréquences  $\nu_{sym}$  et  $\nu_{anti}$  d'un amide NH<sub>2</sub> libre (en fait dans un groupe CO-NH<sub>2</sub>) sont proches de 3438 et 3556  $\text{cm}^{-1}$  respectivement,
- la fréquence d'un groupe NH peptidique (CO-NH) libre est proche de 3478  $\text{cm}^{-1}$ ,
- des fréquences d'élongation NH inférieures à ces valeurs impliquent donc que les groupements NH<sub>2</sub> et NH correspondants sont liés par une liaison hydrogène. Plus la liaison hydrogène est forte, plus les fréquences seront basses et les bandes larges. Des fréquences inférieures à 3400  $\text{cm}^{-1}$  correspondent typiquement à des groupes NH perturbés par des liaisons hydrogène fortes.

Les spectres IR de A et B présentent trois bandes NH correspondant aux trois vibreurs de la molécule ; il manque une bande NH dans le spectre de C, probablement située dans le

domaine spectral où il n'y a pas d'émission à cause de l'absorption du cristal de l'OPO (3470-3500  $\text{cm}^{-1}$ ) (Figure 3.4b). Les bandes de plus basse et de plus haute fréquences appartiennent respectivement aux modes d'élongation symétrique ( $\nu_{sym}$ ) et antisymétrique ( $\nu_{anti}$ ) du groupe  $\text{NH}_2$ . La bande située vers 3430  $\text{cm}^{-1}$  appartient au groupement amide central  $\text{NH}_{Phe}$ . Les bandes ont toutes des fréquences inférieures à 3550  $\text{cm}^{-1}$ , ce qui implique que les groupes  $\text{NH}_2$  et  $\text{NH}$  sont perturbés [19, 20]. Les bandes situées dans un domaine « intermédiaire » compris entre 3420-3460  $\text{cm}^{-1}$  indiquent une perturbation plus faible. Les bandes correspondant aux modes d'élongation symétrique et antisymétrique du conformère A sont peu décalées vers le rouge (3426/3543  $\text{cm}^{-1}$ ), ce qui montre que le groupement  $\text{NH}_2$  n'est pas impliqué dans une liaison hydrogène. Les fréquences sont néanmoins sensiblement plus basses que celles du système 2-phénylacétamide (3438/3556  $\text{cm}^{-1}$ ) ; il en résulte que le groupe  $\text{NH}_2$  n'est pas complètement libre mais est légèrement perturbé. À l'inverse, les spectres des conformères B et C sont semblables, hormis l'absence de la bande de l'amide central pour C. Ils présentent des bandes d'élongation  $\text{NH}_2$  symétrique intenses et élargies apparaissant dans une région inférieure à 3400  $\text{cm}^{-1}$ . Celles-ci indiquent donc la présence d'une liaison hydrogène forte impliquant l'amide  $\text{NH}_2$  terminal. Parmi les diverses conformations que peut adopter la chaîne, seule une forme repliée  $\gamma$ , peut présenter une liaison hydrogène forte impliquant l'amide  $\text{NH}_2$  terminal - les structures canoniques  $\alpha_L$ ,  $\beta_L$ ,  $\epsilon_L$ ,  $\delta_L$  et  $\gamma_L$  issues du diagramme de Ramachandran sont montrées sur la Figure 3.6 à titre d'illustration ; les formes  $\alpha_L$ ,  $\beta_L$  et  $\gamma_L$  sont, en terme d'angles dièdres ( $\phi$ ,  $\psi$ ), respectivement, des éléments de base des hélices  $\alpha$ , des brins  $\beta$  et des coudes  $\gamma$  inverses. On peut en déduire que les conformères B et C doivent correspondre à des structures  $\gamma$  dont la liaison hydrogène  $\text{C}_7$  lie les groupes  $\text{CO}_{Ace}$  et  $\text{NH}_2$  dans un cycle à sept atomes. La bande située à 3438  $\text{cm}^{-1}$  dans le spectre de B et appartenant au groupe  $\text{NH}_{Phe}$  montre que celui-ci est légèrement perturbé. On peut expliquer cette fréquence assez basse par la présence d'une interaction  $\text{NH}-\pi$  entre le groupe  $\text{NH}_{Phe}$  et le cycle. Ainsi, l'absence de cette même bande dans le spectre de C montre que le groupe  $\text{NH}_{Phe}$  est libre, et par conséquent, que la chaîne principale interagit différemment avec le cycle.

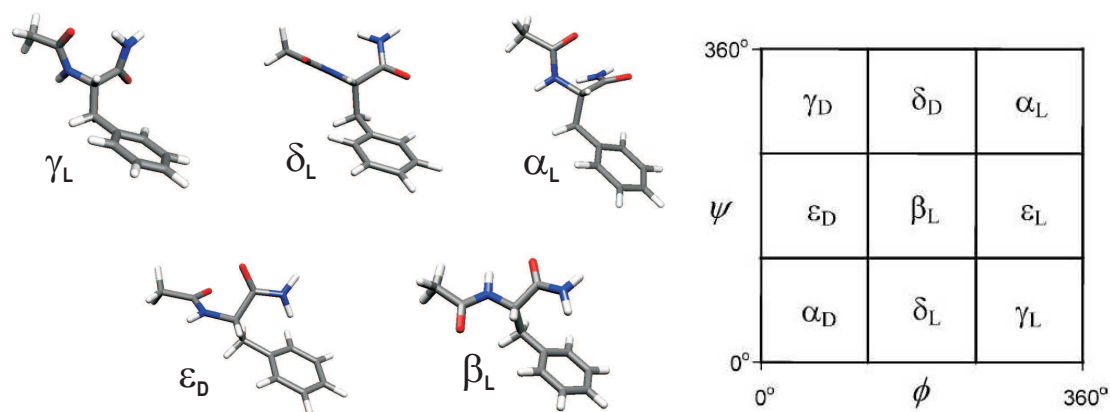


FIG. 3.6 – Géométries idéales  $\alpha_L$ ,  $\beta_L$ ,  $\epsilon_D$ ,  $\delta_L$  et  $\gamma_L$  du système Ac-Phe-NH<sub>2</sub> engendrées avec Hyperchem.

Ac-Phe-NH <sub>2</sub>			
Conformères	A	B	C
Fréquences UV	37503	37610	37533
Fréquences IR expérimentales			
Amide A (mode d'élongation NH)			
NH <sub>2</sub> sym	3426	3342	3348
NH Phe	3434	3438	a
NH <sub>2</sub> anti	3543	3517	3516
Amide I (mode d'élongation CO)			
CO Ace	1694	1682	1680
CO Phe	1725	1718	1719
Amide II (mode de déformation NH)			
NH Phe	1492	1503	1494
NH <sub>2</sub> sym	1592	1589	1580
NH <sub>2</sub> anti	1393		

TAB. 3.2 – Fréquences UV et IR (cm<sup>-1</sup>) des conformères A, B et C du système Ac-Phe-NH<sub>2</sub>.



## Amides I et II

Dans le domaine spectral sondé par le laser à électrons libres FELIX (1300-1850  $\text{cm}^{-1}$ ) peuvent être actifs les modes d'élongation CO, de déformation NH et de déformation CH. Les modes d'élongation CO sont couplés avec des modes de déformation NH, il n'est donc pas aisé d'attribuer une bande à un groupement amide donné, même en phase gazeuse. On peut en revanche estimer que la fréquence d'élongation est majoritairement composée de la vibration de ce même groupement CO. Pareillement, les modes de déformation NH sont couplés aux élongations CO et CN ainsi qu'aux modes de déformation CH ; on ne leur donnera donc qu'une attribution en termes de contribution majoritaire. Néanmoins, en général, malgré les effets du mélange avec les modes de déformation NH, les carbonyles liés par une liaison hydrogène apparaissent typiquement à des fréquences inférieures à 1700  $\text{cm}^{-1}$ , les groupes non perturbés à des fréquences supérieures.

Les spectres IR mesurés dans le domaine amide I et II sont présentés dans la Figure 3.4a. Les bandes sont plus larges que dans le domaine amide A car la largeur spectrale de FELIX est de 15  $\text{cm}^{-1}$  dans cette région. Les spectres sont similaires, avec deux bandes CO et quelques bandes distinctes entre 1400 et 1600  $\text{cm}^{-1}$  correspondant aux modes de déformation NH. La bande CO la plus rouge dans les spectres de B et C apparaît à une fréquence de respectivement 1682 et 1680  $\text{cm}^{-1}$ , ce qui confirme la présence d'une liaison hydrogène forte sur l'un des deux carbonyles, en fait celui de la protection acétyle. Dans le cas du conformère A, la fréquence plus élevée (1694  $\text{cm}^{-1}$ ) témoigne d'une perturbation plus faible sur le groupe acétyle. Dans le domaine amide II, le spectre de A diffère des deux autres conformères par la bande intense située à 1393  $\text{cm}^{-1}$  qui semble absente des spectres de B et C ou bien décalée en fréquence et de moindre intensité. À l'inverse, la bande qui apparaît dans la région de 1580  $\text{cm}^{-1}$  est moins intense pour A que pour B et C. Ces différences d'intensité indiquent un environnement différent sur les groupes amides.

Les spectres IR expérimentaux nous ont permis de proposer une attribution préliminaire en termes de structure : les conformères B et C adoptent une forme repliée de type coude  $\gamma$ , tandis que le conformère A ne présente pas de liaison hydrogène. Il est donc à présent indispensable de comparer ces résultats à des calculs de fréquences afin d'obtenir une attribution plus précise.

### 3.1.2 Calculs et attribution

Nous n'avons pas cherché à déterminer tous les minima de la surface d'énergie potentielle du système Ac-Phe-NH<sub>2</sub>. Le diagramme de Ramachandran prédit neuf conformations possibles,  $\alpha_{L/D}$ ,  $\beta_L$ ,  $\gamma_{L/D}$ ,  $\delta_{L/D}$  et  $\epsilon_{L/D}$  (Figure 3.6). Pour chacune de ces conformations, il faudrait alors considérer les trois orientations (g+, g-, a) du cycle aromatique ce qui nous amènerait à calculer 27 structures ! Nous nous sommes alors appuyés sur des études de chimie quantique menées sur des systèmes proches du nôtre. En l'occurrence, les travaux de Perczel et collaborateurs sur la molécule For-Phe-NH<sub>2</sub> [21], de Barone et collaborateurs sur la molécule Ac-Tyr-NH<sub>2</sub> [13, 14] et de Csizmadia et collaborateurs sur les systèmes Ac-Tyr-NHMe et Ac-Phe-NHMe [15] ont montré que les conformations les plus stables étaient les géométries étendues  $\beta_L$  équivalentes au brin  $\beta$  et repliées  $\gamma_L$  et  $\gamma_D$ . Les autres conformations du diagramme de Ramachandran ont des énergies plus élevées d'environ 2 kcal/mol suivant la méthode utilisée. Notre sélection de structures a alors consisté à ne retenir que les formes les plus stables citées ci-dessus. Les calculs de structures et de fréquences harmoniques ont en général été réalisés à Saclay en collaboration avec l'équipe de Chimie Théorique. Nous citerons également les travaux de G. Chass *et al.* qui, dans le cadre d'une collaboration, se sont attachés à caractériser le paysage conformationnel de la molécule Ac-Phe-NH<sub>2</sub> [22]. Les résultats de cette étude théorique pourront être consultés dans l'annexe B. Les fréquences harmoniques ont été calculées à partir de géométries optimisées avec la méthode B3LYP/6-31+G(d). Un facteur d'échelle différent a été appliqué à chaque gamme de vibration amide afin de tenir compte de l'anharmonicité du potentiel (voir section 2.3 du chapitre précédent), il sont de 0.96 pour les modes d'élongation NH, 0.98 pour les modes d'élongation CO et 0.97 pour les modes de déformation NH. Les coudes  $\gamma_L$  avec les trois orientations (g+, g-, a) du cycle aromatique, le coude  $\gamma_D(g^-)$  et la forme étendue  $\beta_L(a)$  et  $\beta_L(g^-)$  ont été calculés (Figure 3.7) Afin de tenir compte des effets de dispersion, les énergies relatives ont été calculées à partir des géométries optimisées avec la méthode MP2/6-31+G(d); elles ont été corrigées de l'énergie vibrationnelle de point zéro (ZPE) au niveau B3LYP/6-31+G(d).

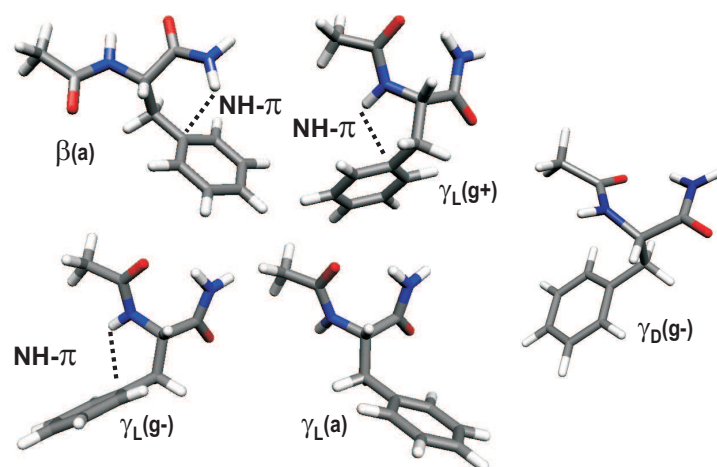


FIG. 3.7 – Géométries  $\beta_L(a)$ ,  $\gamma_L(g+)$ ,  $\gamma_L(g-)$ ,  $\gamma_L(a)$  et  $\gamma_D(g-)$  du système Ac-Phe-NH<sub>2</sub> calculées avec la méthode B3LYP/6-31+G(d).

**Énergies des conformations** Trois structures sont trouvées dans un intervalle d'énergie de 1 kcal/mol : les structures  $\beta_L(a)$ ,  $\gamma_L(g+)$  et  $\gamma_L(g-)$ , ces formes sont considérées comme isoénergétiques, compte tenu de la précision de la méthode. Les formes  $\gamma_L(a)$  et  $\gamma_D(g-)$  ont des énergies légèrement plus élevées que le minimum le plus stable  $\beta_L(a)$  (Tableau 3.3). La forme  $\beta_L(g-)$  n'est pas stable et converge vers la forme  $\gamma_L(g-)$ . On aperçoit sur la Figure 3.7 que les structures de plus basse énergie présentent une interaction NH- $\pi$ . Celle-ci relie l'amide NH<sub>2</sub> terminal au cycle phényle dans le cas de la conformation  $\beta_L(a)$ , tandis que c'est le groupe NH<sub>Phe</sub> qui interagit avec le cycle dans les conformations  $\gamma_L(g+)$  et  $\gamma_L(g-)$ . Cette interaction stabilisante est absente des formes  $\gamma_L(a)$  et  $\gamma_D$ , ce qui explique leur énergie plus élevée.

**Fréquences** L'interaction NH- $\pi$  est mesurable dans la région amide A comme le montre la Figure 3.8. En effet, la bande NH correspondante apparaît dans le domaine que l'on a qualifié d'« amides légèrement perturbés » ( $\nu \sim 3420 - 3470 \text{ cm}^{-1}$ ) par opposition aux fréquences d'élongation de groupes amides libres ( $\nu > 3470 \text{ cm}^{-1}$ , voir discussion sur les systèmes modèles étudiés par Simons et collaborateurs, p.73), et de groupements liés par une liaison hydrogène ( $\nu < 3400 \text{ cm}^{-1}$ ). La comparaison entre les spectres IR mesurés et ceux calculés confirme l'attribution des espèces B et C à des formes repliées coude  $\gamma$  (Figure 3.8) : la forme  $\gamma_L(g+)$  reproduit très bien le spectre IR de B, en particulier la bande à  $3438 \text{ cm}^{-1}$  associée au groupe

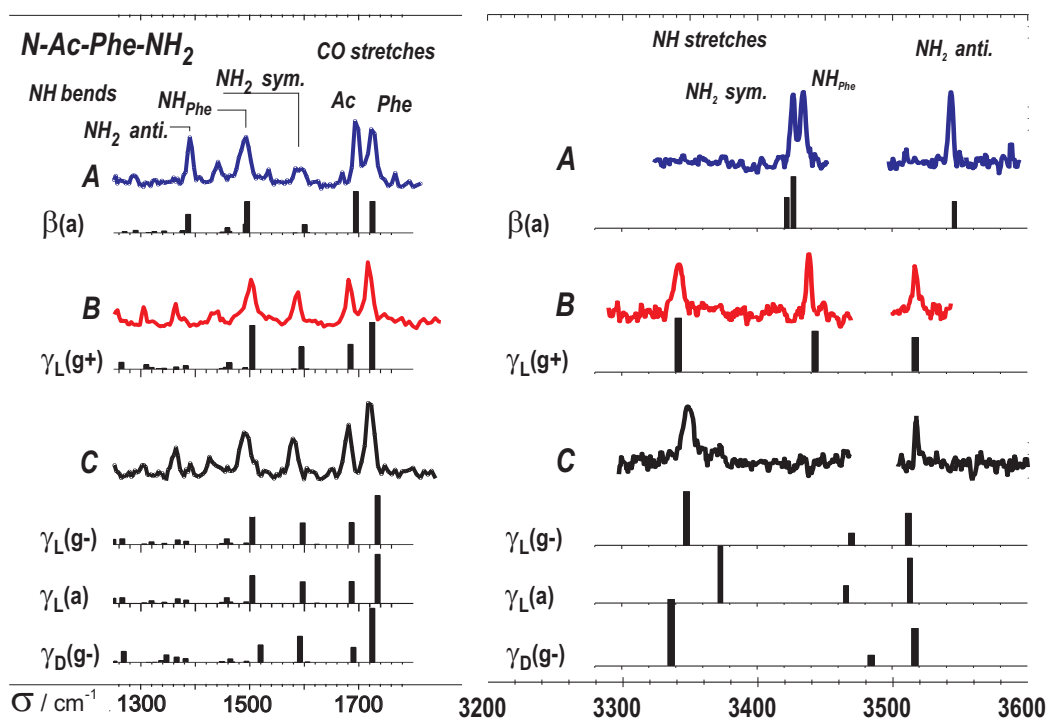


FIG. 3.8 – Comparaison entre les fréquences expérimentales des conformères A, B et C de  $\text{Ac-Phe-NH}_2$  et les fréquences théoriques des structures  $\beta_L(a)$ ,  $\gamma_L(g+)$ ,  $\gamma_L(g-)$ ,  $\gamma_L(a)$  et  $\gamma_D(g-)$ . Les fréquences ont été calculées avec la méthode B3LYP/6-31+G(d) et sont corrigées respectivement des facteurs 0.96, 0.98 et 0.97 dans les domaines amide A, amide I et amide II afin de tenir compte de l'anharmonicité et de la taille de la base.

$\text{NH}_{\text{Phe}}$  interagissant avec le cycle. Le meilleur accord pour le conformère C est obtenu avec la forme  $\gamma_L(g-)$ , en l'occurrence, la fréquence théorique  $\text{NH}_{\text{Phe}}$  donne un très bon accord avec la bande amide II expérimentale située à  $1494 \text{ cm}^{-1}$ . Le spectre vibrationnel théorique de la forme  $\beta_L(a)$  reproduit très bien celui du conformère A. Les fréquences d'élongation  $\text{NH}_2$  symétrique et antisymétrique indiquent que le groupe  $\text{NH}_2$  n'est pas complètement libre mais interagit avec le cycle aromatique via une interaction  $\text{NH}-\pi$ . Ceci explique les fréquences plus basses ( $3426/3543 \text{ cm}^{-1}$ ) que celles mesurées pour le système 2-phénylacétamide étudié par le groupe de Simons ( $3438/3556 \text{ cm}^{-1}$ ). La bande NH située à  $3434 \text{ cm}^{-1}$  est attribuée à une interaction  $\text{C}_5$  entre les groupes NH et CO du résidu Phe, cette interaction est spécifique à la conformation  $\beta_L$ .

Conformations	$\beta_L(a)$	$\gamma_L(g+)$	$\gamma_L(g-)$	$\gamma_L(a)$	$\gamma_D(g-)$
$\Delta E$ MP2	0.00	0.56	0.93	1.31	1.70
Attribution	A	B	C		
<b>Fréquences IR théoriques</b>					
Amide A (mode d'élongation NH)					
NH <sub>2</sub> sym	3422	3342	3348	3373	3337
NH Phe	3427	3443	3470	3466	3484
NH <sub>2</sub> anti	3546	3517	3512	3513	3517
Amide I (mode d'élongation CO)					
CO Ace	1695	1685	1687	1690	1690
CO Phe	1725	1725	1735	1735	1724
Amide II (mode de déformation NH)					
NH Phe	1495	1504	1504	1498	1520
NH <sub>2</sub> sym	1601	1595	1597	1601	1592
NH <sub>2</sub> anti	1383				
Distances (Å)					
C <sub>5</sub>	2.17				
C <sub>7</sub>		2.20	2.03	2.03	1.94
NH- $\pi$	2.66	2.45	2.83		

TAB. 3.3 – Énergies relatives MP2 (kcal/mol) et fréquences théoriques (B3LYP/6-31+G(d)) des conformations  $\beta_L(a)$ ,  $\gamma_L(g+)$ ,  $\gamma_L(g-)$ ,  $\gamma_L(a)$  et  $\gamma_D(g-)$  de Ac-Phe-NH<sub>2</sub>. Les fréquences sont corrigées avec le facteur 0.96 pour les modes amide A, 0.98 pour les modes amide I et 0.97 pour les modes amide II pour tenir compte de l'anharmonicité.

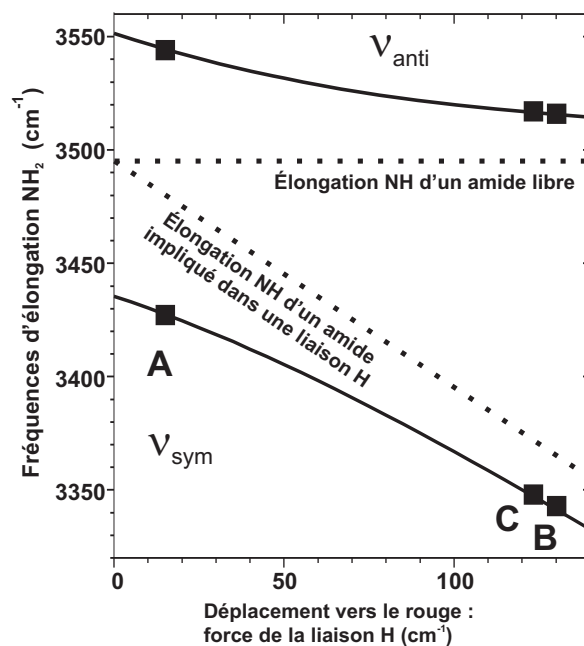


FIG. 3.9 – Modèle de couplage vibrationnel paramétré sur les fréquences des trois conformères du système Ac-Phe-NH<sub>2</sub>.

### 3.1.3 Établissement d'un modèle phénoménologique

En plus des données déduites de la spectroscopie IR, la protection C-terminale  $\text{NH}_2$  peut apporter des informations structurales supplémentaires. Plus précisément, cette protection permet d'indiquer si le groupement amino est impliqué dans une liaison hydrogène ou non. Dans un groupement amino libre, parce qu'ils partagent le même atome N, les deux vibrateurs NH sont liés par couplage vibrationnel menant à l'existence de deux modes normaux symétrique et antisymétrique. Toute perturbation sur l'une des composantes NH induit une perturbation sur l'autre, brisant la symétrie originelle : plus la perturbation est forte, plus l'écart énergétique entre les deux composantes sera grand. Nous avons établi un modèle phénoménologique impliquant un système à deux niveaux vibrationnels pour décrire la perturbation induite par la liaison hydrogène sur le groupement  $\text{NH}_2$ . Le modèle a été paramétré sur les fréquences  $\nu_{sym}$  et  $\nu_{anti}$  des trois conformères de Ac-Phe- $\text{NH}_2$ , il est discuté en détail dans l'annexe A, seule une description rapide sera donnée dans ce paragraphe. En raison de la corrélation (persistante) entre les modes symétrique et antisymétrique, la fréquence d'élongation antisymétrique permet d'identifier la composante symétrique. En outre, l'écart entre ces deux fréquences, défini par  $\Delta\nu = \nu_{anti} - \nu_{sym}$ , renseigne sur la force de la liaison hydrogène (ou éventuellement de la perturbation). On constate ainsi que les écarts  $\Delta\nu$  de 174 et 169  $\text{cm}^{-1}$  pour B et C indiquent des liaisons hydrogène fortes (Figure 3.9). Les courbes du modèle indiquent que les fréquences d'un groupement  $\text{NH}_2$  totalement libre devraient être de 3436 et 3553  $\text{cm}^{-1}$ , ce qui est très proche des valeurs trouvées par Simons et collaborateurs pour le système 2-phénylacétamide [19]. Ainsi, en plaçant les fréquences  $\nu_{sym}$  et  $\nu_{anti}$  du conformère A, on constate que le groupe  $\text{NH}_2$  n'est pas complètement libre mais légèrement perturbé, en l'occurrence par le cycle aromatique dans une interaction de type  $\text{NH}-\pi$  (Figure 3.7). Ce type d'interaction entre la chaîne latérale en position (i) et l'amide en position (i+1) est assez fréquent [23].

### 3.1.4 Effet du cycle aromatique

Quelques travaux portant sur des systèmes similaires à Ac-Phe- $\text{NH}_2$  (NAPA) sont parus au cours de ces trois années de thèse. L'attribution des trois conformères de NAPA à des structures

$\beta_L$  et  $\gamma_L$  est en accord avec les travaux de Zwier et collaborateurs sur les systèmes Ac-Trp-NHMe (NATMA) et Ac-Trp-NH<sub>2</sub> (NATA) - le résidu aromatique est le tryptophane - qui ont mis en évidence la formation des structures  $\beta_L$  avec une orientation (a) du cycle indole et des structures repliées  $\gamma_L$  [24]. On constate par conséquent que la nature du résidu aromatique (Trp vs. Phe) n'influe pas sur la conformation de la chaîne principale. Il en est de même lorsque la protection C-terminale NH<sub>2</sub> est remplacée par la fonction NHMe puisque les mêmes types de structures sont observés pour les systèmes NATA et NATMA. Ceci est également confirmé par les récents travaux de Gerhards et collaborateurs sur le système Ac-Phe-NHMe qui indiquent la formation de structures repliées  $\gamma_L$  et étendues  $\beta_L$ , cette dernière étant également la forme la plus stable [25]. En revanche, les études menées par le même groupe sur l'acide aminé protégé Ac-Phe-O-Me ont montré que seule la forme  $\beta_L$  était observée [26, 27, 28]. Or, la substitution du terminal par la fonction -O-Me empêche en effet la formation d'une liaison hydrogène entre les extrémités de la chaîne, c'est-à-dire la formation de la structure repliée  $\gamma_L$ . Par défaut, le système adopte alors une conformation étendue.

La formation de coudes  $\gamma$  pour notre système Ac-Phe-NH<sub>2</sub> est donc conforme aux nombreux calculs sur des acides aminés non aromatiques trouvant que le minimum global est la forme  $\gamma_L$ . En revanche, l'observation de la structure  $\beta_L$  peut paraître plus surprenante car elle est trouvée très haute en énergie pour les résidus non aromatiques, démontrant que l'interaction C<sub>5</sub> seule n'est pas suffisante pour stabiliser la forme étendue. Afin d'illustrer ce propos, nous avons calculé les énergies des conformations  $\beta_L$  et  $\gamma$  pour les systèmes Ac-Gly-NH<sub>2</sub> et Ac-Ala-NH<sub>2</sub> avec la fonctionnelle B3LYP et au niveau MP2 en utilisant la base 6-31+G(d) (Tableau 3.4). On constate que les résultats en DFT surestiment la stabilité de la structure  $\beta_L$  qui est trouvée quasi isoénergétique du coude  $\gamma$ . En revanche, les calculs au niveau MP2 montrent que la structure repliée  $\gamma$  est la forme la plus stable pour ces deux systèmes. On peut donc conclure que c'est la présence d'une interaction NH- $\pi$  qui permet de stabiliser la forme  $\beta_L$ (a) pour le système Ac-Phe-NH<sub>2</sub>. Ce type d'interaction entre la chaîne latérale et l'amide de la liaison peptidique est assez fréquent dans les protéines [29, 30, 31]. Il faut souligner que, parmi les structures  $\beta_L$  calculées, seule l'orientation (a) du phényle permet cette interaction NH- $\pi$ , la forme  $\beta_L$ (g+) est trouvée plus haute en énergie et la forme  $\beta_L$ (g-) n'est pas un minimum mais

Conformations	$\Delta E$ B3LYP + ZPE	$\Delta E$ MP2 + ZPE
Ac-Phe-NH <sub>2</sub>		
$\beta_L$ (a)	0.00	0.00
$\gamma_L$ (g+)	0.69	0.56
$\gamma_L$ (g-)	0.56	0.93
$\gamma_L$ (a)	0.75	1.31
$\gamma_D$ (g-)	2.24	1.70
Ac-Gly-NH <sub>2</sub>		
$\gamma_{L/D}$	-0.08	-1.39
$\beta_L$	0.00	0.00
Ac-Ala-NH <sub>2</sub>		
$\gamma_L$	-0.74	-1.27
$\beta_L$	0.00	0.00

TAB. 3.4 – Énergies relatives (kcal/mol) calculées au niveau B3LYP et MP2 avec la base 6-31+G(d) pour les systèmes Ac-Phe-NH<sub>2</sub>, Ac-Gly-NH<sub>2</sub> et Ac-Ala-NH<sub>2</sub>.

converge vers le coude  $\gamma_L$ (g-) [22]. On peut alors estimer que, d'un point de vue général, la forme la plus stable est la structure repliée coude  $\gamma$  et que la stabilité particulière de la forme  $\beta_L$  est un « épiphénomène » résultant de la présence du cycle aromatique. La formation de deux coudes  $\gamma_L$  (B et C) pour le système Ac-Phe-NH<sub>2</sub> semble conforter cette idée. On peut en ce sens citer les travaux de Lavrich et collaborateurs sur le système Ac-Ala-NHMe [10] ainsi que Compagnon et collaborateurs sur le système Z-Pro-NHMe [11] qui ont montré respectivement par spectroscopie micro-onde et par spectroscopie de double résonance IR-UV que l'unique conformère observé adoptait la forme coude  $\gamma_L$ .

### 3.1.5 Conclusion

Ces travaux sur le système Ac-Phe-NH<sub>2</sub> sont finalement très prometteurs. À l'aide des techniques de spectroscopie de double résonance couplées à une détente supersonique, nous avons pu former des structures typiques de la biologie. Malgré la complexité de la molécule qui en théorie peut présenter 27 conformations selon le diagramme de Ramachandran, seules trois espèces sont observées : deux conformères  $\gamma_L$  équivalents au coude  $\gamma$  et un conformère étendu  $\beta_L$  équivalent au brin  $\beta$ . La formation des structures repliées  $\gamma_L$  n'est pas surprenante car c'est la forme qu'adopte spontanément l'acide aminé lorsqu'il est isolé, on peut considérer que c'est sa **préférence conformationnelle locale**. En revanche, la formation de la structure



$\beta_L$  est spécifique aux résidus aromatiques et est favorisée par un effet de chaîne latérale qui se manifeste par une interaction de dispersion entre le cycle phényle et le groupe  $\text{NH}_2$  terminal de la chaîne principale. Cette interaction est suffisante pour que la forme  $\beta_L(a)$  soit le minimum le plus stable et devienne l'espèce majoritaire. De même, une stabilisation par la présence d'une interaction  $\text{NH}-\pi$  est observée pour les conformères  $\gamma_L(g+)$  et  $\gamma_L(g-)$ , cependant dans une moindre mesure, ce qui explique l'inversion de l'ordre de stabilité par rapport aux résidus non aromatiques Gly et Ala. En réalité, la forme  $\beta_L$  peut être considérée comme un « biais » introduit par le cycle aromatique ou encore comme une conformation  $\gamma$  très déformée. En effet, les conformations  $\beta_L$  et  $\gamma_L$  sont topologiquement proches dans le diagramme de Ramachandran et sont probablement séparées par de faibles barrières. Le fait que la forme  $\beta_L(g-)$  converge vers la forme  $\gamma_L(g-)$  illustre bien cette hypothèse. On s'aperçoit donc qu'il existe une compétition entre la formation d'une structure repliée et d'une structure étendue pour les résidus aromatiques.

Par ailleurs, nous avons identifié la signature IR de plusieurs types d'interactions, faibles telles que l'interaction  $C_5$  de la structure  $\beta_L$  et l'interaction  $\text{NH}-\pi$ , et fortes telles que les liaisons hydrogène  $C_7$  des coudes  $\gamma$ . Le conformère A de Ac-Phe- $\text{NH}_2$  adopte la structure étendue  $\beta_L(a)$  qui est caractérisée par l'interaction  $C_5$  et dont la signature amide A apparaît dans le domaine spectral compris entre 3420 et 3470  $\text{cm}^{-1}$ . Dans ce domaine apparaît aussi l'interaction  $\text{NH}-\pi$  entre le squelette et le cycle aromatique. Les conformères B et C adoptent une conformation repliée coude  $\gamma_L$  dont la signature se traduit par des bandes d'élongation CO et NH très décalées vers le rouge apparaissant à des fréquences inférieures respectivement à 1700 et 3400  $\text{cm}^{-1}$ . Nous avons également élaboré un modèle décrivant le couplage entre les deux composantes du groupement  $\text{NH}_2$ . Ce modèle permet, à partir de la position de la bande d'élongation  $\text{NH}_2$  antisymétrique, de prédire la position approximative de la composante symétrique et d'évaluer la force de la perturbation sur le groupe  $\text{NH}_2$ . Ainsi, lorsque la différence  $\Delta\nu$  entre les fréquences  $\nu_{anti}$  et  $\nu_{sym}$  est de l'ordre 118  $\text{cm}^{-1}$ , le groupement est considéré comme libre ou quasi libre, alors qu'une valeur de  $\Delta\nu$  plus élevée, de l'ordre de 160  $\text{cm}^{-1}$ , signifie que le groupe  $\text{NH}_2$  est impliqué dans une liaison hydrogène forte. Ces caractéristiques spectrales des liaisons hydrogène vont pouvoir être utilisées dans l'étude de

molécules plus grandes afin de proposer des attributions structurales qualitatives en terme de liaison hydrogène.

## 3.2 Dipeptides Ac-Phe-Xxx-NH<sub>2</sub> : effet de l'ajout d'un résidu

Les techniques de spectroscopie laser de double résonance IR-UV sont suffisamment sélectives pour identifier les différentes espèces observées dans les acides aminés protégés. La résolution spectrale des lasers couplée à la détente supersonique nous a en effet permis d'accéder aux caractéristiques spectroscopiques de deux types de structures. L'étape suivante consiste donc à étudier des systèmes à deux résidus, c'est-à-dire des dipeptides (on parlera également de modèles de tripeptides car le système comprend trois liaisons peptidiques). L'étude des dipeptides Ac-Phe-Xxx-NH<sub>2</sub> se veut une approche systématique nous permettant de déterminer l'effet de la nature du résidu Xxx sur la conformation de la chaîne principale. Nous avons choisi des résidus Xxx hydrophobes de taille croissante (Gly, Ala, Val, Pro) afin qu'ils ne forment pas de liaison hydrogène avec la chaîne principale, le résidu Phe est conservé et sert de chromophore UV. Les extrémités acide et basique des peptides ont été protégées afin d'éviter des effets de bord, aussi parlerons-nous parfois de la séquence Phe-Xxx pour faire référence aux dipeptides Ac-Phe-Xxx-NH<sub>2</sub>. Comme pour l'étude de la molécule Ac-Phe-NH<sub>2</sub> (NAPA), les techniques de spectroscopie UV et IR ont été appliquées aux dipeptides. La spectroscopie électronique sera présentée en première partie et la présentation des spectres IR suivra dans une seconde partie. Nous verrons que les données expérimentales nous permettront de proposer une première attribution. La discussion des résultats s'appuiera sur les conclusions déduites de l'étude de NAPA et l'attribution finale reposera sur la comparaison avec des calculs de fréquences harmoniques et de stabilité relative.

### 3.2.1 Spectroscopie électronique

Les spectres d'excitation en ionisation de Ac-Phe-Xxx-NH<sub>2</sub> sont montrés sur la Figure 3.10, par ordre de taille de résidu croissante. Les spectres électroniques des quatre systèmes sont similaires; ils présentent des bandes fines formant une progression vibrationnelle accompagnées

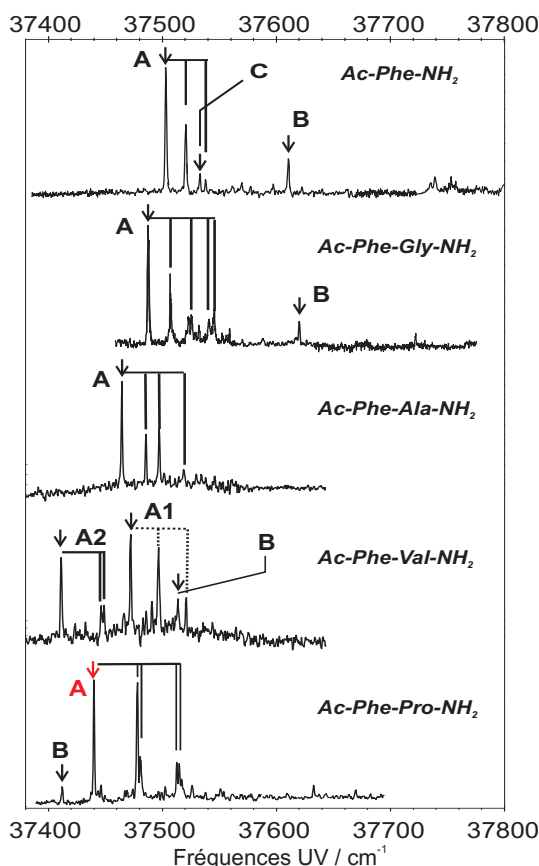


FIG. 3.10 – Spectres d’excitation IR2P de la série de dipeptides  $\text{Ac-Phe-Xxx-NH}_2$ ,  $\text{Xxx} = \text{Gly, Ala, Val, Pro}$ . Les origines de chaque conformère sont identifiées par des flèches.

de bandes isolées. Les techniques de double résonance IR-UV ont identifié deux conformères principaux pour les séquences Phe-Gly et Phe-Pro, trois pour Phe-Val et un seul pour Phe-Ala. Leur spectroscopie IR sera détaillée dans la partie suivante.

À ce stade, on peut remarquer que les progressions vibrationnelles observées ressemblent fortement à celle de l’espèce A,  $\beta_L(a)$ , de  $\text{Ac-Phe-NH}_2$ , dont le spectre UV est rappelé dans le haut de la Figure 3.10. Or, la spectroscopie électronique étant très sensible à l’environnement du cycle aromatique, des profils UV similaires suggèrent des arrangements similaires autour du cycle. Ainsi, on peut estimer que la progression vibrationnelle appartient à un conformère ayant probablement une forme étendue  $\beta_L$  avec l’interaction caractéristique  $C_5$ . Par analogie avec NAPA A, les conformères responsables de la progression vibrationnelle seront donc notés A. À l’inverse, nous pouvons déduire des bandes UV isolées que les conformères correspondants, notés B, n’adoptent pas la forme  $\beta_L(a)$  pour le résidu Phe.

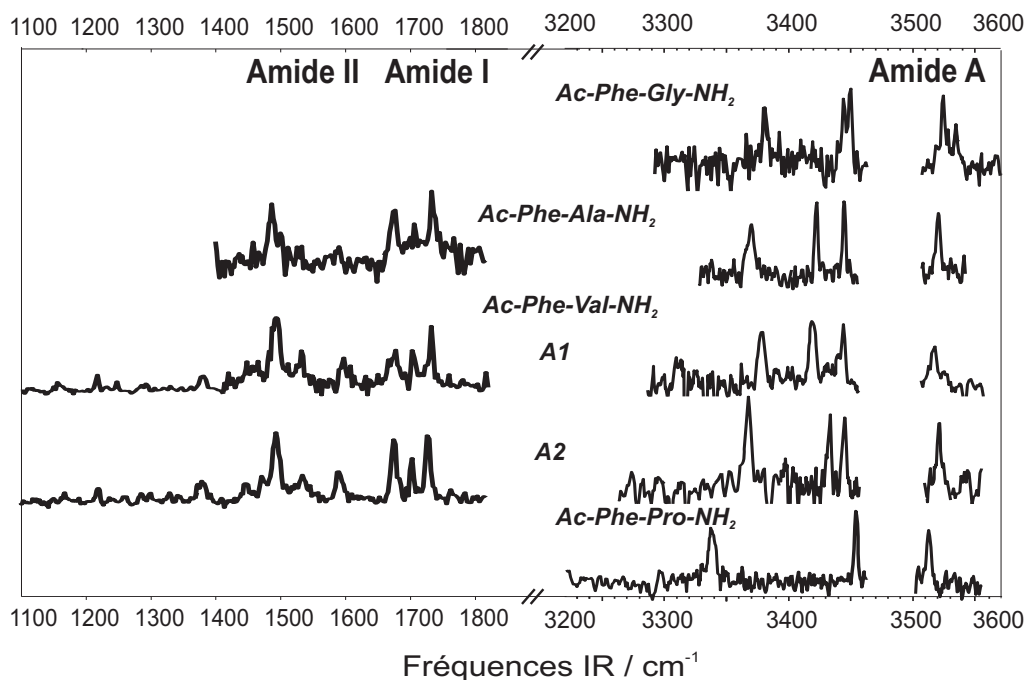


FIG. 3.11 – Spectres de dépopulation IR-UV des conformères A (A1 et A2 pour Phe-Val) de la série Ac-Phe-Xxx-NH<sub>2</sub>.

### 3.2.2 Spectroscopie IR

Les spectres de dépopulation IR-UV ont été mesurés en fixant le laser de sonde UV sur les transitions origine des conformères A et B, indiquées par des flèches sur les spectres électroniques de la Figure 3.10. Les bandes UV de moindre intensité ont été identifiées en vérifiant la présence ou l'absence des déplétions observées pour les bandes origine. Ainsi, on a pu vérifier que les bandes des progressions vibrationnelles appartenaient bien aux conformères A. Nous discuterons essentiellement des signatures IR des modes d'élongation NH car, comme nous l'avons mentionné dans la partie précédente, ces modes sont purs contrairement aux modes d'élongation CO et de déformation (pliage) NH qui sont couplés. Les spectres IR dans ce domaine spectral (1300-1800 cm<sup>-1</sup>) ont été mesurés pour Phe-Ala, Phe-Val A1 et A2, ils seront discutés à la fin de cette partie pour conforter l'attribution proposée.

Les spectres IR dans le domaine d'absorption des modes d'élongation NH des cinq conformères mis en évidence (deux formes A1 et A2 pour Phe-Val) sont montrés dans la Figure 3.11. Ils présentent quatre bandes conformément au nombre de vibrateurs NH de la molécule - ceux des deux liaisons peptidiques et les deux composantes du groupement NH<sub>2</sub> - excepté pour la

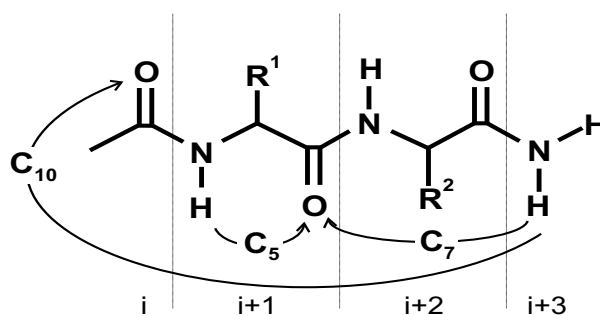


FIG. 3.12 – Schéma montrant les liaisons hydrogène possibles pour le dipeptide Ac-Phe-Xxx-NH<sub>2</sub>. R<sub>1</sub> représente le résidu Phe et R<sub>2</sub> le résidu Xxx (Gly, Ala, Val, Pro). Les notations C<sub>5</sub>, C<sub>7</sub> et C<sub>10</sub> correspondent respectivement aux liaisons hydrogène des conformations  $\beta_L$ , du coude  $\gamma$  sur Xxx et du coude  $\beta$ .

séquence Phe-Pro qui possède trois vibrateurs NH et ne présente donc que trois bandes. La position des bandes IR les plus rouges ( $\nu < 3400 \text{ cm}^{-1}$ ) indique la présence d'une liaison hydrogène relativement forte. Le léger décalage de la bande antisymétrique ( $\nu_{anti} < 3550 \text{ cm}^{-1}$ ) traduit le fait que le groupement NH<sub>2</sub> est impliqué dans la liaison hydrogène, la bande de plus basse fréquence est donc attribuée au mode d'élongation NH<sub>2</sub> symétrique. Deux classes de structures peuvent reproduire une telle liaison hydrogène : une conformation repliée coude  $\gamma$  sur le résidu Xxx reliant les groupements CO<sub>Phe</sub> et NH<sub>2</sub> dans un cycle C<sub>7</sub>, ou bien une conformation coude  $\beta$  dans laquelle les groupes CO<sub>Ace</sub> et NH<sub>2</sub> sont reliés par une liaison C<sub>10</sub> (Figure 3.12).

Deux bandes IR sont situées dans la région des groupements NH faiblement perturbés ( $3420\text{-}3470 \text{ cm}^{-1}$ ) - une bande pour Phe-Pro - elles correspondent soit à une interaction C<sub>5</sub> caractéristique de la forme  $\beta_L$ , soit à une interaction NH- $\pi$  entre le groupe NH<sub>Xxx</sub> et le cycle aromatique. La similitude entre les spectres électroniques des dipeptides Ac-Phe-Xxx-NH<sub>2</sub> et du conformère A  $\beta_L(a)$  de NAPA suggère que le résidu Phe adopte une conformation étendue  $\beta_L$ . Or, cette forme n'est pas compatible avec le coude  $\beta$  qui, bien que moins compacte que le coude  $\gamma$ , est aussi une forme repliée. On peut donc supposer que les conformères A adoptent une conformation  $\beta_L\text{-}\gamma$ , les bandes dans le domaine intermédiaire ( $3420\text{-}3460 \text{ cm}^{-1}$ ) sont attribuées à des interactions C<sub>5</sub> et NH- $\pi$ , impliquant respectivement les amides NH<sub>Phe</sub> et NH<sub>Xxx</sub>. Comme pour NAPA A, seule l'orientation (a) du cycle phényle permet à l'amide NH<sub>Xxx</sub> de pointer vers le cycle, stabilisant ainsi la forme  $\beta_L$ . Pour Phe-Pro, cette orientation permet la formation

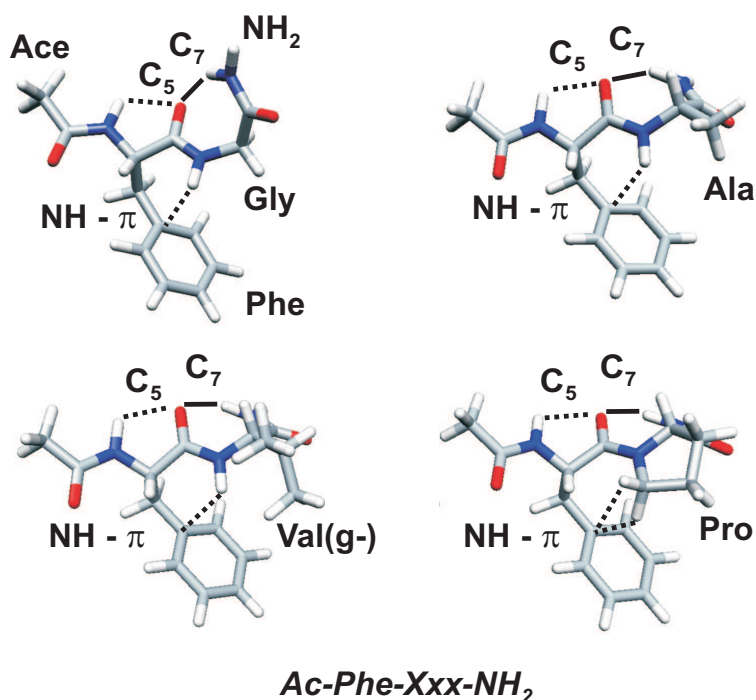


FIG. 3.13 – Conformations  $\beta_L(a)$ - $\gamma_L$  calculées pour les systèmes Ac-Phe-Xxx-NH<sub>2</sub>, Xxx = Gly, Ala, Val, Pro, au niveau B3LYP/6-31+G(d).

d'une interaction de dispersion entre le cycle et un groupement CH<sub>2</sub> de Pro, la bande située à 3454 cm<sup>-1</sup> est donc associée à une interaction C<sub>5</sub> (Figure 3.13).

**Modèle de couplage vibrationnel** En plaçant les fréquences  $\nu_{sym}$  et  $\nu_{anti}$  des conformères  $\beta_L(a)$ - $\gamma$  sur le modèle phénoménologique présenté dans la section précédente, on constate que les fréquences des dipeptides Ac-Phe-Xxx-NH<sub>2</sub> s'ajustent parfaitement aux courbes du modèle (carrés sur la Figure 3.14), ce qui confirme l'attribution des coudes  $\gamma$  sur la partie Xxx, hormis pour Phe-Val A1 dont les fréquences s'écartent de la courbe. Nous reviendrons dans une prochaine section sur les conformères A1 et A2 de Phe-Val. Les liaisons hydrogène C<sub>7</sub> correspondantes révèlent des forces plus faibles que celles des coudes  $\gamma_L(g+)$  et  $\gamma_L(g-)$  de NAPA, excepté pour Phe-Pro A qui montre une force de liaison similaire. On note en outre que la force de la liaison C<sub>7</sub> n'est pas constante mais semble se renforcer lorsque l'on passe du résidu Gly au résidu Pro, si l'on excepte le conformère A1 de Phe-Val. Nous discuterons de cet effet ultérieurement.

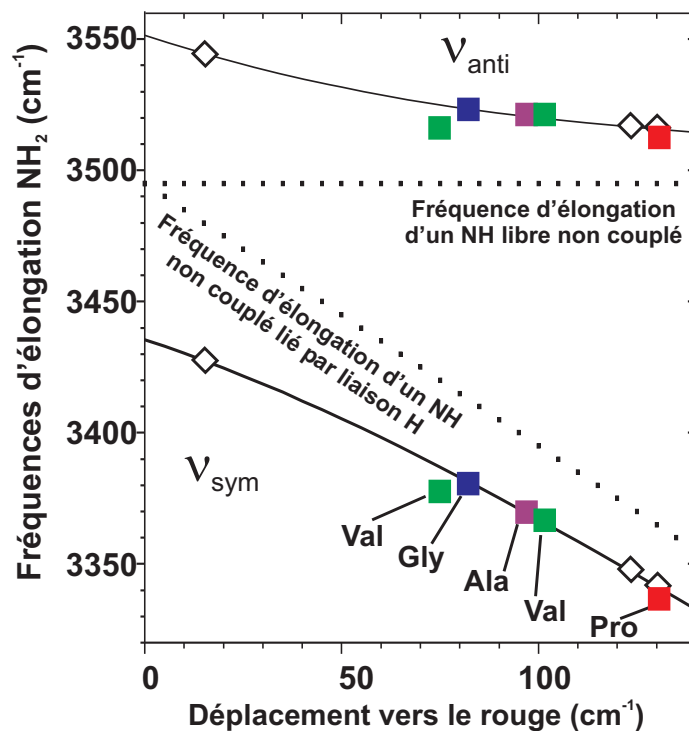


FIG. 3.14 – Fréquences d'élongation  $\nu_{sym}$  et  $\nu_{anti}$  des conformères A de Ac-Phe-Xxx-NH<sub>2</sub> (carrés) placées sur le modèle de couplage vibrationnel établi avec le système Ac-Phe-NH<sub>2</sub> (losanges).

### Spectroscopie amide I et II

Les spectres IR des modes d'élongation CO et de déformation NH ont été mesurés avec le laser à électrons libres FELIX pour les conformères A de Phe-Ala et A1 et A2 de Phe-Val (Figure 3.11). Ils sont relativement similaires : ils présentent trois bandes distinctes dans la région amide I appartenant aux groupements CO de la protection acétyle et des deux liaisons peptidiques. Deux ou trois bandes apparaissent dans la région amide II et des bandes de faible intensité sont situées à des fréquences inférieures à  $1400\text{ cm}^{-1}$  dans les spectres de Phe-Val. La fréquence des bandes CO les plus rouges ( $\sim 1675\text{ cm}^{-1}$ ) est très proche de celle observée pour les coudes  $\gamma_L$  de NAPA ( $\sim 1680\text{ cm}^{-1}$ ) confirmant l'hypothèse d'une liaison hydrogène forte sur un carbonyle et donc compatible avec une conformation repliée  $\gamma$ . La spectroscopie IR dans la région amide I conforte en particulier l'attribution pour le conformère Phe-Val A1 dont la spectroscopie dans la région amide A est sensiblement différente des autres conformères  $\beta_L$ - $\gamma$ . Les deux autres bandes CO indiquent que les groupes correspondants sont peu perturbés, voire libres pour les bandes les plus bleues situées à  $\sim 1730\text{ cm}^{-1}$ .

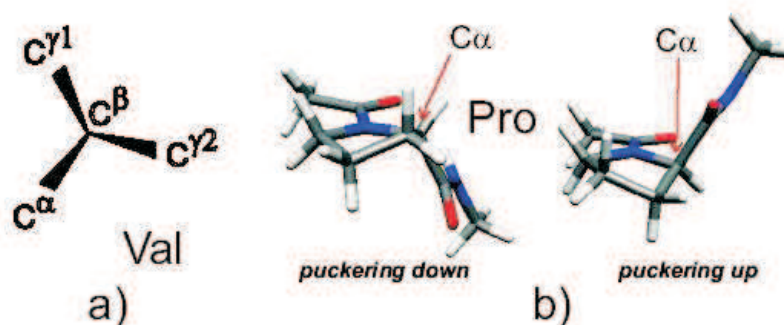


FIG. 3.15 – a) L'orientation de la chaîne latérale de Val est définie par le dièdre  $\chi_1$  C $\alpha$ -C $\beta$ -C $\gamma$ 1 selon : (a) :  $\chi_1 \approx 180^\circ$ , (g-) :  $\chi_1 \approx -60^\circ$ , (g+) :  $\chi_1 \approx 60^\circ$ . b) Configurations « pucker down » et « pucker up » du cycle à 5 atomes de Pro.

### 3.2.3 Résultats théoriques et attributions

Afin de confirmer cette première attribution, nous avons calculé les fréquences harmoniques de plusieurs conformations  $\beta_L(a)$ - $\gamma$  pour chaque dipeptide avec la méthode B3LYP/6-31+G(d) (Tableau 3.5). Nous avons considéré les formes  $\gamma_L$  et  $\gamma_D$  pour Phe-Gly car ces deux formes sont isoénergétiques dans le résidu Gly isolé [32, 3]. En revanche, la forme  $\gamma_D$  étant beaucoup plus haute en énergie pour les résidus isolés Ala (+ 2.6 kcal/mol au dessus du minimum selon la méthode employée [5, 4]) et Val [33, 21], et n'étant pas un minimum pour Pro [6], seule la forme  $\gamma_L$  a été prise en compte pour les séquences Phe-Ala, Phe-Val et Phe-Pro. Les trois orientations possibles de la chaîne latérale Val (a, g+, g-) ont été calculées (Figure 3.15a), la notation est conforme aux recommandations de la Commission IUPAC-IUB (1970) [34]. Les calculs de Barone et collaborateurs [6] ainsi que Kang [7] sur l'acide aminé Pro, naturel et protégé, ont montré que le cycle à cinq atomes formé par Pro n'était pas plan mais présentait un pli ou « pucker ». Nous avons alors retenu la configuration trouvée comme étant la plus stable en phase gazeuse c'est-à-dire la configuration dite « pucker down ». La forme  $\gamma_L$  « up » est trouvée plus haute en énergie d'environ 1.5 kcal/mol [6, 7] (Figure 3.15b).

Il y a un très bon accord entre les spectres expérimentaux et théoriques (Figure 3.16). Dans la région amide I, les calculs reproduisent de façon très satisfaisante les positions des bandes d'élongation CO. Comme nous l'avons supposé, la bande la plus rouge qui traduit



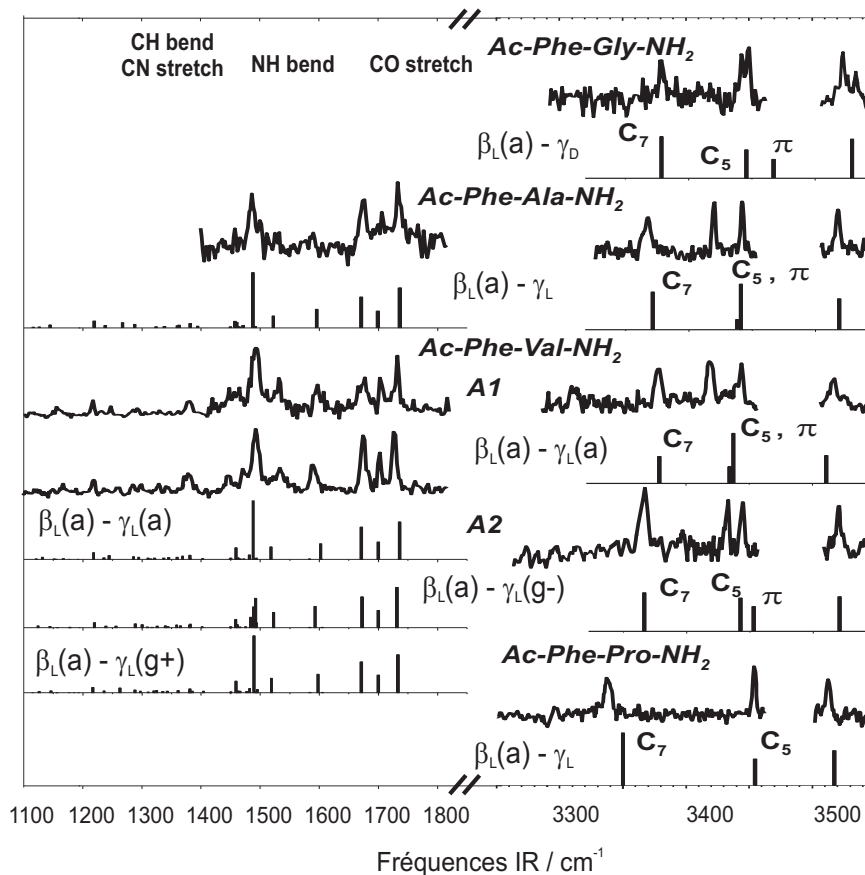


FIG. 3.16 – Comparaison entre les spectres expérimentaux des conformères A de Ac-Phe-Xxx-NH<sub>2</sub> (A1 et A2 pour Ac-Phe-Val-NH<sub>2</sub>) et les fréquences harmoniques calculées pour des structures de type  $\beta_L$ - $\gamma$ . Les fréquences sont exprimées en cm<sup>-1</sup>, les fréquences théoriques (B3LYP/6-31+G(d)) sont corrigées avec le facteur 0.96 pour les modes amide A, 0.98 pour les modes amide I et 0.97 pour les modes amide II pour tenir compte de l'anharmonicité.

une forte perturbation sur le  $\text{CO}_{Phe}$  correspond à une liaison hydrogène  $\text{C}_7$ . Les deux autres bandes sont attribuées aux carbonyles  $\text{CO}_{Ace}$  et  $\text{CO}_{Xxx}$  libres. Les bandes amide II les plus intenses sont attribuées aux modes de déformation  $\text{NH}_2$  symétrique et  $\text{NH}_{Phe}$ . Dans la région amide A, les fréquences calculées reproduisent très bien les variations de la liaison hydrogène  $\text{C}_7$  le long de la série : par exemple, la forme  $\beta_L(a)$ - $\gamma_L(a)$  calculée pour Phe-Val convient particulièrement bien au conformère Phe-Val A1 car elle correspond à la liaison hydrogène la plus faible ( $\nu_{th} = 3379 \text{ cm}^{-1}$  et  $\nu_{exp} = 3378 \text{ cm}^{-1}$ ) et la fréquence  $\nu_{sym}$  de la forme  $\beta_L(a)$ - $\gamma_L$  calculée pour Phe-Pro correspond à la liaison hydrogène la plus forte, conformément à ce qui est mesuré.

En revanche, le calcul des modes correspondant aux interactions plus faibles  $C_5$  et  $NH-\pi$  montre un accord moins satisfaisant. Hormis pour les géométries  $\beta_L(a)-\gamma_L(g-)$  et  $\beta_L(a)-\gamma_L(g+)$  de Phe-Val, les écarts entre les bandes  $C_5$  et  $NH-\pi$  des spectres expérimentaux ne sont pas retrouvés par les calculs (Tableau 3.5). Ils indiquent pour Phe-Ala et Phe-Val A1 un fort couplage des élongations  $NH$  impliquées dans les interactions  $C_5$  et  $NH-\pi$ . Les fréquences théoriques sont en effet quasi dégénérées et en conséquence très couplées - la fréquence la plus basse est associée à une élongation en phase des modes  $C_5$  et  $NH-\pi$  et la fréquence la plus élevée, qui est aussi la plus intense, à l'élongation en opposition de phase (Figure 3.17). À l'inverse pour Phe-Gly, l'écart entre les bandes  $C_5$  et  $NH-\pi$  est surestimé par les calculs pour les formes  $\beta_L(a)-\gamma_L$  et  $\beta_L(a)-\gamma_D$  (Tableau 3.5). Cette différence entre résultats expérimentaux et théoriques provient vraisemblablement de la méthode de calcul. En effet, bien que la fonctionnelle B3LYP donne de bons résultats pour les systèmes possédant des liaisons hydrogène, elle ne décrit pas correctement les interactions de dispersion [35, 36, 37, 38], en particulier les interactions  $NH-\pi$ .

Néanmoins, le bon accord trouvé pour les liaisons  $C_7$  nous permet d'attribuer les conformères A de Phe-Ala (Tableau 3.6) et Phe-Pro aux formes  $\beta_L(a)-\gamma_L$ . Les deux formes  $\beta_L(a)-\gamma_L$  et  $\beta_L(a)-\gamma_D$  peuvent convenir pour Phe-Gly. On peut en outre remarquer que cette dernière structure semble mieux reproduire les bandes centrales à 3444 et 3449  $\text{cm}^{-1}$ . Dans le cas de Phe-Val, la forme  $\beta_L(a)-\gamma_L(a)$  peut être attribuée au conformère A1; l'attribution du conformère A2 est un peu plus délicate puisque les deux formes  $\beta_L(a)-\gamma_L(g-)$  et  $\beta_L(a)-\gamma_L(g+)$  peuvent convenir (Tableau 3.5 et 3.6). L'attribution finale des conformères de Phe-Val fait l'objet du paragraphe qui suit.

### Effet de la chaîne latérale de la valine

Outre les calculs de fréquences harmoniques pour Ac-Phe-Val-NH<sub>2</sub>, nous avons calculé les énergies relatives des trois conformations  $\beta_L(a)-\gamma_L(a)$ ,  $\beta_L(a)-\gamma_L(g-)$  et  $\beta_L(a)-\gamma_L(g+)$  au niveau MP2. Les calculs montrent ainsi que les structures, c'est-à-dire à la fois les énergies et les fréquences, sont sensibles à l'orientation de la chaîne latérale de Val. Parmi les trois conformations, la structure  $\beta_L(a)-\gamma_L(a)$  est trouvée comme étant la plus stable, les deux autres sont

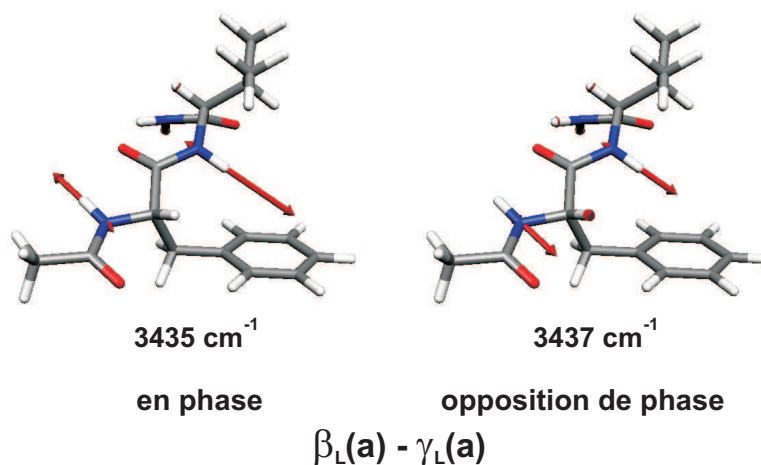


FIG. 3.17 – Vibrations d'élongation en phase et en opposition de phase des amides  $\text{NH}_{Phe}$  et  $\text{NH}_{Val}$  calculées pour la structure  $\beta_L(a)-\gamma_L(a)$  de Ac-Phe-Val- $\text{NH}_2$ . Les fréquences harmoniques sont calculées au niveau B3LYP/6-31+G(d), elles sont corrigées avec le facteur 0.96 pour tenir compte de l'anharmonicité.

moins stables de 0.70 kcal/mol pour  $\beta_L(a)-\gamma_L(g^-)$  et 1.70 kcal/mol pour  $\beta_L(a)-\gamma_L(g^+)$ . De façon surprenante, le minimum possède la distance de liaison  $\text{C}_7$  la plus grande (2.18 Å) ainsi que la fréquence  $\nu_{sym}$  la plus faible et est donc associé à la liaison hydrogène la plus faible, ce qui est confirmé par les calculs de fréquences (Tableau 3.5). Ceci démontre que la force de la liaison hydrogène n'est pas l'unique facteur déterminant la stabilité de la structure. En revanche, les orientations ( $g^-$ ) et ( $g^+$ ) de la chaîne latérale de Val mènent à des liaisons hydrogène plus fortes comme le montrent les distances  $\text{C}_7$  plus courtes (respectivement 2.05 et 2.11 Å) et les fréquences symétriques plus basses (Tableau 3.7). On peut également remarquer de légères différences au niveau des interactions  $\text{C}_5$  et  $\text{NH}-\pi$ . En l'occurrence, la faible liaison  $\text{C}_7$  de  $\beta_L(a)-\gamma_L(a)$  est associée à de courtes distances de contact  $\text{C}_5$  et  $\text{NH}-\pi$  (distance entre le groupe  $\text{NH}$  et le carbone  $\text{C}_\gamma$  du cycle aromatique), et inversement, la liaison  $\text{C}_7$  la plus forte pour  $\beta_L(a)-\gamma_L(g^-)$  semble mener aux interactions  $\text{C}_5$  et  $\text{NH}-\pi$  les plus faibles. Enfin, les modes d'élongation  $\text{C}_5$  et  $\text{NH}-\pi$  sont purs pour les formes  $\beta_L(a)-\gamma_L(g^-)$  et  $\beta_L(a)-\gamma_L(g^+)$  alors qu'ils sont couplés pour  $\beta_L(a)-\gamma_L(a)$ . On peut donc se demander si la variation de la liaison hydrogène  $\text{C}_7$ , observée expérimentalement et théoriquement, provient d'un effet de chaîne latérale du résidu Val ou bien du résidu Phe. Nous avons alors comparé les motifs  $\text{C}_7$  des formes  $\beta_L(a)-\gamma_L$  de Phe-Val à des coudes  $\gamma_L$  simples du système Ac-Val-NHMe. Ce système permet

Ac-Phe-Xxx-NH <sub>2</sub>	Fréq. UV (cm <sup>-1</sup> )	Fréquences IR expérimentales (cm <sup>-1</sup> )							
Xxx =									
Gly	37486	3381	3444	3449	3523				
Ala	37464	3370	3422	3445	3521				
Val A1	37472	3378	3419	3443	3516				
Val A2	37410	3367	3433	3445	3521				
Pro	37441	3337	3454						

	Fréquences IR calculées (cm <sup>-1</sup> )								
	NH <sub>2</sub> sym.	NH Phe		NH Xxx		NH <sub>2</sub> anti.			
Ac-Phe-Gly-NH <sub>2</sub>									
$\beta_L(a)-\gamma_L$	3378	C <sub>7</sub>	3432	C <sub>5</sub>	3457	$\pi$	3527	lib.	
$\beta_L(a)-\gamma_D$	3379	C <sub>7</sub>	3446	C <sub>5</sub>	3468	$\pi$	3529	lib.	
Ac-Phe-Ala-NH <sub>2</sub>									
$\beta_L(a)-\gamma_L$	3372	C <sub>7</sub>	3439 <sup>1</sup>	C <sub>5</sub>	3441 <sup>2</sup>	$\pi$	3519	lib.	
Ac-Phe-Val-NH <sub>2</sub>									
$\beta_L(a)-\gamma_L(a)$	3379	C <sub>7</sub>	3435 <sup>1</sup>	C <sub>5</sub>	3437 <sup>2</sup>	$\pi$	3511	lib.	
$\beta_L(a)-\gamma_L(g^-)$	3366	C <sub>7</sub>	3443	C <sub>5</sub>	3453	$\pi$	3521	lib.	
$\beta_L(a)-\gamma_L(g^+)$	3372	C <sub>7</sub>	3440	C <sub>5</sub>	3453	$\pi$	3517	lib.	
Ac-Phe-Pro-NH <sub>2</sub>									
$\beta_L(a)-\gamma_L$	3350	C <sub>7</sub>	3455	C <sub>5</sub>	-		3517	lib.	

TAB. 3.5 – Fréquences expérimentales des conformères A de Ac-Phe-Xxx-NH<sub>2</sub> (Xxx = Gly, Ala, Val, Pro) et fréquences théoriques des formes  $\beta_L(a)-\gamma$ . Les fréquences sont corrigées avec un facteur d'échelle de 0.96 pour tenir compte de l'anharmonicité. L'attribution en termes d'interaction est donnée par C<sub>5</sub>, C<sub>7</sub>, NH- $\pi$  et lib. (qui signifie que le groupe NH correspondant est libre).<sup>1,2</sup> Élongation en phase et en opposition de phase des modes C<sub>5</sub> et NH- $\pi$ .

de plus de vérifier que la protection NH<sub>2</sub> dans nos modèles n'influence pas la conformation globale de la chaîne principale.

Nous avons calculé les structures  $\gamma_L(a)$ ,  $\gamma_L(g^-)$  et  $\gamma_L(g^+)$  de Ac-Val-NHMe ainsi que leurs fréquences harmoniques avec la méthode B3LYP/6-31+G(d); les énergies relatives ont été obtenues au niveau MP2 à partir des géométries DFT (Tableau 3.7). Les paramètres géométriques - distances des liaisons C<sub>7</sub> et angles dièdres  $\phi$ ,  $\psi$  et  $\chi_1$  - et les fréquences des vibreurs NH impliqués dans les liaisons hydrogène C<sub>7</sub> indiquent des valeurs relativement proches pour les formes  $\gamma_L$  de Ac-Val-NHMe et  $\beta_L(a)-\gamma_L$  de Ac-Phe-Val-NH<sub>2</sub> et montrent ainsi que la nature des coudes  $\gamma(\text{Val})$  dans le système Phe-Val ne dépend pas de la présence du résidu Phe :

	Déformation NH			Élongation CO		
	NH Phe	NH Xxx	NH <sub>2</sub> sym	CO Phe	CO Ace	CO Xxx
Ac-Phe-Ala-NH <sub>2</sub>						
Conf. A	1487	a	1591	1675	1706	1733
$\beta_L(a)-\gamma_L$	1488	1523	1596	1672	1700	1738
Ac-Phe-Val-NH <sub>2</sub>						
Conf. A1	1493	1532	1598	1678	1703	1733
Conf. A2	1493	1534	1590	1675	1703	1726
$\beta_L(a)-\gamma_L(a)$	1488	1518	1602	1671	1700	1736
$\beta_L(a)-\gamma_L(g^-)$	1489	1523	1593	1672	1700	1731
$\beta_L(a)-\gamma_L(g^+)$	1489	1519	1598	1671	1700	1733

TAB. 3.6 – Comparaison entre les fréquences expérimentales de Phe-Ala A et Phe-Val A1 et A2 et les fréquences calculées pour les structures  $\beta_L(a)-\gamma_L$  dans la région amide I et II (modes d'élongation CO et de pliage NH). Les fréquences sont corrigées avec un facteur d'échelle de 0.98 pour les modes CO et 0.97 pour les modes NH pour tenir compte de l'anharmonicité.

Conformations	$\Delta E$ (kcal/mol)	$\nu$ C <sub>7</sub> (cm <sup>-1</sup> )	Distances (Å)			Dièdres (°)		
			C <sub>7</sub>	C <sub>5</sub>	NH- $\pi$	$\phi$	$\psi$	$\chi_1$
Ac-Phe-Val-NH <sub>2</sub>								
$\beta_L(a)-\gamma_L(a)$	0.00	3379	2.18	2.15	2.59	-82	85	-177
$\beta_L(a)-\gamma_L(g^-)$	0.70	3366	2.05	2.19	2.71	-83	58	-71
$\beta_L(a)-\gamma_L(g^+)$	1.70	3372	2.11	2.17	2.66	-83	76	60
Ac-Val-NHMe								
$\gamma_L(a)$	0.00	3398	2.21	-	-	-84	85	-178
$\gamma_L(g^-)$	1.82	3359	2.04	-	-	-85	65	-73
$\gamma_L(g^+)$	1.94	3378	2.09	-	-	-84	76	60

TAB. 3.7 – Comparaison entre les coudes  $\gamma_L$  de Ac-Val-NHMe et les formes  $\beta_L(a)-\gamma_L$  du dipeptide Ac-Phe-Val-NH<sub>2</sub>. Les fréquences de la liaison hydrogène C<sub>7</sub> sont corrigées par un facteur d'échelle de 0.96 pour tenir compte de l'anharmonicité. Les distances correspondent à la liaison C<sub>7</sub> et aux interactions C<sub>5</sub> et NH- $\pi$  dans le cas de Phe-Val. Les angles dièdres  $\phi$  et  $\psi$  correspondent à la conformation  $\gamma_L$ .  $\chi_1$  donne l'orientation de la chaîne latérale de Val par rapport au squelette peptidique.

la chaîne latérale de Phe, et de fait l'interaction NH- $\pi$  n'intervient ni dans la force de la liaison C<sub>7</sub>, ni dans l'ordre énergétique puisque qui est conservé : l'orientation (a) est le minimum le plus stable et l'orientation (g+) la forme la moins stable. Ceci semble traduire le fait qu'il n'y a pas ou peu d'interaction entre les résidus Phe et Val dans la chaîne et nous amène à suggérer que le comportement des coudes  $\gamma$  est similaire dans les systèmes Phe-Val et Ac-Val-NHMe, c'est pourquoi nous pouvons estimer que le système Phe-Val ( $\beta_L$ - $\gamma$ ) est un bon modèle de coude  $\gamma$  sur la valine.

Les coudes  $\gamma_L(a)$ ,  $\gamma_L(g^-)$  et  $\gamma_L(g^+)$  sont représentés sur la Figure 3.18. Pour les formes  $\gamma_L(g^-)$  et  $\gamma_L(g^+)$ , on remarque que les faibles distances entre les groupes CH<sub>3</sub> de la chaîne latérale et les groupes NH et CO du squelette conduisent à des interactions de contact défavorables ( $\sim 2.10$  et  $2.50$  Å, respectivement) pouvant expliquer des valeurs d'énergie plus élevées. À l'inverse, la configuration décalée de la chaîne latérale dans l'orientation (a) n'entraîne aucune gêne stérique mais elle provoque une forte distorsion du squelette - en témoigne la valeur anormalement élevée de  $\psi$  [les valeurs canoniques du couple ( $\phi$ ,  $\psi$ ) de la forme  $\gamma_L$  étant ( $-80^\circ$ ,  $60^\circ$ )] - menant à l'affaiblissement de la liaison hydrogène C<sub>7</sub>. Les valeurs de  $\phi$  et  $\psi$  des formes  $\gamma_L(g^-)$  et  $\gamma_L(g^+)$  sont plus proches des valeurs théoriques et sont associées à des liaisons hydrogène C<sub>7</sub> plus fortes. Il semble donc y avoir un effet de compensation : la formation d'une liaison hydrogène faible due à la déformation du squelette est compensée par l'absence de contacts défavorables entre la chaîne latérale et la chaîne principale. La chaîne est flexible : elle peut prendre diverses conformations. Ce sont les interactions entre le squelette peptidique et les chaînes latérales, interactions spécifiques dans le cas présent, qui permettent que différentes orientations de la chaîne latérale soient isoénergétiques car elles sont compensées par des déformations du squelette. Cette flexibilité est un facteur primordial de stabilité puisque, contrairement aux autres séquences Phe-Gly, Phe-Ala et Phe-Pro, deux conformères  $\beta_L(a)$ - $\gamma_L$  sont formés pour Phe-Val. Nous avons déjà souligné l'importance des effets de chaîne latérale lors de l'étude de Ac-Phe-NH<sub>2</sub> qui se manifestent par des interactions NH- $\pi$  permettant de stabiliser des formes isoénergétiques (dans un domaine de  $\sim 1$  kcal/mol) et topologiquement très proches, c'est-à-dire les formes  $\beta_L$  et  $\gamma_L$ . Dans le cas de Phe-Val, les effets de chaîne latérale se manifestent par un effet de gêne stérique entre la chaîne latérale de Val et la chaîne

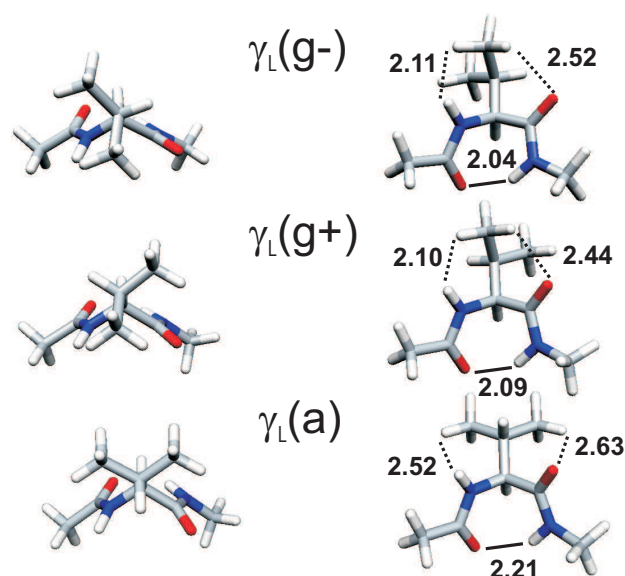


FIG. 3.18 – Conformations (B3LYP/6-31+G(d))  $\gamma_L(g-)$ ,  $\gamma_L(g+)$  et  $\gamma_L(a)$  du système Ac-Val-NHMe. Les interactions de contact entre la chaîne latérale Val et le squelette sont indiquées en pointillés et les distances des liaisons hydrogène C<sub>7</sub> sont représentées en traits pleins. Les distances sont données en  $\text{\AA}$ .

principale conduisant à la distorsion de celle-ci. Cet effet est suffisamment important pour que les deux conformations  $\beta_L(a)-\gamma_L(a)$  et  $\beta_L(a)-\gamma_L(g-)$ , qui appartiennent au même bassin conformationnel, ne s'interconvertissent pas dans la détente de sorte que, bien que qualifié d'interaction « mineure » ou faible, cet effet doit être considéré comme un facteur pouvant influencer notablement la conformation de la chaîne peptidique au même titre que les liaisons hydrogène.

### Force des coudes $\gamma$ dans la conformation $\beta_L(a)-\gamma$

Les composantes symétrique et antisymétrique de  $\text{NH}_2$  donnent accès, grâce au couplage vibrationnel, à la force de la liaison hydrogène et donc à l'environnement immédiat de  $\text{NH}_2$ . Ainsi, les coudes  $\gamma_L$  du système Ac-Phe- $\text{NH}_2$  sur lequel le modèle de couplage vibrationnel est paramétré sont associés à des forces de liaison hydrogène de l'ordre de  $130 \text{ cm}^{-1}$  (Figure 3.14). La forme  $\beta_L$  complètement dépliée, mais dont le groupe  $\text{NH}_2$  est perturbé par une interaction  $\text{NH}-\pi$ , coïncide avec une force plus faible de  $15 \text{ cm}^{-1}$ . Dans le cas des conformères A  $\beta_L(a)-\gamma$ , le modèle indique que les liaisons hydrogène C<sub>7</sub> sont plus faibles que celles des coudes  $\gamma_L(g+)$  et  $\gamma_L(g-)$  de NAPA, excepté pour Phe-Pro A qui montre une force de liaison similaire. Il

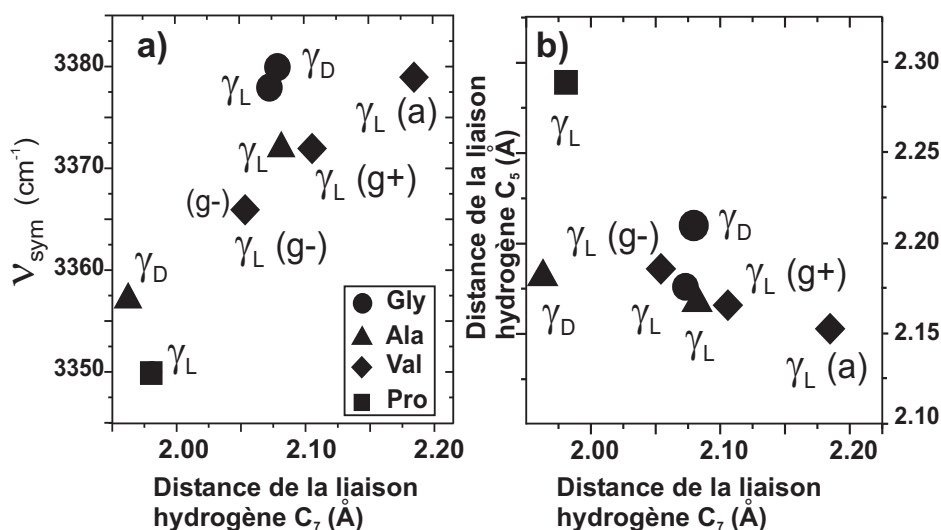


FIG. 3.19 – a) Fréquences  $\nu_{sym}$  et b) distances des interactions  $C_5$  des structures  $\beta_L(a)-\gamma$  tracées en fonction des distances des liaisons hydrogène  $C_7$ .

montre de plus que la liaison  $C_7$  se renforce en fonction de la taille du résidu, excepté pour Phe-Val A1 dont la liaison  $C_7$  est au contraire la plus faible.

Afin d'illustrer ce propos, nous avons tracé les fréquences théoriques  $\nu_{sym}$  des formes  $\beta_L(a)-\gamma$  en fonction des distances calculées pour les liaisons  $C_7$  (Figure 3.19a). On remarque que la distance  $C_7$  varie en fait très peu en fonction du résidu, elle est comprise entre 2.05 et 2.11 Å, hormis pour les formes  $\beta_L(a)-\gamma_L(a)$  de Phe-Val et  $\beta_L(a)-\gamma_L$  de Phe-Pro pour lesquelles elle prend des valeurs extrêmes (2.18 et 1.98 Å). On peut notamment souligner qu'à ces structures sont associées les distances  $C_5$  la plus courte et la plus longue, respectivement (2.15 et 2.28 Å), comme l'illustre la Figure 3.19b. Ceci semble témoigner d'un effet de compensation, du moins d'interaction entre les configurations  $C_5$  et  $C_7$ . Or, ceci semble en contradiction avec les observations faites dans le paragraphe précédent. Nous avons donc comparé les fréquences théoriques et les paramètres géométriques des formes  $\beta_L(a)-\gamma$  de la série Ac-Phe-Xxx-NH<sub>2</sub> avec des coudes  $\gamma$  simples de la série Ac-Xxx-NHMe, comme nous l'avons fait pour le résidu Val dans le paragraphe précédent (Tableau 3.8). Le tableau indique des valeurs relativement proches pour les coudes  $\gamma$  dans les deux séries, signalant que la conformation  $\gamma$  est peu affectée par la présence de Phe dans les dipeptides Ac-Phe-Xxx-NH<sub>2</sub>, conformément au résultat trouvé pour Ac-Phe-Val-NH<sub>2</sub>. Dans ce cas, les distances  $C_7$  très longue et  $C_5$  très courte de Phe-Val A1 peuvent s'expliquer par la distorsion de la chaîne principale qui résulte de la contrainte



X <sub>xxx</sub> =		Ac-Phe-X <sub>xxx</sub> -NH <sub>2</sub>				Ac-X <sub>xxx</sub> -NHMe			
		$\nu_{sym}$	$\beta_L(a)$	$\gamma$		$\nu_{C_7}$	$\gamma_L$		
		d C <sub>7</sub>	$\phi$	$\psi$		d C <sub>7</sub>	$\phi$	$\psi$	
Gly	$\gamma_L$	3378	2.07	-80	62	3385	2.07	-82	69
	$\gamma_D$	3379	2.08	84	-63	3385	2.07	82	-69
Ala	$\gamma_L$	3372	2.08	-81	74	3378	2.08	-83	75
Val	$\gamma_L(g-)$	3366	2.05	-83	58	3359	2.04	-85	65
	$\gamma_L(g+)$	3372	2.11	-83	76	3378	2.09	-84	76
	$\gamma_L(a)$	3379	2.18	-82	85	3398	2.21	-84	85
Pro	$\gamma_L$	3350	1.98	-83	71	3348	1.98	-83	70

TAB. 3.8 – Fréquences et paramètres géométriques (B3LYP/6-31+G(d)) des structures  $\gamma_L$  de Ac-X<sub>xxx</sub>-NHMe et  $\beta_L(a)$ - $\gamma$  de Ac-Phe-X<sub>xxx</sub>-NH<sub>2</sub> (X<sub>xxx</sub> = Gly, Ala, Val and Pro). Les fréquences (cm<sup>-1</sup>) sont corrigées avec le facteur d'échelle 0.96, les angles dièdres ( $\phi$ ,  $\psi$ ) des coudes  $\gamma$  sont donnés en degrés et les distances en angstroms.

induite par la chaîne latérale de Val (voir paragraphe précédent). La distance C<sub>7</sub> très courte trouvée pour Phe-Pro est elle aussi probablement due à la contrainte exercée sur la chaîne par l'absence d'un atome H.

## Conclusion

Finalement, les conformères A adoptent des structures qui correspondent à la succession des préférences conformationnelles locales pour chaque résidu le long de la chaîne. Les résidus hydrophobes Gly, Ala, Val et Pro prennent une conformation  $\gamma$  qui est leur minimum le plus stable et le résidu aromatique Phe a une conformation étendue  $\beta_L(a)$ , comme pour l'espèce A du système Ac-Phe-NH<sub>2</sub>. Nous avons également montré que la nature du résidu - plus précisément la taille du résidu hydrophobe - n'influçait pas le comportement général de la chaîne peptidique puisque chaque résidu adopte sa forme la plus stable, signalant ainsi qu'il interagit peu avec son voisin dans la séquence. Néanmoins nous avons mis en évidence un effet de chaîne latérale, qui dans notre cas, se manifeste par la stabilisation de certaines formes. En effet, seule l'orientation (a) du cycle aromatique permet la formation de l'interaction NH-

$\pi$  essentielle à la stabilisation de la structure étendue  $\beta_L$ . Dans le cas de Phe-Pro, cette orientation mène à une interaction de dispersion entre le cycle phényle et un groupement méthylène du cycle de Pro. Dans le cas Phe-Val, les effets de la chaîne latérale volumineuse de Val permettent de stabiliser une deuxième structure  $\beta_L$ - $\gamma_L$ . Bien que l'on soit tenté de les qualifier de mineures, ces interactions plus faibles montrent bien que les effets de chaîne latérale participent intégralement au repliement de la chaîne peptidique.

### 3.3 Dipeptides Ac-Xxx-Phe-NH<sub>2</sub> : effet de l'ordre de la séquence

Après l'étude de la série de dipeptides Ac-Phe-Xxx-NH<sub>2</sub>, il nous a paru logique d'étendre nos travaux à la série Ac-Xxx-Phe-NH<sub>2</sub> qui présente la séquence inverse. Alors que la nature du résidu hydrophobe influence peu le comportement de la chaîne, on peut alors se demander si l'ordre des résidus est important. Comme pour la série précédente, nous avons inséré un second résidu Xxx (Xxx = Gly, Ala, Val, Pro) dans la séquence. Les techniques expérimentales utilisées lors de l'étude des systèmes Ac-Phe-NH<sub>2</sub> et Ac-Phe-Xxx-NH<sub>2</sub> ont été employées pour les dipeptides Ac-Xxx-Phe-NH<sub>2</sub>. Les résultats expérimentaux seront accompagnés de calculs de fréquences et d'énergie.

#### 3.3.1 Spectroscopie électronique

Les spectres électroniques des dipeptides Ac-Xxx-Phe-NH<sub>2</sub> sont montrés dans la Figure 3.20. Ils sont très différents des systèmes précédents : à la place de bandes fines et espacées - isolées ou appartenant à une progression vibrationnelle d'environ 20 cm<sup>-1</sup> - les spectres de Ac-Xxx-Phe-NH<sub>2</sub> sont congestionnés et présentent une très forte activité Franck-Condon, conséquence d'un changement de géométrie entre l'état fondamental S<sub>0</sub> et le premier état électronique excité S<sub>1</sub>. Pour un résidu donné, on remarque aussi que les spectres sont décalés vers les hautes énergies par rapport aux spectres des systèmes Ac-Phe-Xxx-NH<sub>2</sub>. On peut donc déjà déduire de la signature UV, domaine spectral différent et forte activité Franck-Condon, que la perturbation du nuage électronique du cycle aromatique est différente et par conséquent que le résidu Phe n'adopte pas la conformation étendue  $\beta_L$ .

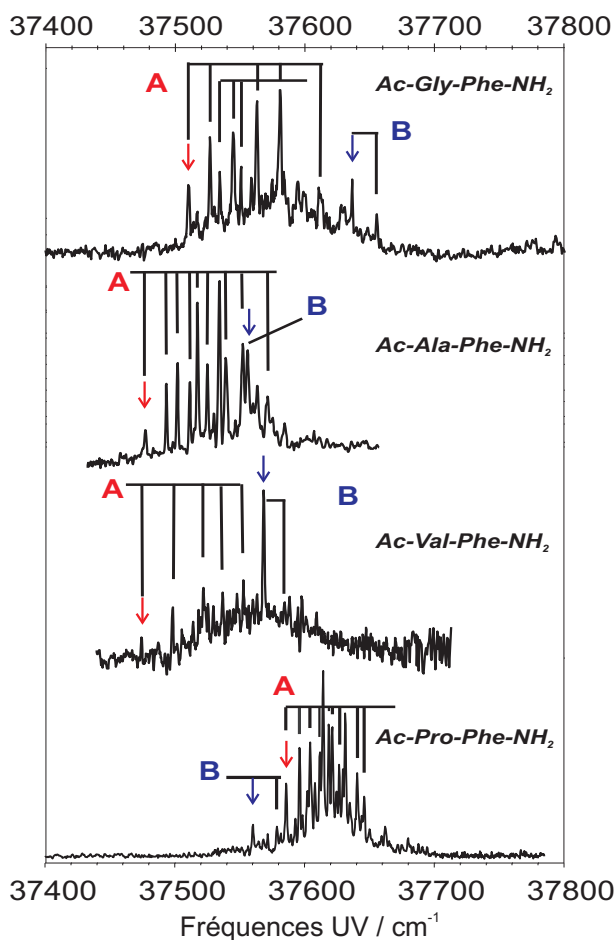


FIG. 3.20 – Spectres d'excitation des dipeptides  $\text{Ac-Xxx-Phe-NH}_2$ ,  $\text{Xxx} = \text{Gly, Ala, Val, Pro}$ . Les bandes origine sont représentées par des flèches.

En appliquant les techniques de double résonance IR-UV, nous avons identifié deux conformères significativement peuplés, notés A et B, dont les bandes origine sont indiquées par des flèches sur la Figure 3.20. Nous verrons dans le paragraphe qui suit que les conformères A possèdent tous une signature IR semblable. Pour les séquences Gly-Phe, Ala-Phe et Pro-Phe, le conformère A montre les bandes les plus intenses, c'est donc l'espèce majoritaire. En revanche, c'est le conformère B qui est le plus abondant pour Val-Phe. Nous avons vérifié que les bandes des progressions vibrationnelles avaient les mêmes déplétions IR que la bande origine de A. Les conformères B montrent des bandes plus isolées, une seule bande pour Ala-Phe ou bien une progression vibrationnelle relativement courte pour les autres systèmes. Hormis pour Pro-Phe B dont les résultats seront exposés dans ce paragraphe, les conformères B seront discutés dans une autre partie car leur signature UV et IR semble révéler une géométrie totalement

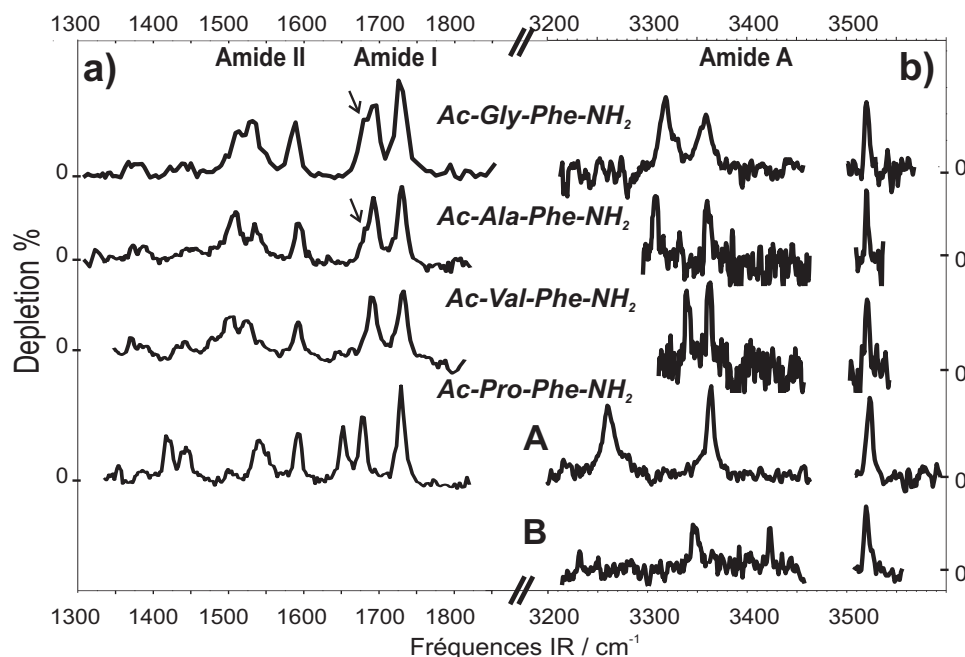


FIG. 3.21 – Spectres IR dans les domaines amide A et amide I et II des conformères A de la série de dipeptides  $\text{Ac-Xxx-Phe-NH}_2$  et du conformère B pour  $\text{Ac-Pro-Phe-NH}_2$ .

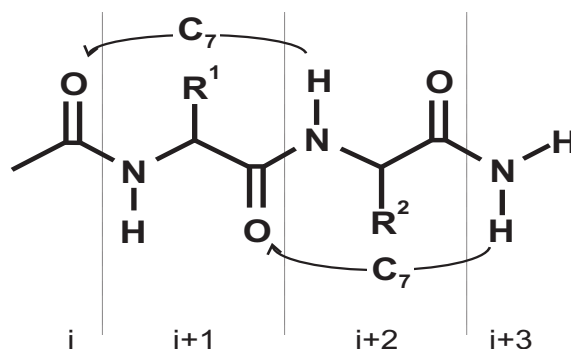
différente de celle de A.

### 3.3.2 Spectres IR

Les spectres IR dans les régions amide A (mode d'élongation NH) et amide I et II (modes d'élongation CO et de déformation dans le plan CNH) ont été mesurés en fixant le laser de sonde UV sur les bandes intenses des conformères A, ils sont présentés sur la Figure 3.21.

#### Amide A

Les dipeptides  $\text{Ac-Xxx-Phe-NH}_2$  possèdent quatre vibrateurs NH, trois vibrateurs pour  $\text{Ac-Pro-Phe-NH}_2$  (ou Pro-Phe), on doit donc voir apparaître quatre bandes, respectivement trois bandes. On observe effectivement trois bandes NH pour Pro-Phe mais une bande est manquante dans les spectres de Gly-Phe, Ala-Phe et Val-Phe (Figure 3.21b). Elle est probablement située dans le domaine spectral qui n'est pas couvert par le laser OPO ( $3470\text{-}3500\text{ cm}^{-1}$ ), ce qui indiquerait que les trois dipeptides ont un groupement NH libre. On observe deux bandes très décalées vers le rouge et une bande dans un domaine de fréquences plus élevées apparte-

FIG. 3.22 – Schéma d'une structure présentant deux liaisons hydrogène  $C_7$ .

nant au mode d'élongation antisymétrique du groupement  $NH_2$ . La position des deux bandes  $NH$  ( $\nu < 3400\text{ cm}^{-1}$ ) révèle la présence de **deux** liaisons hydrogène fortes, contrairement aux dipeptides  $Ac-Phe-Xxx-NH_2$  qui possèdent une liaison hydrogène **unique**. La bande antisymétrique, légèrement décalée vers le rouge ( $\sim 3520\text{ cm}^{-1}$ ) par rapport à un groupement amino libre ( $\sim 3550\text{ cm}^{-1}$  [19]), indique que celui-ci est impliqué dans une des deux liaisons hydrogène. Or, d'après la Figure 3.22 schématisant les liaisons hydrogène possibles dans le système, seule une conformation de type  $\gamma$ - $\gamma$ , c'est-à-dire possédant deux coudes  $\gamma$ , peut présenter deux liaisons hydrogène fortes.

### Amide I et II

Les spectres IR des systèmes Gly-Phe, Ala-Phe et Val-Phe montrent des profils d'absorption similaires (Figure 3.21a) ; le spectre de Pro-Phe sera traité dans une section différente. Ils présentent deux bandes CO et trois bandes larges dans les domaines amide I et II. La bande CO de plus basse fréquence est accompagnée d'un épaulement sur le côté rouge pour Gly-Phe et Ala-Phe (indiqué par une flèche sur le spectre). Tous montrent des bandes situées à des fréquences inférieures à  $1700\text{ cm}^{-1}$  indiquant que les carbonyles correspondants sont impliqués dans des liaisons hydrogène fortes. Les épaulements observés dans le cas de Gly-Phe et Ala-Phe pourraient correspondre à une deuxième liaison hydrogène  $C_7$  compatible avec une structure coude  $\gamma$  double.

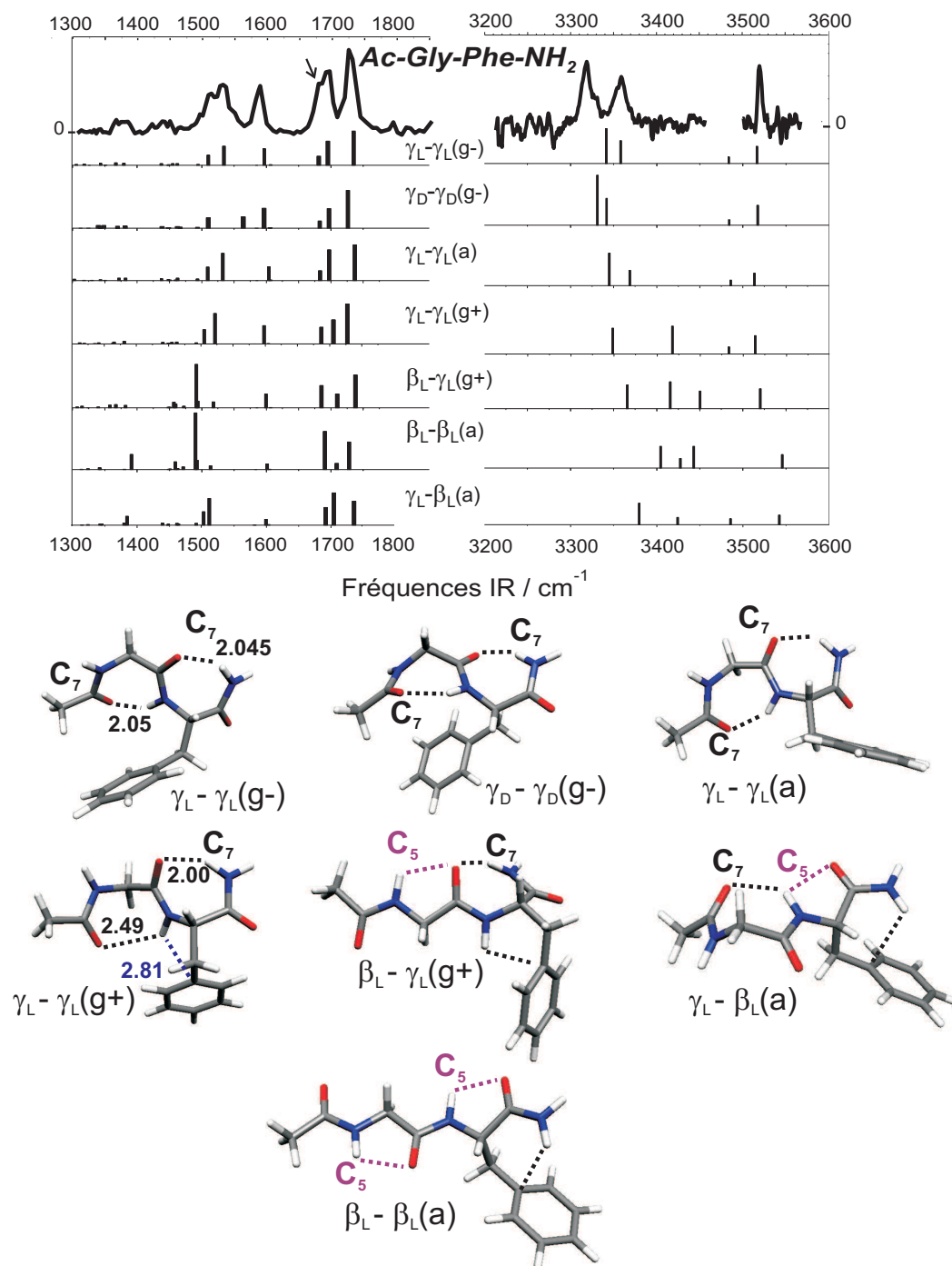


FIG. 3.23 – Spectre IR de Ac-Gly-Phe-NH<sub>2</sub> A comparé aux fréquences calculées (B3LYP/6-31+G(d)) pour une sélection de structures. Les fréquences sont corrigées avec le facteur 0.96 dans la région amide A, 0.98 dans la région amide I et 0.97 dans la région amide II.

	Ac-Xxx-Phe-NH <sub>2</sub>		
	Ac-Gly-Phe-NH <sub>2</sub>	Ac-Ala-Phe-NH <sub>2</sub>	Ac-Val-Phe-NH <sub>2</sub>
Fréquences UV	37510	37477	37474
Fréquences IR expérimentales			
Amide A (mode d'élongation NH)			
	3319	3308	3340
	3359	3360	3362
	<sup>a</sup>	<sup>a</sup>	<sup>a</sup>
	3520	3520	3520
Amide I (mode d'élongation CO)			
	1678 <sup>b</sup>	1680 <sup>b</sup>	
	1695	1691	1689
	1725	1729	1733
Amide II (mode de déformation NH)			
	1510	1508	1502
	1531	1534	1525
	1588	1592	1593

TAB. 3.9 – Fréquences expérimentales UV et IR (cm<sup>-1</sup>) des conformères A de Ac-Xxx-Phe-NH<sub>2</sub>, Xxx = Gly, Ala, Val. <sup>a</sup> les bandes manquantes sont situées dans la zone spectrale non couverte par le cristal de l'OPO. <sup>b</sup> ces bandes correspondent à l'épaule observé pour Ac-Gly-Phe-NH<sub>2</sub> et Ac-Ala-Phe-NH<sub>2</sub>.

### 3.3.3 Calculs : recherche des conformations pour Ac-Gly-Phe-NH<sub>2</sub>

Afin de confirmer cette attribution, nous avons comparé les spectres expérimentaux avec ceux calculés pour différentes conformations de la chaîne peptidique. Les calculs ont été menés pour le système Ac-Gly-Phe-NH<sub>2</sub> car l'absence de chaîne latérale permet de réduire les temps de calcul. En revanche, la petite taille du résidu Gly, et donc l'absence de gêne stérique implique a priori un paysage conformationnel accessible plus vaste. Néanmoins, nous n'avons pas cherché à explorer la surface d'énergie potentielle de façon exhaustive mais, par analogie avec les systèmes Ac-Phe-NH<sub>2</sub> et Ac-Phe-Xxx-NH<sub>2</sub> étudiés dans les sections précédentes, nous avons sélectionné des combinaisons de formes locales, c'est-à-dire les formes les plus stables pour chaque résidu telles que les conformations  $\gamma$ - $\gamma$  (C<sub>7</sub>-C<sub>7</sub>),  $\beta_L$ - $\gamma$  (C<sub>5</sub>-C<sub>7</sub>) et  $\gamma$ - $\beta_L$  (C<sub>7</sub>-C<sub>5</sub>).

Parmi ces différentes conformations, les formes  $\gamma$ - $\gamma$  reproduisent le mieux les spectres IR des conformères A (Figure 3.23). L'accord est très satisfaisant dans les domaines amide I et II : la bande CO la plus rouge, ainsi que les épaulements dans le cas de Gly-Phe et Ala-Phe correspondent aux groupements carbonyles ( $\text{CO}_{Xxx}$  et  $\text{CO}_{Ace}$ ) impliqués dans les liaisons hydrogène  $\text{C}_7$ . L'absence d'une deuxième bande CO vers  $1680\text{ cm}^{-1}$  pour Val-Phe peut, elle, s'expliquer par le fait que les fréquences calculées ne sont séparées que d'une dizaine de  $\text{cm}^{-1}$ . Elles ne sont pas résolues en raison de la largeur spectrale du laser FELIX qui est de  $15\text{ cm}^{-1}$  dans ce domaine. Dans la région amide A, les calculs attribuent les bandes les plus rouges des spectres expérimentaux à la vibration d'élongation du groupe  $\text{NH}_{Phe}$ , c'est-à-dire au coude  $\gamma$  sur le résidu Xxx, et les bandes situées à environ  $3360\text{ cm}^{-1}$  au mode  $\text{NH}_2$  symétrique, c'est-à-dire au coude  $\gamma$  sur le résidu Phe. On remarque ainsi que le mode d'élongation  $\text{NH}_2$  symétrique absorbe toujours à une fréquence d'élongation d'environ  $3360\text{ cm}^{-1}$  indépendamment de la nature du résidu Xxx. À l'inverse, le mode  $\text{NH}_{Phe}$  absorbe à des fréquences différentes, la liaison hydrogène  $\text{C}_7$  étant plus ou moins forte suivant la nature de Xxx. L'accord entre les spectres expérimentaux et théoriques est moins bon dans la région amide A : les fréquences des modes d'élongation NH des liaisons hydrogène  $\text{C}_7$  sont surestimées d'une vingtaine de  $\text{cm}^{-1}$  par les calculs et l'écart entre les deux bandes n'est pas retrouvé. Or, les deux liaisons hydrogène ont en commun la liaison peptidique centrale, entraînant la polarisation de celle-ci. Les deux liaisons  $\text{C}_7$  sont en fait en compétition, la conformation résulte de l'équilibre entre les deux forces. Cet équilibre subtil peut alors ne pas être correctement décrit par les calculs. Notamment, alors que les calculs réalisés sur plusieurs coudes  $\gamma$ - $\gamma$  indiquent que la fréquence d'élongation la plus basse appartient au groupement amide  $\text{NH}_{Phe}$ , on observe une inversion de fréquence entre les modes d'élongation  $\text{NH}_{Phe}$  et  $\text{NH}_2$  symétrique pour la conformation  $\gamma_L$ - $\gamma_L(g+)$ . Or, on constate que cette conformation montre une signature IR très différente, le mode d'élongation  $\text{NH}_{Phe}$  dont la fréquence est anormalement élevée ( $3418\text{ cm}^{-1}$ ) ne peut en fait rendre compte d'une liaison  $\text{C}_7$  mais correspond vraisemblablement à une interaction  $\text{NH}-\pi$  favorisée par l'orientation ( $g+$ ) du phényle. L'orientation ( $g+$ ) du cycle conduit donc à un phénomène de compétition entre la formation d'une liaison  $\text{C}_7$  sur le résidu Gly et celle d'une interaction  $\text{NH}-\pi$ , en faveur de cette dernière. On peut ainsi souligner l'affaiblissement de la liaison  $\text{C}_7$



au profit de l'interaction NH- $\pi$ , les distances NH $\cdots$ O et NH $\cdots$ C $\gamma$  étant respectivement de 2.81 et 2.00 Å. À l'inverse, ces distances sont assez proches pour les deux liaisons C $\gamma$  dans le cas des autres coudes  $\gamma$ - $\gamma$  : elles ne diffèrent que d'environ 0.05 Å. On constate par ailleurs que les diverses conformations  $\gamma$ - $\gamma$  calculées, hormis  $\gamma_L$ - $\gamma_L(g+)$ , ont une signature IR très semblable, il est donc difficile de donner une attribution précise quant à l'orientation du cycle aromatique ou en terme de configuration  $\gamma_L$  ou  $\gamma_D$  pour Gly-Phe. On peut néanmoins écarter la forme  $\gamma_D$ - $\gamma_D(g-)$  dont la fréquence de déformation NH<sub>Phe</sub> est significativement plus élevée (1565 cm<sup>-1</sup>) que la valeur expérimentale (1525 cm<sup>-1</sup>). Les autres structures calculées peuvent également être écartées : la conformation  $\beta_L$  sur le résidu Phe pour les structures  $\beta_L$ - $\beta_L$  et  $\gamma_L$ - $\beta_L$  mène à un groupement NH<sub>2</sub> libre non compatible avec le spectre expérimental amide A. Enfin, la forme  $\beta_L$ - $\gamma_L$  ne présente qu'une seule liaison hydrogène (Figure 3.23).

Nous n'avons pas effectué de calculs pour les systèmes Ala-Phe et Val-Phe mais par analogie avec Gly-Phe, seules les conformations  $\gamma$ - $\gamma$  sont à retenir. En outre, les spectres expérimentaux ne présentant pas de bande dans la région 3420-3470 cm<sup>-1</sup>, où apparaissent les interactions NH- $\pi$  et C<sub>5</sub>, la conformation  $\gamma_L$ - $\gamma_L(g+)$  peut aussi être écartée. Enfin, la configuration  $\gamma_D$  étant plus haute en énergie que la configuration  $\gamma_L$  pour les résidus Ala [5], Val [21] et Phe (cf section 3.1), les dipeptides adoptent donc probablement une conformation  $\gamma_L$ - $\gamma_L$ .

### 3.3.4 Conformères A et B de Ac-Pro-Phe-NH<sub>2</sub>

Comme nous l'avons mentionné précédemment, la présence de deux bandes très décalées vers les basses fréquences dans la région des modes d'élongation NH pour Pro-Phe A suggère fortement que le conformère adopte une structure double coude  $\gamma$ . Cette hypothèse est soutenue par le spectre dans la région amide I qui présente trois bandes CO distinctes, dont deux apparaissent à des fréquences très basses (1652 et 1677 cm<sup>-1</sup>) (Figure 3.21). Celles-ci indiquent la présence de deux liaisons hydrogène fortes compatibles avec une structure  $\gamma$ - $\gamma$ . Le spectre IR de Pro-Phe B dans la région amide A montre trois bandes, dont la plus rouge correspond au mode d'élongation  $\nu_{sym}$  et indique que le groupement NH<sub>2</sub> est impliqué dans une liaison hydrogène forte. Celle-ci correspond probablement à une liaison hydrogène C $\gamma$ . La bande à 3422 cm<sup>-1</sup> pourrait alors être associée à une interaction NH- $\pi$ .

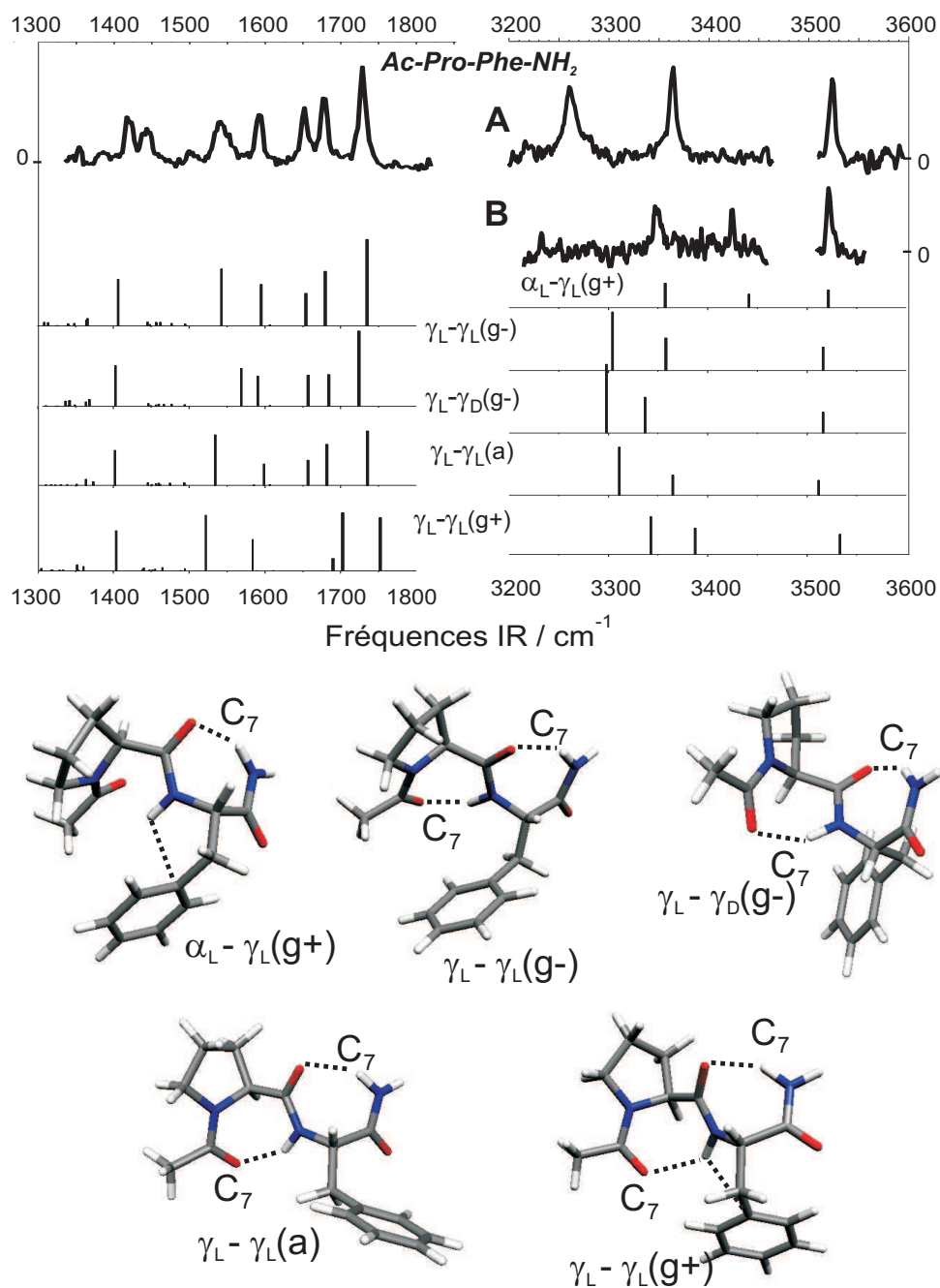


FIG. 3.24 – Spectres IR de Ac-Pro-Phe-NH<sub>2</sub> A comparés aux fréquences calculées pour une sélection de structures. Les fréquences sont corrigées avec le facteur 0.96 dans la région amide A, 0.98 dans la région amide I et 0.97 dans la région amide II.

Nous avons réalisé des calculs de fréquences harmoniques pour des structures  $\gamma$ - $\gamma$ , mais nous avons aussi considéré la structure  $\alpha_L$ - $\gamma_L$  car la forme  $\alpha_L$  est la deuxième forme trouvée comme étant la plus stable pour le résidu Pro isolé [6] par les calculs de chimie quantique (Figure 3.24 et Tableau 3.10). La comparaison entre les spectres expérimentaux de Pro-Phe A et les fréquences théoriques des coudes  $\gamma$ - $\gamma$  montre un accord très satisfaisant. Dans la région amide I, les fréquences théoriques des bandes CO ( $\nu < 1700 \text{ cm}^{-1}$ ) indiquent bien la présence de deux liaisons hydrogène fortes. Dans la région amide II, ce sont les conformations  $\gamma_L$ - $\gamma_L$ (g-) et  $\gamma_L$ - $\gamma_L$ (a) qui semblent le mieux reproduire le spectre de Pro-Phe A. Dans le domaine amide A, les calculs reproduisent correctement la bande située à  $3364 \text{ cm}^{-1}$  qui appartient à la liaison C<sub>7</sub> sur le résidu Phe. En revanche, comme précédemment, l'effet coopératif entre les deux liaisons hydrogène, dû à la polarisation de la liaison amide centrale, n'est pas retrouvé par les calculs. La bande la plus rouge appartenant à la liaison C<sub>7</sub> sur le résidu Pro est surestimée d'une soixantaine de  $\text{cm}^{-1}$ . On peut toutefois souligner que les calculs de fréquences reproduisent bien le fait que le conformère A de Pro-Phe possède la liaison hydrogène la plus forte de la série Ac-Xxx-Phe-NH<sub>2</sub>. Un tel résultat a déjà été observé dans la section précédente où le conformère  $\beta_L$ (a)- $\gamma_L$  du dipeptide Ac-Phe-Pro-NH<sub>2</sub> présentait la liaison hydrogène C<sub>7</sub> la plus forte de la série Ac-Phe-Xxx-NH<sub>2</sub>. De la même façon, cette liaison hydrogène très forte peut s'expliquer par la contrainte induite sur la chaîne principale par le cycle à cinq atomes du résidu Pro. En conséquence, en se référant uniquement aux calculs de fréquences harmoniques, il est difficile d'attribuer une structure précise au conformère Pro-Phe A. Cependant, les calculs d'énergies relatives au niveau MP2/6-31+G(d) single point à partir des géométries optimisées en DFT indiquent que la structure  $\gamma_L$ - $\gamma_L$ (g-) est la forme la plus stable parmi les formes calculées (Tableau 3.10). On est donc tenté d'attribuer le conformère Pro-Phe A à cette structure. Dans le cas du conformère B de Pro-Phe, les fréquences théoriques de la forme  $\alpha_L$ - $\gamma_L$ (g+) donnent un accord correct avec le spectre expérimental. Néanmoins, en raison de l'énergie relativement élevée de la structure, aucune attribution certaine ne peut être proposée. On peut en revanche dire que le résidu Phe adopte très probablement une structure coude  $\gamma$ .

	Ac-Pro-Phe-NH <sub>2</sub>		Structures calculées				
	Conf. A	Conf. B	$\gamma_L\text{-}\gamma_L(g^-)$	$\gamma_L\text{-}\gamma_D(g^-)$	$\gamma_L\text{-}\gamma_L(a)$	$\gamma_L\text{-}\gamma_L(g^+)$	$\alpha_L\text{-}\gamma_L(g^+)$
Fréq.	37541	37515	0.00	2.35	2.56	2.98	2.70
UV					Énergies MP2		
			Fréquences IR				
			Amide A (mode d'élongation NH)				
NH Phe	3261	3422	3304	3298	3311	3395	3442
NH <sub>2</sub> sym	3364	3348	3358	3337	3365	3340	3357
NH <sub>2</sub> anti	3522	3519	3516	3516	3512	3505	3521
			Amide I (mode d'élongation CO)				
CO Pro	1652		1654	1657	1657	1667	
CO Ace	1677		1680	1682	1684	1685	
CO Phe	1729		1741	1736	1724	1725	
			Amide II (mode de déformation NH)				
NH Phe	1541		1543	1534	1568	1526	
NH <sub>2</sub> sym	1592		1595	1599	1591	1599	

TAB. 3.10 – Fréquences expérimentales UV et IR ( $\text{cm}^{-1}$ ) des conformères A et B de Ac-Pro-Phe-NH<sub>2</sub> et fréquences théoriques (B3LYP/6-31+G(d)) d'une sélection de structures. Les fréquences sont corrigées avec le facteur 0.96 dans la région amide A, 0.98 dans la région amide I et 0.97 dans la région amide II. Les énergies relatives (kcal/mol) sont calculées au niveau MP2/6-31+G(d) single point à partir des géométries B3LYP.

Le modèle de couplage vibrationnel (Annexe A) montre que les fréquences  $\nu_{sym}$  et  $\nu_{anti}$  des quatre conformères A de la série Ac-Xxx-Phe-NH<sub>2</sub> apparaissent tous dans la même région montrant que la force de la liaison hydrogène C<sub>7</sub> est quasi identique pour chaque espèce ( $\Delta\nu \sim 160 \text{ cm}^{-1}$ , force de la liaison  $\sim 110 \text{ cm}^{-1}$ ). Les fréquences s'ajustent parfaitement au modèle, ce qui corrobore la formation d'un coude  $\gamma$  sur le résidu Phe. De plus, ces fréquences définissent une nouvelle région pour laquelle la liaison hydrogène impliquant le groupe NH<sub>2</sub> est significativement plus forte que celle des coudes  $\gamma$  simples ( $\beta_L\text{-}\gamma$ ) observés pour la série Ac-Phe-Xxx-NH<sub>2</sub>.

### 3.3.5 Conclusion : effet de l'ordre de la séquence

Tous les conformères A de la série de dipeptides Ac-Xxx-Phe-NH<sub>2</sub> adoptent donc une structure coude  $\gamma$  double, c'est-à-dire une structure  $\gamma\text{-}\gamma$ , suggérant qu'il n'y a pas d'effet de la nature du résidu Xxx sur la conformation générale de la chaîne principale. Les coudes  $\gamma$  formés sur la partie Phe montrent tous des signatures IR très similaires. Les groupements

amides correspondants apparaissent en effet tous à des fréquences d'absorption très proches, de l'ordre de  $1685\text{ cm}^{-1}$  pour les modes d'élongation CO et de l'ordre de  $3360\text{ cm}^{-1}$  pour les modes d'élongation NH. On retrouve ainsi la préférence conformationnelle locale coude  $\gamma$  pour les résidus Gly, Ala, Val et Pro. La formation du coude  $\gamma$  est en revanche plus surprenante pour le résidu Phe car la structure étendue  $\beta_L$  est la forme la plus stable pour le résidu isolé, cette stabilité particulière résultant d'une interaction NH- $\pi$  entre le groupement  $\text{NH}_2$  terminal et le cycle aromatique. En outre, cette interaction est aussi responsable de la formation de la structure  $\beta_L$  pour les conformères  $\beta_L$ - $\gamma$  de la série Ac-Phe-Xxx-NH<sub>2</sub>. En revanche, on s'aperçoit qu'elle n'est pas suffisante dans le cas des dipeptides Ac-Xxx-Phe-NH<sub>2</sub> pour stabiliser la forme  $\gamma$ - $\beta_L$ , qui correspond en principe à la succession des formes locales pour chaque résidu. Afin d'illustrer ce propos, nous avons donc comparé les données énergétiques et géométriques des conformations  $\gamma$ - $\gamma$  et  $\gamma$ - $\beta_L$  pour le dipeptide Ac-Gly-Phe-NH<sub>2</sub>.

Les calculs effectués sur les diverses structures coudes doubles  $\gamma$  ont montré que les distances NH...O des liaisons hydrogène étaient semblables pour tous les coudes  $\gamma$ , de l'ordre de  $2.05\text{ \AA}$ , excepté pour la forme  $\gamma_L$ - $\gamma_L(g+)$  (cf paragraphe précédent), indiquant que les coudes  $\gamma$  sur chaque résidu étaient en compétition. Le coude  $\gamma$  de la conformation  $\gamma_L$ - $\beta_L(a)$ , lui, possède une distance de  $2.33\text{ \AA}$ , ce qui correspond à une liaison beaucoup plus faible que pour les coudes  $\gamma$ - $\gamma$ . En fait, les calculs indiquent que l'interaction C<sub>5</sub> présente une distance NH...O similaire de  $2.32\text{ \AA}$ . Or, les calculs effectués sur la conformation  $\beta_L$ - $\gamma_D$  de Phe-Gly montrent des distances d'interaction significativement plus faibles,  $2.21\text{ \AA}$  pour l'interaction C<sub>5</sub> et  $2.08\text{ \AA}$  pour la liaison C<sub>7</sub>. Cette différence provient du caractère « double donneur » de  $\text{NH}_{Phe}$  qui établit à la fois une interaction C<sub>5</sub> sur le résidu Phe et une liaison C<sub>7</sub> sur le résidu Gly (Figure 3.23). Le coude  $\gamma_L$  est alors déformé [le couple  $(\phi, \psi)$  vaut  $(-86^\circ, 79^\circ)$ , au lieu de  $(-82^\circ, 62^\circ)$  pour la forme  $\gamma_L$ - $\gamma_L(g-)$ ] et il en résulte que ni la liaison C<sub>7</sub>, ni l'interaction C<sub>5</sub> ne sont finalement optimales. Ainsi, si l'on compare les énergies relatives entre la conformation  $\gamma_L$ - $\beta_L(a)$  et diverses conformations  $\gamma$ - $\gamma$ , on constate que la première est notablement déstabilisée par rapport aux secondes d'environ  $3\text{ kcal/mol}$  (Tableau 4.3 en annexe).

### 3.4 Conclusion du chapitre 3

En conclusion, nous nous sommes attachés à montrer dans ce chapitre que de petits peptides protégés pouvaient mimer des petits segments de chaîne peptidique puisque certaines structures typiques de la biologie sont présentes. Nous avons ainsi mis en évidence la formation de structures relativement compactes et de structures moins repliées. Notre étude sur le système « test », l'acide aminé protégé Ac-Phe-NH<sub>2</sub>, a montré que la **préférence conformationnelle locale** du résidu Phe isolé était la forme étendue  $\beta_L$ , forme qu'adoptent les résidus dans un brin  $\beta$ , suivie des formes repliées  $\gamma$ . Nous avons démontré que la stabilité particulière de la structure  $\beta_L$  résultait de la présence d'une interaction NH- $\pi$  entre le cycle aromatique et un groupe amide du squelette peptidique. Une telle structure n'est donc stable que pour des résidus **aromatiques**. Ainsi, cette conclusion reste en accord avec les nombreux travaux qui indiquent que la forme locale des résidus **non aromatiques** est la forme  $\gamma$ .

L'étude de dipeptides nous a permis de constater que les formes locales  $\beta_L$  et  $\gamma$  étaient conservées lorsque l'on associait deux résidus dans la chaîne peptidique. La série de dipeptides Ac-**Phe-Xxx**-NH<sub>2</sub>, pour laquelle nous avons fait varier la nature du résidu hydrophobe Xxx en deuxième position dans la séquence, montre que les conformères majoritaires adoptent tous une structure  $\beta_L(a)$ - $\gamma$ . La chaîne peptidique se comporte donc comme une succession de briques élémentaires interagissant peu entre elles, c'est-à-dire la forme étendue  $\beta_L$  d'une part et le coude  $\gamma$  d'autre part. La comparaison entre les calculs effectués sur ces structures et ceux effectués pour des coudes  $\gamma$  de résidus isolés Ac-Xxx-NHMe a permis de confirmer cette tendance.

Nous avons montré avec l'étude des dipeptides Ac-**Xxx-Phe**-NH<sub>2</sub>, dont l'ordre des résidus dans la séquence est inversé, que les conformères principaux adoptaient une structure  $\gamma$ - $\gamma$ . Comme précédemment, le résidu hydrophobe Xxx conserve sa forme locale repliée  $\gamma$ , mais le résidu Phe adopte ici sa deuxième forme locale. Dans ce cas, les résidus ne sont plus indépendants dans la séquence et conduisent à des effets coopératifs (partage de l'amide central). Néanmoins, les formes  $\gamma_L$  et  $\beta_L$  sont topologiquement très proches dans le diagramme de Ramachandran et appartiennent en fait au même bassin conformationnel, suggérant que les

barrières les séparant ne sont pas très élevées. Le fait que les calculs menés sur Ac-Phe-NH<sub>2</sub> montrent que la structure  $\beta_L(g^-)$  converge vers la structure  $\gamma_L(g^-)$  en est une bonne illustration. On peut par ailleurs mentionner que les coudes  $\gamma_L$  sont trouvés à l'intérieur de feuillets  $\beta$  antiparallèles lorsque ceux-ci subissent des contraintes de compression longitudinale [8]. Les deux séries de dipeptides adoptent donc des structures relativement similaires ( $\beta_L\text{-}\gamma$  et  $\gamma\text{-}\gamma$ ). Nous avons donc montré que l'ordre des résidus dans la séquence influençait finalement peu la conformation globale de la chaîne peptidique puisque celle-ci adopte une structure dite « ouverte », c'est-à-dire pour laquelle ses extrémités N- et C-terminales ne sont pas reliées par une liaison hydrogène.

Enfin, il faut souligner que malgré la complexité du paysage conformationnel, nous n'avons observé que quelques espèces significativement peuplées pour les deux séries de dipeptides. Ce chapitre s'est ainsi focalisé sur les formes majoritaires,  $\beta_L\text{-}\gamma$  et  $\gamma\text{-}\gamma$ , mais d'autres conformères ont également été identifiés. Ils font l'objet du chapitre suivant.

**Références bibliographiques du Chapitre 3**

- [1] G. Endrédi, A. Perczel, O. Farkas, M. A. McAllister, G. I. Csonka, J. Ladik and I. G. Csizmadia, *J. Molec. Struct.* **391**, 15 (1997).
- [2] K. Möhle, M. Gussman, A. Rost, R. Cimiraglia and H.-J. Hofmann, *J. Phys. Chem. A* **101**, 8571 (1997).
- [3] M. D. Beachy, D. Chasman, R. B. Murphy, T. A. Halgern and R. A. Friesner, *J. Am. Chem. Soc.* **119**, 5908 (1997), et réf. citées.
- [4] K. J. Jalkanen, M. Elstner and S. Suhai, *J. Mol. Struct. (Theochem)* **675**, 61 (2004), et réf. citées.
- [5] R. Vargas, J. Garza, B. P. Hay and D. Dixon, *J. Phys. Chem. A* **106**, 3213 (2002).
- [6] C. Benzi, R. Improta, G. Scalmani and V. Barone, *J. Comput. Chem.* **23**, 341 (2002), et réf. citées.
- [7] Y. K. Kang, *J. Mol. Struct. (Theochem)* **675**, 37 (2004), et réf. citées.
- [8] E. Milner-White, B. R. Moss, R. Ismail, K. Belhjad-Mostefa and R. Poet, *J. Mol. Biol.* **204**, 777 (1988).
- [9] E. Milner-White, *J. Mol. Biol.* **216**, 385 (1990).
- [10] R. J. Lavrich, D. F. Plusquellic, R. D. Suenram, G. T. Fraser, A. R. H. Walker and M. J. Tubergen, *J. Chem. Phys.* **118**, 1253 (2003).
- [11] I. Compagnon, J. Oomens, J. Bakker, G. Meijer and G. von Helden, *Phys. Chem. Chem. Phys.* **7**, 13 (2005).
- [12] A. Perczel, O. Farkas, A. G. Császár and I. G. Csizmadia, *Can. J. Chem.-Rev. Can. Chim.* **75**, 1120 (1997).
- [13] E. Langella, N. Rega, R. Improta, O. Crescenzi and V. Barone, *J. Comput. Chem.* **23**, 650 (2002).
- [14] E. Langella, N. Rega, R. Improta and V. Barone, *J. Am. Chem. Soc.* **124**, 11531 (2002).
- [15] G. A. Chass, S. Lovas, R. F. Murphy and I. G. Csizmadia, *Eur. Phys. J. D.* **20**, 481 (2002).



- [16] B. C. Dian, A. Longarte and T. S. Zwier, *Science* **296**, 2369 (2002), et ref. citées.
- [17] T. S. Zwier, *J. Phys. Chem. A* **105**, 8827 (2001), et ref. citées.
- [18] M. Schauer, K. Law and E. Bernstein, *J. Chem. Phys.* **81**, 49 (1981).
- [19] E. G. Robertson, M. R. Hockridge, P. D. Jelfs and J. P. Simons, *Phys. Chem. Chem. Phys.* **3**, 786 (2001).
- [20] E. G. Robertson, M. R. Hockridge, P. D. Jelfs and J. P. Simons, *J. Phys. Chem. A* **?**, ? (2001).
- [21] P. Hudáky, I. Jákli, A. G. Császár and A. Perczel, *J. Comput. Chem.* **22**, 732 (2001), et réf. citées.
- [22] G. A. Chass, R. S. Mirasol, D. H. Setiadi, T.-H. Tang, W. Chin, M. Mons, I. Dimicoli, J.-P. Dognon, B. Viskolcz, S. Lovas, B. Penke and I. G. Csizmadia, *J. Phys. Chem. A* **109**, 5289 (2005).
- [23] G. Toth, C. R. Watts, R. F. Murphy and S. Lovas, *Proteins* **43**, 373 (2001).
- [24] B. C. Dian, A. Longarte, S. Mercier, D. A. Evans, D. J. Wales and T. S. Zwier, *J. Chem. Phys.* **117**, 10688 (2002).
- [25] M. Gerhards, C. Unterberg, A. Gerlach and A. Jansen, *Phys. Chem. Chem. Phys.* **6**, 2682 (2004).
- [26] C. Unterberg, A. Gerlach, T. Schrader and M. Gerhards, *Eur. Phys. J. D* **20**, 543 (2002).
- [27] M. Gerhards, C. Unterberg and A. Gerlach, *Phys. Chem. Chem. Phys.* **4**, 5563 (2002).
- [28] M. Gerhards and C. Unterberg, *Phys. Chem. Chem. Phys.* **4**, 1760 (2002).
- [29] S. K. Burley and G. A. Petsko, *FEBS Lett.* **203**, 139 (1986).
- [30] M. Levitt and M. F. Perutz, *J. Mol. Biol.* **201**, 751 (1988).
- [31] T. Steiner, *Angew. Chem., Int. Ed.* **41**, 48 (2002).
- [32] R. F. Frey, J. Coffin, S. Q. Newton, M. Ramek, V. K. W. Cheng, F. A. Momany and L. Schafer, *J. Am. Chem. Soc.* **114**, 5369 (1992).
- [33] W. Viviani, J.-L. Rivail, A. Perczel and I. G. Csizmadia, *J. Am. Chem. Soc.* **115**, 8321 (1993).

- [34] R. A. IUPAC-IUB Commission on Biochemical nomenclatur 1970., *J. Mol. Biol.* **52**, 1 (1970).
- [35] A. D. Rabuck and G. E. Scuseria, *Theor. Chem. Acc.* **104**, 439 (2000).
- [36] A. K. Rappé and E. R. Bernstein, *J. Phys. Chem. A* **104**, 6117 (2000).
- [37] R. Vargas, J. Garza, D. Dixon and B. P. Hay, *J. Am. Chem. Soc.* **122**, 4750 (2000).
- [38] J. Ireta, J. Neugebauer and M. Sceffler, *J. Phys. Chem. A* **108**, 5692 (2004).



## Chapitre 4

# Coudes $\beta$

Les coudes  $\beta$  (ou  $\beta$ -turns) sont les premières structures secondaires pouvant se former dans de petites chaînes peptidiques, en l'occurrence des dipeptides protégés. Ce sont les structures secondaires non périodiques les plus fréquentes dans les protéines qui permettent le changement de direction de la chaîne peptidique, donnant ainsi à la protéine sa forme plus ou moins compacte. Les coudes  $\beta$  sont constitués de quatre résidus définis par les positions  $i, i+1, i+2, i+3$  dans la chaîne et présentent une liaison hydrogène  $C_{10}$  reliant les groupes  $CO(i) \leftarrow NH(i+3)$  dans un cycle à dix atomes [1]. Ils sont classés par « types » et sont définis par les angles  $\phi$  et  $\psi$  des résidus en position  $i+1$  et  $i+2$ . Une brève description de ces types est donnée dans l'introduction générale mais on pourra trouver dans la revue de Rose et collaborateurs une description plus détaillée [2].

Les calculs de chimie quantique sur les coudes  $\beta$  en phase gazeuse sont très nombreux ; toutefois, suivant la méthode de calcul utilisée, ils aboutissent parfois à des conclusions différentes : les calculs d'énergie relative de petits peptides protégés effectués avec la méthode HF favorisent la préférence conformationnelle locale sur chaque résidu, la forme double  $\gamma$  [3, 4] ; les calculs en DFT stabilisent la structure étendue  $\beta_L$  équivalente au brin  $\beta$  [5] ; les calculs réalisés au niveau MP2 trouvent en revanche que les coudes  $\beta$  deviennent les minimas les plus stables [6].

Toutefois, il n'existe pour l'instant aucune contrepartie expérimentale en phase gazeuse pouvant valider ces nombreux résultats théoriques. Nous nous proposons donc d'appliquer les

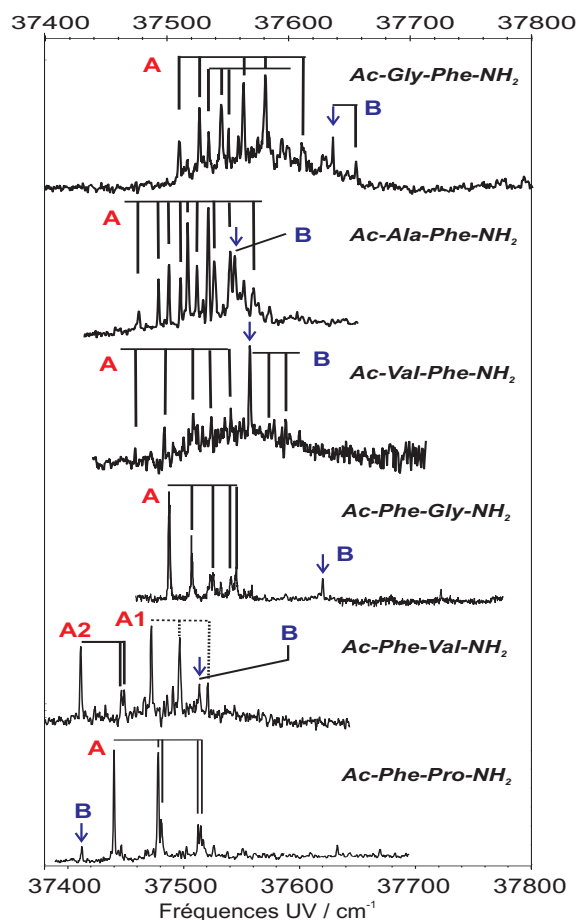


FIG. 4.1 – Spectres d’excitation des systèmes  $\text{Ac-Xxx-Phe-NH}_2$  ( $\text{Xxx} = \text{Gly, Ala, Val}$ ) et  $\text{Ac-Phe-Xxx-NH}_2$  ( $\text{Xxx} = \text{Gly, Val, Pro}$ ). Les bandes origines des conformères B sont indiquées par des flèches.

techniques de double résonance IR-UV à la recherche des coudes  $\beta$ ; celles-ci se sont révélées très efficaces pour la caractérisation des dipeptides  $\text{Ac-Phe-Xxx-NH}_2$  et  $\text{Ac-Xxx-Phe-NH}_2$  présentés dans le chapitre précédent puisque nous avons pu mettre en évidence la formation de structures qui correspondaient à la succession des préférences conformationnelles locales sur chaque résidu pour ces deux séries de dipeptides. Ces structures se sont avérées être les formes les plus peuplées. Toutefois, un autre conformère, noté B, a également été observé dans les deux séries. Ce chapitre sera donc consacré à la caractérisation de ces conformères B.

## 4.1 Spectroscopie électronique

En appliquant les techniques de double résonance IR-UV, nous avons donc identifié un deuxième conformère B dont la signature spectroscopique était différente de celle des conformères A. Rappelons que le système Phe-Ala ne présente qu'un seul conformère  $\beta_L(a)$ - $\gamma_L$  et que le conformère B de Pro-Phe a été attribué dans le chapitre précédent à une structure coude  $\gamma$  simple  $\alpha_L$ - $\gamma_L$ . Les spectres d'excitation des conformères B montrent des bandes UV isolées ou, dans le cas de Gly-Phe et Val-Phe, présentant une progression vibrationnelle courte (Figure 4.1). La spectroscopie électronique étant très différente des espèces A qui montraient une forte activité Franck-Condon, les espèces B adoptent vraisemblablement des structures différentes de A. Les bandes origine sont décalées vers les hautes fréquences par rapport aux origines des conformères A, hormis pour Phe-Pro dont l'origine de B est plus rouge.

## 4.2 Signature IR

Dans ce paragraphe, nous discuterons d'abord simultanément des résultats obtenus pour les deux séries de dipeptides Ac-Phe-Xxx-NH<sub>2</sub> et Ac-Xxx-Phe-NH<sub>2</sub>. Nous verrons que les signatures IR des modes d'élongation NH et l'utilisation du modèle de couplage vibrationnel nous permettront de proposer une première attribution. Dans des parties distinctes, nous nous appuierons ensuite sur les calculs - fréquences et énergies relatives - effectués sur les systèmes Phe-Gly et Gly-Phe, qui sont les systèmes les plus simples et les moins coûteux en temps de calculs. Les spectres IR de Gly-Phe et Val-Phe des modes d'élongation CO seront accompagnés de calculs de fréquences, afin de conforter l'attribution déduite de la spectroscopie amide A. Enfin, nous traiterons dans une dernière partie de la difficulté d'interpréter les résultats aussi bien expérimentaux que théoriques en nous intéressant à l'effet éventuel de la nature du résidu sur la conformation générale de la chaîne. Ceci nous amènera à discuter les travaux réalisés sur le modèle Ac-Val-Tyr(Me)-NHMe du groupe de Gerhards à Düsseldorf.

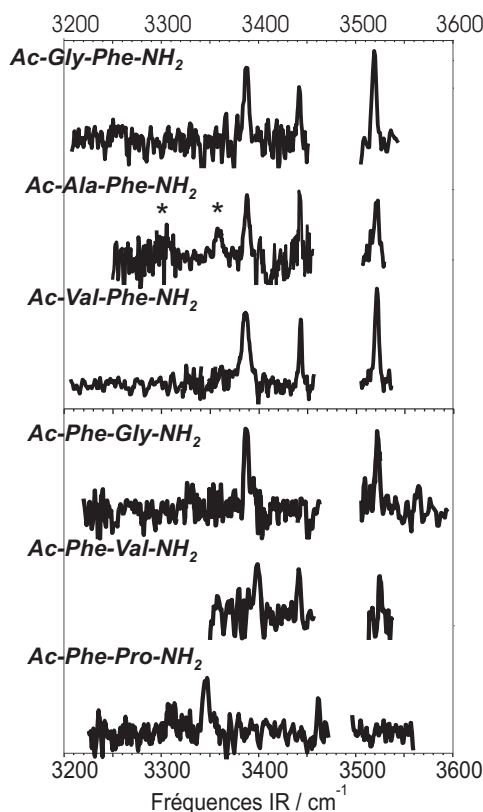


FIG. 4.2 – Spectres de dépopulation IR-UV des conformères B des dipeptides  $\text{Ac-Xxx-Phe-NH}_2$  ( $\text{Xxx} = \text{Gly, Ala, Val}$ ) et  $\text{Ac-Phe-Xxx-NH}_2$  ( $\text{Xxx} = \text{Gly, Val, Pro}$ ) dans la région amide A.

#### 4.2.1 Amide A

Les spectres IR dans le domaine amide A montrent cinq bandes pour Ala-Phe, trois bandes pour Gly-Phe, Val-Phe et Phe-Val, et deux bandes pour Phe-Gly et Phe-Pro (Figure 4.2). Excepté pour Phe-Pro dont la bande d'élongation  $\text{NH}_2$  antisymétrique est absente, les bandes symétrique et antisymétrique apparaissent aux mêmes fréquences  $\nu_{\text{sym}} \sim 3390 \text{ cm}^{-1}$  et  $\nu_{\text{anti}} \sim 3520 \text{ cm}^{-1}$  pour tous les dipeptides. Leur position indique que le groupement  $\text{NH}_2$  est impliqué dans une liaison hydrogène de force modérée, rappelant celle des conformères  $\beta_L(\text{a})-\gamma$  de  $\text{Ac-Phe-Xxx-NH}_2$  (voir Chapitre 3). La faible absorption centrée à  $3308 \text{ cm}^{-1}$  ainsi que la bande d'intensité moyenne à  $3360 \text{ cm}^{-1}$  dans le spectre de Ala-Phe indiquées par des astérisques sur le spectre) sont attribuées à des absorptions du conformère A ; les bandes UV de A et B se recouvrent en effet partiellement (Figure 4.1). Tous les spectres, excepté celui de Phe-Gly, montrent une bande IR située à  $\sim 3440 \text{ cm}^{-1}$  pouvant correspondre à une interaction  $\text{C}_5$

ou NH- $\pi$ . La bande manquante située dans la région non couverte par le cristal OPO - on devrait normalement mesurer quatre bandes, trois pour Phe-Pro - indique que les dipeptides possèdent un groupement amide libre. Dans le cas de Phe-Gly, on peut en déduire que les deux amides NH<sub>Phe</sub> et NH<sub>Gly</sub> sont libres. Le spectre IR de Phe-Pro B ne présente que deux bandes, la bande correspondant au mode d'élongation symétrique et celle de l'amide central NH<sub>Phe</sub>. Le décalage très marqué vers les basses fréquences de la bande symétrique (3342 cm<sup>-1</sup>) et l'absence du mode antisymétrique révèlent que le groupe NH<sub>2</sub> est impliqué dans une liaison hydrogène très forte.

Conformères B	Ac-Phe-Xxx-NH <sub>2</sub>			Ac-Xxx-Phe-NH <sub>2</sub>		
	Phe-Gly	Phe-Val	Phe-Pro	Gly-Phe	Ala-Phe	Val-Phe
Fréquences UV	37620	37513	37413	37636	37560	37569
Fréquences IR expérimentales						
Amide A (mode d'élongation NH)						
NH <sub>2</sub> sym	3387	3398	3345	3388	3389 3308* 3360*	3395
NH Phe	a	a	3460	3442	3443	3443
NH Xxx	a	3442	a	a	a	a
NH <sub>2</sub> anti	3522	3525	a	3519	3523	3521
Amide I (mode d'élongation CO)						
CO Ace				1687		1694
CO Phe				1707		1721
CO Xxx				1719		
Amide II (mode de déformation NH)						
NH Xxx				1504		1501
NH <sub>2</sub> sym				1588		1595

TAB. 4.1 – Fréquences UV et IR des conformères B des dipeptides Ac-Phe-Xxx-NH<sub>2</sub> (Xxx = Gly, Val, Pro) et Ac-Xxx-Phe-NH<sub>2</sub> (Xxx = Gly, Ala, Val). Les fréquences marquées d'une astérisque pour Ala-Phe correspondent à des bandes du conformère A. a : les bandes manquantes sont situées dans la région spectrale non couverte par le cristal de l'OPO.

Pour chaque série de dipeptides, deux structures sont compatibles avec la liaison hydrogène observée : un coude  $\gamma$  sur le résidu en deuxième position reliant les groupes CO<sub>Xxx/Phe</sub> et NH<sub>2</sub>



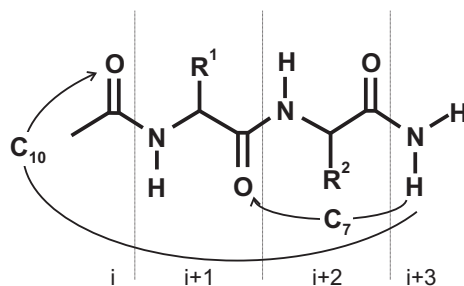


FIG. 4.3 – Schéma des liaisons hydrogène pouvant reproduire la liaison hydrogène impliquant  $\text{NH}_2$ .

dans une liaison  $\text{C}_7$ , ou bien un coude  $\beta$  reliant les groupes  $\text{CO}_{\text{Ace}}$  et  $\text{NH}_2$  dans une liaison hydrogène  $\text{C}_{10}$  (Figure 4.3).

#### 4.2.2 Utilisation du modèle de couplage vibrationnel

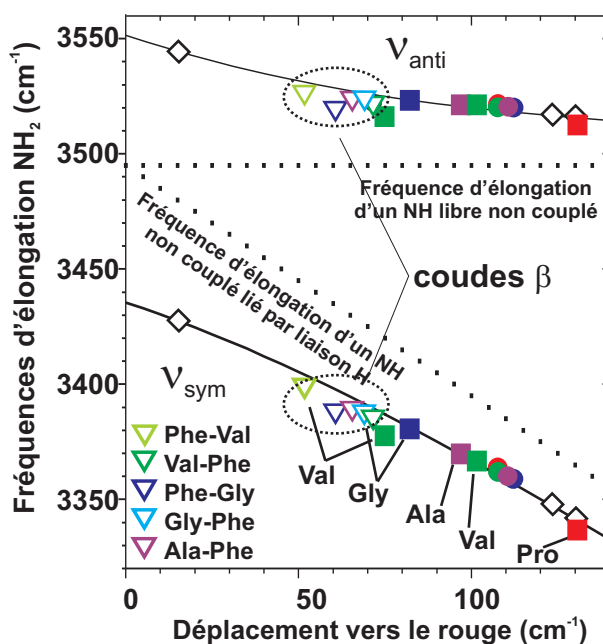


FIG. 4.4 – Modèle de couplage vibrationnel du groupement  $\text{NH}_2$ . Les coudes  $\beta$  sont symbolisés par des triangles inversés, les losanges correspondent aux conformères du système  $\text{Ac-Phe-NH}_2$ , les autres symboles correspondent aux conformères  $\beta_{L-\gamma}$  et  $\gamma-\gamma$  des séries  $\text{Ac-Phe-Xxx-NH}_2$  et  $\text{Ac-Xxx-Phe-NH}_2$ .

D'après le modèle de couplage vibrationnel exposé dans le Chapitre 3 (ou en Annexe ??), les fréquences d'élongation  $\text{NH}_2$  symétrique et antisymétrique des conformères B (en triangles sur la Figure 4.4) indiquent que la liaison hydrogène impliquant le groupe  $\text{NH}_2$  est plus faible que la liaison  $\text{C}_7$  des conformères  $\beta_{L(a)-\gamma}$ . Les fréquences ainsi que l'écart  $\Delta\nu$  définissent en

fait une région de force de liaison hydrogène distincte des formes  $\gamma$ - $\gamma$  et coudes  $\gamma$  simples : alors que le déplacement vers le rouge, indicateur de la force de la liaison hydrogène, est supérieur à  $80 \text{ cm}^{-1}$  pour les conformères  $\beta_L(a)$ - $\gamma$  (en carrés sur la Figure) et est de l'ordre de  $110 \text{ cm}^{-1}$  pour les conformères  $\gamma$ - $\gamma$  (en disques sur la Figure 4.4), il est évaluée à environ  $60 \text{ cm}^{-1}$  pour les conformères B. En outre, contrairement aux coudes  $\gamma$  dont les liaisons  $C_7$  sont très bien décrites par le modèle, les fréquences correspondant aux conformères B s'écartent de façon significative des courbes du modèle. Rappelons que celui-ci a été paramétré à partir des coudes  $\gamma_L(g+)$  et  $\gamma_L(g-)$  du système Ac-Phe-NH<sub>2</sub>, ce qui explique le bon accord entre les conformères  $\beta_L$ - $\gamma$  et  $\gamma$ - $\gamma$  des dipeptides Ac-Phe-Xxx-NH<sub>2</sub> et Ac-Xxx-Phe-NH<sub>2</sub> et les courbes du modèle. En revanche, le fait que le modèle ne décrive pas correctement la liaison hydrogène du groupe NH<sub>2</sub> pour les conformères B suggère fortement que ceux-ci n'adoptent pas une conformation  $\gamma$  mais une autre conformation caractérisée par une force de liaison plus faible telle que la liaison  $C_{10}$  de la conformation coude  $\beta$ . On peut dès lors écarter la possibilité de former une structure  $\beta_L$  qui n'est pas compatible avec la formation d'un coude  $\beta$  et attribuer la bande située à environ  $3440 \text{ cm}^{-1}$  à une interaction NH- $\pi$  avec le cycle phényle.

### 4.2.3 Amides I et II

Les spectres IR dans la région des amides I et II des conformères B n'ont été mesurés que pour les dipeptides Gly-Phe et Val-Phe. Les spectres sont relativement similaires, ils présentent deux bandes larges dans le domaine amide I avec un épaulement vers le bleu pour Gly-Phe (représenté par une flèche), et des bandes peu résolues dans le domaine amide II (Figure 4.5). La position des bandes CO confirme la présence d'une seule liaison hydrogène. Nous avons comparé les spectres de Gly-Phe B et Val-Phe B à ceux des conformères A  $\gamma$ - $\gamma$  et  $\beta_L(a)$ - $\gamma_L$ , nous avons ici considéré les systèmes Val-Phe et Phe-Val : on constate que le profil de bandes CO des conformères B ne correspond pas à celui de coudes  $\gamma$  car il est compact et centré à  $1710 \text{ cm}^{-1}$ , alors que les bandes CO des coudes  $\gamma$  sont très décalées vers les basses fréquences ( $\nu < 1680 \text{ cm}^{-1}$ ). De plus l'écart expérimental entre les bandes extrêmes pour les coudes  $\gamma$  est en général de l'ordre de  $50 \text{ cm}^{-1}$ , ce qui est beaucoup plus grand que ce qui est mesuré pour les conformères B. En s'appuyant sur la spectroscopie IR dans les régions amide A et I et II,

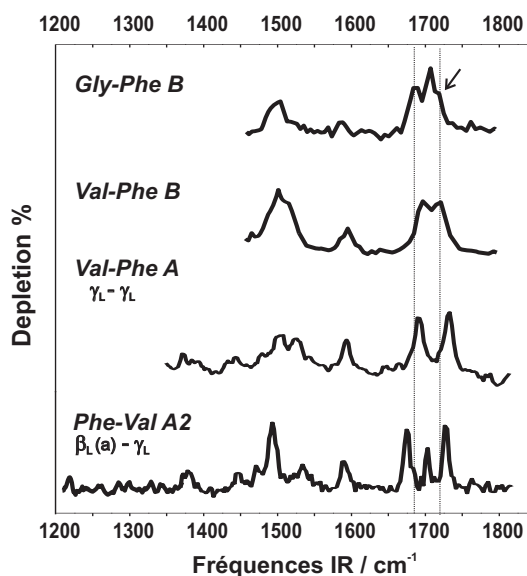


FIG. 4.5 – Spectres IR dans la région amide I et II des conformères B de Ac-Gly-Phe-NH<sub>2</sub> et Ac-Val-Phe-NH<sub>2</sub> et spectres IR des conformères A  $\gamma_L$ - $\gamma_L$  de Ac-Val-Phe-NH<sub>2</sub> et  $\beta_L$ - $\gamma_L$  de Ac-Phe-Val-NH<sub>2</sub>, tirés du chapitre 3.

on peut par conséquent attribuer les conformères Gly-Phe B et Val-Phe B à des coudes  $\beta$ .

#### 4.2.4 Comparaison avec les calculs

Afin de confirmer cette attribution, nous avons comparé les fréquences des conformères B aux fréquences calculées pour différentes structures. Les calculs les plus complets ont été menés sur les systèmes Phe-Gly et Gly-Phe en raison de leur petite taille, ce qui permet de réduire les temps de calculs. Ont été considérés les structures secondaires coudes  $\beta$  (C<sub>10</sub>) avec les types I, I', II et II', et le brin  $\beta$   $\beta_L$ - $\beta_L$ , et des formes correspondant à la succession de formes locales pour chaque résidu présentant une liaison hydrogène, c'est-à-dire la forme  $\beta_L$ - $\gamma_L$  (C<sub>5</sub>-C<sub>7</sub>) dont le coude  $\gamma$  peut aussi former la liaison hydrogène sur le groupe NH<sub>2</sub> terminal et et la forme  $\gamma_L$ - $\beta_L$  (C<sub>7</sub>-C<sub>5</sub>). Les orientations du cycle aromatique ont été choisies afin de prendre en compte les interactions faibles NH- $\pi$  et C<sub>5</sub>. Les calculs de géométrie et de fréquences ont été réalisés par la méthode B3LYP/6-31+G(d) et les énergies relatives ont été calculées au niveau MP2/6-31+G(d) mais aussi avec une plus grande base 6-311+G(d,p) dans le cas du système Gly-Phe afin d'évaluer l'effet de la taille de la base. Les énergies relatives de Phe-Gly et Gly-Phe sont regroupées dans les Tableaux 4.4 et 4.3 en Annexe.

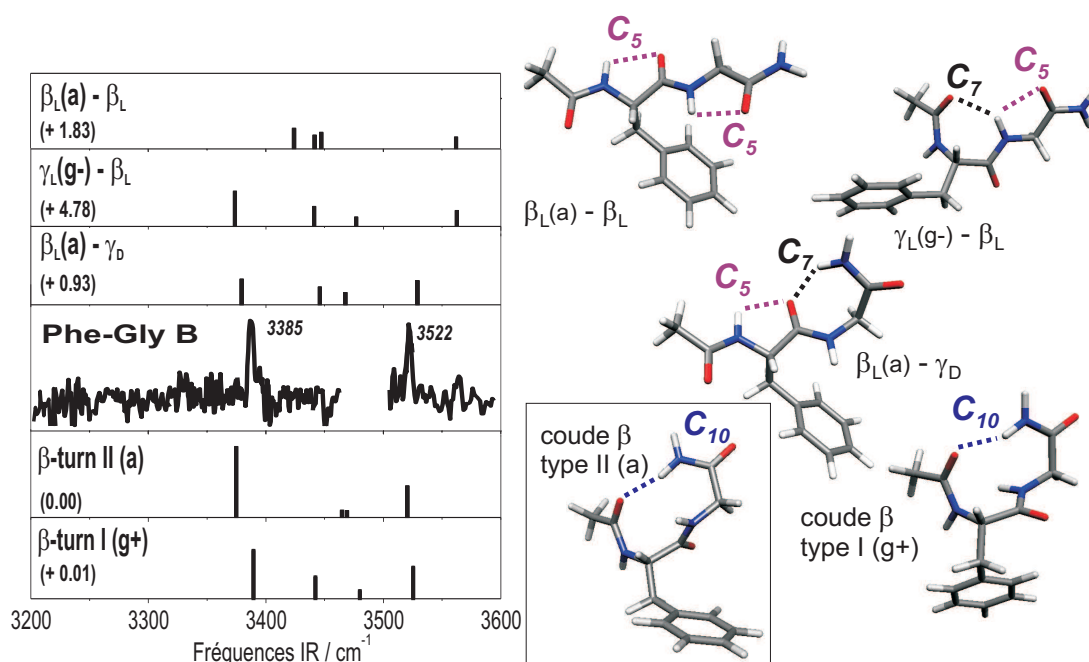


FIG. 4.6 – Spectre IR du conformère B de Ac-Phe-Gly-NH<sub>2</sub> comparé aux fréquences calculées (B3LYP/6-31+G(d)) pour les structures  $\beta_L(a)-\beta_L$ ,  $\gamma_L(g^-)-\beta_L$ ,  $\beta_L(a)-\gamma_D$  et deux coudes  $\beta$  type II (a) et type I (g+). Les fréquences théoriques sont corrigées avec le facteur 0.96 pour tenir compte de l'anharmonicité, les énergies relatives calculées au niveau MP2 sont exprimées en kcal/mol.

### Ac-Phe-Xxx-NH<sub>2</sub>

**Phe-Gly** Nous avons comparé le spectre IR de Phe-Gly dans la région amide A avec les conformations les plus stables dans chaque famille de structures (Figure 4.6). Comme on peut s'y attendre, les formes  $\beta_L-\beta_L$  et  $\gamma_L-\beta_L$  ne conviennent pas car les fréquences d'élongation antisymétrique correspondent à des groupes NH<sub>2</sub> libres. Le spectre théorique de la forme  $\beta_L-\gamma_D$  peut convenir, en particulier pour la fréquence de la liaison C<sub>7</sub>. En revanche, l'absence de bande dans la zone 3420-3460 cm<sup>-1</sup> où apparaît le NH peptidique de l'interaction C<sub>5</sub>, qui est mesurable expérimentalement, montre clairement que la forme  $\beta_L-\gamma$  n'est pas compatible avec ce qui est observé. En outre, l'absence de progression vibrationnelle dans le spectre UV est un argument supplémentaire pour écarter une telle structure.

Un très bon accord est obtenu avec les coudes  $\beta$ . Le spectre IR de Phe-Gly ne présentant pas de bandes NH dans la région des interactions NH- $\pi$  (et C<sub>5</sub>), on ne doit considérer que les structures dont les groupes NH centraux sont libres. Parmi les coudes les plus stables répondant à ce critère, les coudes  $\beta$  type I' (a), type II (g-) et type II (a) peuvent correspondre

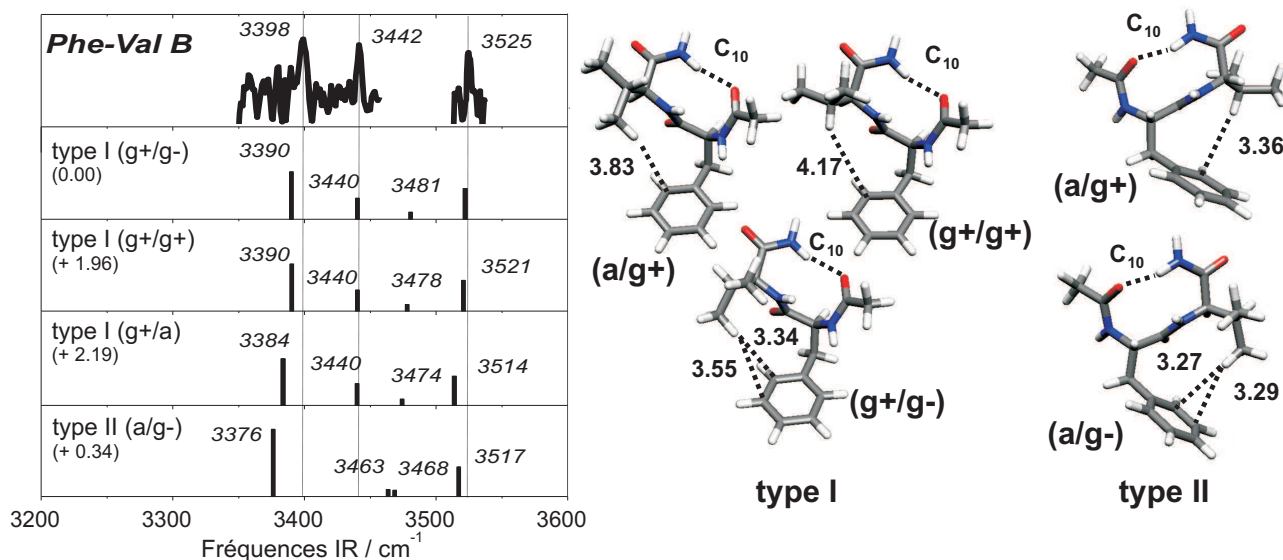


FIG. 4.7 – Comparaison entre le spectre IR de Phe-Val B dans la région amide A avec les fréquences théoriques (corrigées avec le facteur 0.96) des coudes  $\beta$  type I et type II calculées avec la méthode B3LYP/6-31+G(d). Les énergies relatives MP2/6-31+G(d) single point sont exprimées en kcal/mol.

au conformère B de Phe-Gly. Néanmoins, le type II (a) semble être le meilleur candidat car c'est le minimum global, les deux autres coudes apparaissent à 5.22 et 1.09 kcal/mol au-dessus du minimum absolu.

**Phe-Val** Les calculs réalisés sur le système Phe-Gly ont trouvé que les coudes  $\beta$  de type II (a), I (g+) et I (g-) étaient les plus stables parmi les coudes  $\beta$  calculés. Par analogie avec Phe-Gly, nous avons sélectionné les coudes type I (g+) et type II (a) pour les calculs de fréquences harmoniques sur le système Phe-Val. Dans la configuration type I, seule l'orientation (g+) du cycle phényle mène à une interaction NH- $\pi$  reproduisant la bande expérimentale à 3442  $\text{cm}^{-1}$ . Afin de déterminer si la chaîne latérale de Val avait une influence sur le spectre vibrationnel du système Val-Phe, les trois orientations de Val ont été prises en compte pour le type I (g+). Les calculs des coudes  $\beta$  de type I reproduisent correctement les fréquences expérimentales (Figure 4.7). Le type II (a/g-) indique des fréquences décalées d'une vingtaine de  $\text{cm}^{-1}$  vers le rouge, néanmoins l'écart avec la valeur mesurée est encore acceptable. En revanche, le type II (a/g-) ne reproduit pas l'interaction NH- $\pi$  observée mais indique que les deux amides centraux sont libres. Ce coude  $\beta$  peut donc être écarté.

On constate que l'orientation de la chaîne latérale de Val a peu d'influence sur les fréquences. À l'inverse, les stabilités relatives sont beaucoup plus sensibles à l'orientation. La forme la plus stable est le coude type I (g+/g-) pour laquelle la chaîne latérale de Phe a une orientation (g+) et celle de Val une orientation (g-). On remarque que le groupe isopropyle de Val interagit avec le cycle aromatique, les plus courtes distances entre les deux chaînes latérales sont de l'ordre de 3.40 Å. Ces distances sont plus grandes pour les orientations (a) et (g+) de la chaîne latérale de Val, ce qui conduit à des énergies plus élevées. Le coude  $\beta$  type II (a/g-) est aussi très stable. Comme pour le type I, et comme pour les deux conformères A1 et A2 de Phe-Val qui adoptent des structures  $\beta_L(a)-\gamma_L$ , on peut supposer que sa stabilité est due à des interactions entre les chaînes latérales de Phe et Val, les distances d'interaction CH...C sont de 3.27 Å. Cette interaction de dispersion est plus faible pour le type II (a/g+), qui se trouve de ce fait beaucoup plus haut en énergie (+ 7.12 kcal/mol).

Par conséquent, le conformère B de Phe-Val est compatible avec le coude  $\beta$  type I (g+/g-). À ce stade, une attribution plus précise ne peut être proposée, eu égard à l'insuffisance de l'exploration du paysage conformationnel.

**Phe-Pro** Contrairement à la majorité des liaisons peptidiques qui ont une configuration *trans*, le résidu Pro peut présenter une configuration *cis*. En raison de la contrainte exercée sur la chaîne par le cycle à cinq atomes de Pro, l'isomérisation *cis/trans* est en effet fréquemment observée pour ce résidu [7, 8]. La configuration *cis* peut ainsi mener à la formation d'un coude  $\beta$  type VI spécifique au résidu Pro. Le spectre IR de Phe-Pro B est assez singulier en raison de l'absence de la bande NH<sub>2</sub> antisymétrique, probablement située dans la région non accessible au laser OPO. Pareillement aux autres dipeptides, doivent être prises en compte les structures possédant une seule liaison hydrogène : la structure  $\beta_L(a)-\gamma_L$ , les coudes  $\beta$  type I et type VI avec une orientation (g+) du phényle afin de reproduire l'interaction NH- $\pi$ . La structure  $\beta_L(a)-\gamma_L$  qu'adopte le conformère Phe-Pro A peut être écartée, le spectre électronique de Phe-Pro B ne montrant pas de progression vibrationnelle. Le coude  $\beta$  type I a une énergie très élevée par rapport au minimum  $\beta_L(a)-\gamma_L$ . De plus, les fréquences théoriques ne rendent pas compte du spectre IR de Phe-Pro B : la position des bandes expérimentales NH<sub>2</sub> symétrique

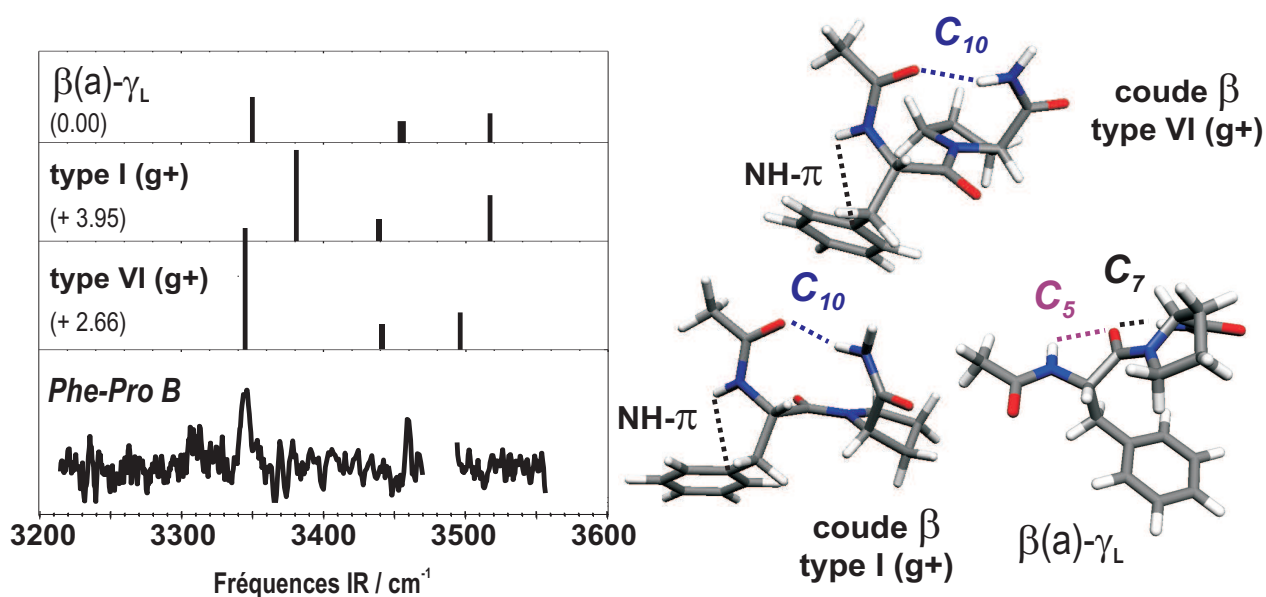


FIG. 4.8 – Comparaison entre le spectre IR de Phe-Pro B dans la région amide A avec les fréquences théoriques (corrigées avec le facteur 0.96) des structures  $\beta_L(a)-\gamma_L$  et coudes  $\beta$  type I (g+) et type VI (g+) calculées avec la méthode B3LYP/6-31+G(d). Les énergies relatives MP2/6-31+G(d) single point sont exprimées en kcal/mol.

et antisymétrique est surestimée (d'une quarantaine de  $\text{cm}^{-1}$  pour la composante symétrique). L'accord entre les fréquences théoriques du coude  $\beta$  type VI (g+) est en revanche très satisfaisant. En outre, la configuration *cis* de l'amide central conduit à une liaison hydrogène  $\text{C}_{10}$  très différente de celles observées et calculées jusqu'à présent, ce dont témoigne la fréquence antisymétrique très décalée vers le rouge. Ce décalage peut expliquer l'absence de cette bande IR dans le spectre expérimental. Malgré une énergie relativement haute, le conformère B de Phe-Pro adopte probablement une conformation coude  $\beta$  type VI pour laquelle l'amide  $\text{NH}_{Phe}$  interagit avec le cycle aromatique dans une interaction  $\text{NH}-\pi$ .

En outre, on peut remarquer que la bande origine UV de Phe-Pro B est très décalée vers les basses fréquences par rapport aux autres coudes  $\beta$ . En effet, les coudes  $\beta$  des systèmes Ala-Phe, Val-Phe et Phe-Val absorbent vers  $37550 \text{ cm}^{-1}$ . Ceux des systèmes Phe-Gly et Gly-Phe absorbent dans la région la plus bleue, au delà de  $37600 \text{ cm}^{-1}$ . Or, le coude  $\beta$  de Phe-Pro présente la bande origine la plus rouge de toute la série. Ceci tend à confirmer que le type de coude  $\beta$  est vraisemblablement différent de ceux des autres systèmes, c'est-à-dire qu'il ne s'agit probablement pas des types I, II ou II', et conforte l'attribution au type VI.

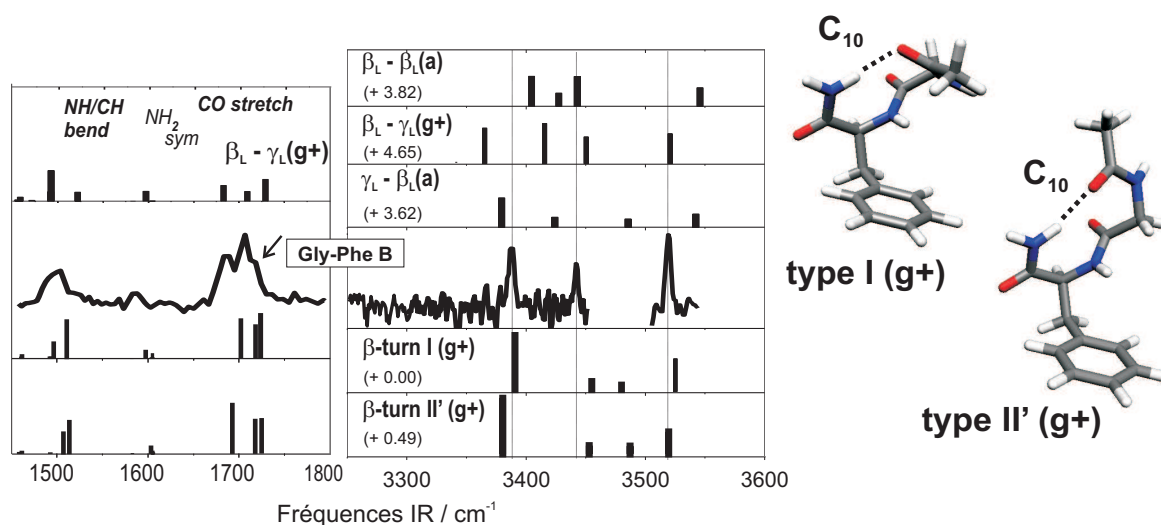


FIG. 4.9 – Comparaison entre le spectre IR de Gly-Phe B avec les fréquences calculées dans la région amide A pour les structures  $\gamma_L$ - $\gamma_L$ (g-), bL-bL(a),  $\beta_L$ - $\gamma_L$ (g+),  $\gamma_L$ - $\beta_L$ (a) et les coudes  $\beta$  de type I (g+) et type II' (g+). Les fréquences sont corrigées avec un facteur 0.96. Géométries B3LYP/6-31+G(d) des coudes  $\beta$  type I (g+) et type II' (g+).

### Ac-Xxx-Phe-NH<sub>2</sub>

**Gly-Phe** Les calculs effectués sur Gly-Phe ont également porté sur plusieurs conformations, les structures  $\beta_L$ - $\beta_L$ (a),  $\beta_L$ - $\gamma_L$ (g+),  $\gamma_L$ - $\beta_L$ (a), et les coudes  $\beta$  de type I, I', II et II'. Les formes  $\beta_L$ - $\beta_L$ (a) et  $\gamma_L$ - $\beta_L$ (a) dont le groupement NH<sub>2</sub> est libre peuvent être écartés (Figure 4.9). La forme  $\beta_L$ - $\gamma_L$ (g+) reproduit correctement le spectre expérimental dans la région amide A mais l'accord est moins satisfaisant dans le domaine amide I. L'écart entre les bandes extrêmes CO de Gly-Phe B de l'ordre de 30 cm<sup>-1</sup> n'est en effet pas retrouvé par les calculs de la forme  $\beta_L$ - $\gamma_L$ (g+); de plus son énergie très élevée nous permet d'écarter cette forme. En revanche, parmi les coudes  $\beta$  les plus stables (cf Tableau 4.3 en annexe), on constate que les coudes type I (g+) et type II' (g+) donnent un très bon accord dans les domaines amide A et I. Les coudes  $\beta$  de type I (g+) et de type II' (g+) peuvent tous deux rendre compte du spectre expérimental du conformère Gly-Phe B (Figure 4.9). À ce stade, les méthodes expérimentales et théoriques dont nous disposons ne nous permettent pas de conclure entre les deux formes.

Au vu des résultats théoriques présentés qui montrent que les coudes  $\beta$  sont des structures très stables, et peuvent être attribués aux espèces B de Phe-Gly, Phe-Val, Phe-Pro et Gly-Phe,



nous attribuons également les conformères B de Ala-Phe et Val-Phe à des coudes  $\beta$  avec une orientation ( $g+$ ) du cycle aromatique afin de reproduire l'interaction NH- $\pi$  observée.

### 4.3 Effet de la séquence : phénylalanine vs. tyrosine et NH<sub>2</sub> vs. NHMe

Nous avons pu constater à quel point il était parfois délicat de proposer une attribution sur la base des spectres IR uniquement. Ainsi, lorsque la liaison hydrogène est de force modérée comme dans le cas des liaisons C<sub>10</sub> ou des liaisons C<sub>7</sub> des formes  $\beta_L$ - $\gamma_L$ , le recours au modèle de couplage vibrationnel entre les composantes symétrique et antisymétrique de NH<sub>2</sub> nous a permis d'évaluer la force de la liaison hydrogène impliquant le groupe NH<sub>2</sub> et donc de donner une attribution claire et non ambiguë. En revanche, lorsque l'on modifie la nature de la protection C-terminale en la remplaçant par exemple par un groupement méthylamide NH-CH<sub>3</sub> (ou NHMe) - ce groupement est en fait plus voisin de la liaison peptidique - on perd l'information structurale qui était apportée par le couplage vibrationnel et il devient parfois plus difficile d'identifier certaines structures.

Ainsi, Gerhards et collaborateurs ont étudié un système très proche des nôtres, le peptide protégé Ac-Val-Tyr(Me)-NHMe - le groupement hydroxyle du résidu tyrosine est méthylé afin de ne pas interagir avec la chaîne principale - et ont attribué ce système à une structure complètement étendue  $\beta_L$ - $\beta_L$  [9]. Cette conclusion étant à l'opposé de notre interprétation, ce paragraphe sera consacré à la discussion des résultats du groupe de Gerhards.

Leurs travaux sur les systèmes protégés Ac-Phe-OMe [10, 11, 12] et Ac-Trp-OMe [13] ont permis à Gerhards et collaborateurs de montrer que ceux-ci adoptaient des géométries étendues  $\beta_L$  équivalentes au brin  $\beta$ . Dans cette perspective, l'équipe a étendu ses investigations à des peptides protégés plus longs : les systèmes Ac-Val-Phe-OMe [14] et Ac-Val-Tyr(Me)-NHMe [9]. À partir des signatures IR, ils ont attribué le conformère principal de chaque peptide protégé à la structure étendue brin  $\beta$  ( $\beta_L$ - $\beta_L$ ) (Figure 4.10). Or, le dipeptide Ac-Val-Tyr(Me)-NHMe est très semblable au nôtre : la protection NH<sub>2</sub> est remplacée par un groupe NHMe, ce qui affecte peu la conformation globale de la chaîne, et le résidu Phe est remplacé par le résidu Tyr. Il est

par conséquent très surprenant d'aboutir à deux interprétations différentes, à savoir un coude  $\beta$  pour notre système Ac-Val-Phe-NH<sub>2</sub> et une forme  $\beta_L$ - $\beta_L$  pour Ac-Val-Tyr(Me)-NHMe.

Afin de comprendre cette différence, nous avons cherché à déterminer l'influence de la protection NHMe et du résidu Tyr(Me). Nous avons donc effectué des calculs de fréquences (au niveau B3LYP/6-31+G(d)) et d'énergie (au niveau MP2 à partir des géométries optimisées en DFT) sur les systèmes Ac-Val-Phe-NH<sub>2</sub>, Ac-Val-Phe-NHMe et sur le système Ac-Val-Tyr(Me)-NHMe. Les structures sélectionnées sont les suivantes : la structure  $\beta_L(g^-)$ - $\beta_L(a)$  identique à celle calculée par le groupe de Gerhards et des coudes  $\beta$  de type I avec une orientation (g+) du cycle aromatique et avec les trois orientations de la chaîne latérale de Val. L'orientation (g+) du cycle aromatique permet de reproduire la bande expérimentale située à 3440 cm<sup>-1</sup> dans le spectre IR de la molécule Ac-Val-Tyr(Me)-NHMe. Les résultats sont présentés dans le Tableau 4.2 et la Figure 4.10.

- La première remarque que l'on puisse faire concerne l'ordre énergétique, celui-ci est conservé pour les trois dipeptides : le coude  $\beta$  type I (g-/g+) est la structure la plus stable, suivie des coudes type I (g+/g+) et type I (a/g+); la forme  $\beta_L$ - $\beta_L$  est de 3 kcal/mol plus haute en énergie.
- Les géométries pour Ac-Val-Phe-NHMe et Ac-Val-Phe-NH<sub>2</sub> sont superposables pour les coudes  $\beta$  mais aussi pour les formes  $\beta_L$ - $\beta_L$  (Figure 4.10), prouvant que la nature de la protection, NH<sub>2</sub> ou NHMe, affecte peu la conformation globale de la chaîne peptidique.
- Les fréquences des modes amide A des coudes  $\beta$  et de la forme  $\beta_L$ - $\beta_L$  sont peu différentes pour les systèmes Ac-Val-Phe-NHMe et Ac-Val-Tyr(Me)-NHMe, la liaison hydrogène C<sub>10</sub> et l'interaction C<sub>5</sub> montrent des forces similaires. La bande C<sub>10</sub> est en revanche beaucoup plus intense que la bande C<sub>5</sub>. Les mêmes classes de structures entre Ac-Val-Phe-NHMe et Ac-Val-Tyr(Me)-NHMe sont comparables, les angles dièdres ayant des valeurs très proches.

Finalement, dans le cas des peptides avec la protection -NHMe, on constate qu'il est difficile de distinguer les coudes  $\beta$  de la forme étendue  $\beta_L$ - $\beta_L$  à partir des seuls spectres IR car les vibreurs de ces deux structures absorbent dans la même région. Néanmoins, on remarque que les bandes amide A sont plus resserrées pour la structure  $\beta_L$ - $\beta_L$  que pour les coudes  $\beta$  qui

		coudes $\beta$ type I				
		(a/g+)	(g-/g+)	(g+/g+)	$\beta_L(g-)-\beta_L(a)$	Exp
Ac-Val- Tyr(Me)-NHMe <sup>a</sup>						
$\Delta E$ (kcal/mol)		1.1	0.0	0.1	3.0	
Amide A	NHMe	3419	3413	3414	3466	<b>3477</b>
	NHTyr	3450	3457	3456	3414	<b>3411</b>
	NHVal	3471	3475	3462	3432	<b>3440</b>
Amide I	COAce	1691	1694	1691	1713	<b>1715</b>
	COVal	1702	1703	1702	1681	<b>1694</b>
	COTyr	1718	1711	1710	1706	<b>1700</b>
Ac-Val- Phe-NHMe						
$\Delta E$ (kcal/mol)		1.5	0.0	0.5	3.7	
Amide A	NHMe	3418	3416	3415	3467	
	NHPhe	3449	3459	3454	3414	
	NHVal	3470	3474	3463	3433	
Amide I	COAce	1691	1691	1691	1713	
	COVal	1701	1700	1702	1681	
	COPhe	1719	1711	1710	1707	
Ac-Val- Phe-NH <sub>2</sub> <sup>b</sup>						
$\Delta E$ (kcal/mol)		1.5	0.0	0.5	3.3	
Amide A	NH <sub>2</sub> sym	3384	3386	3383	3427	<b>3443</b>
	NHPhe	3449	3458	3457	3416	<b>3385</b>
	NHVal	3468	3474	3462	3435	-
	NH <sub>2</sub> anti	3521	3523	3522	3545	<b>3521</b>
Amide I	COAce	1699	1697	1698	1727	<b>1721</b>
	COVal	1714	1711	1710	1682	<b>1694</b>
	COPhe	1722	1719	1719	1712	-

TAB. 4.2 – <sup>a</sup> Les valeurs expérimentales (en gras) de Ac-Val-Tyr(Me)NHMe sont tirées de la réf [9], <sup>b</sup> celles de Ac-Val-Phe-NH<sub>2</sub> sont issues de nos travaux. Les fréquences sont respectivement corrigées par un facteur 0.96 et 0.98 pour les modes NH et CO

montrent une bande  $C_{10}$  relativement isolée. À l'inverse, le motif correspondant aux modes CO semble plus compact : l'écart entre les bandes extrêmes est de l'ordre de  $20 \text{ cm}^{-1}$ , alors qu'il est plus grand pour  $\beta_L-\beta_L$  ( $\sim 30 \text{ cm}^{-1}$ ). En outre, les fréquences d'élongation amide A et amide I d'un groupement amide impliqué dans une liaison hydrogène sont significativement plus basses : plus la liaison hydrogène est forte, plus la bande est intense, élargie et décalée vers le rouge. Or, les bandes associées aux liaisons  $C_{10}$  sont beaucoup plus intenses que celles des interactions  $C_5$  (et des bandes  $\text{NH}-\pi$ ) qui correspondent à une perturbation plus faible. On est donc tenté d'associer la bande NH intense à  $\sim 3440 \text{ cm}^{-1}$  du spectre de Ac-Val-Tyr(Me)-NHMe à une liaison hydrogène  $C_{10}$ .

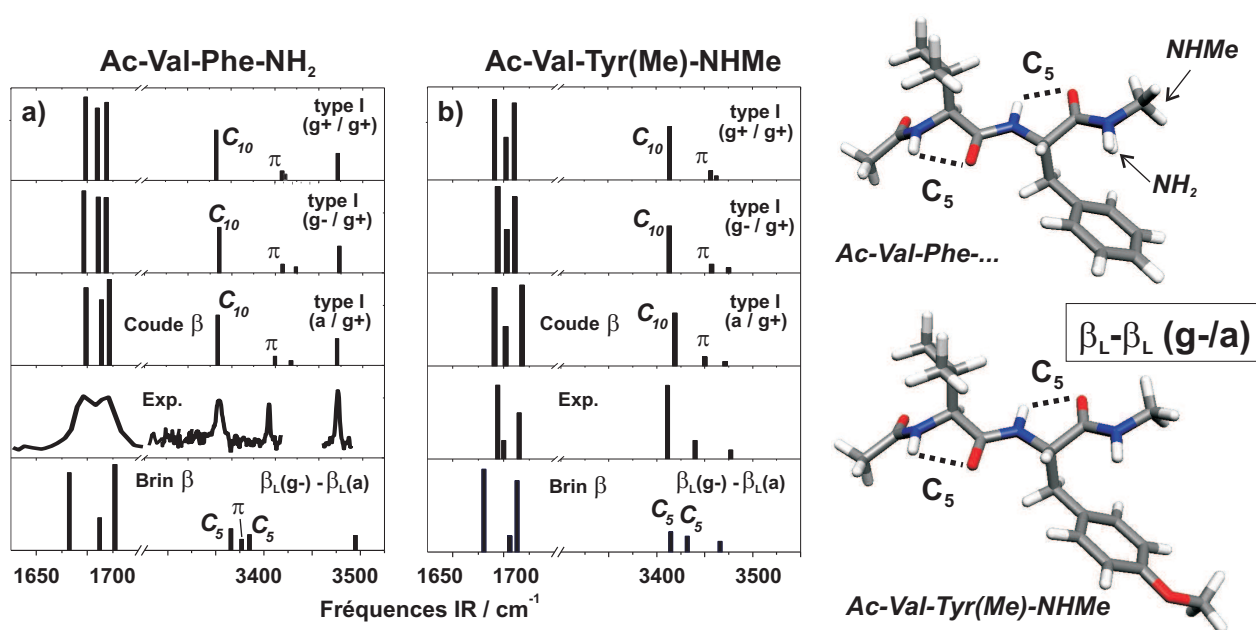


FIG. 4.10 – Spectres IR des modes d'élongation NH et CO mesurés et calculés pour les systèmes Ac-Val-Phe-NH<sub>2</sub> et Ac-Val-Tyr(Me)-NHMe. Les valeurs expérimentales de Ac-Val-Tyr(Me)NHMe sont tirées de la Ref. [9].

L'attribution dans le cas de Ac-Val-Phe-NH<sub>2</sub> est plus simple. En plus des informations déduites des spectres IR, c'est-à-dire la position des bandes et leur intensité, le couplage vibrationnel entre les deux composantes du groupe NH<sub>2</sub> permet d'évaluer la force de la perturbation induite sur le groupe amino. Ainsi, dans le cas de la conformation étendue  $\beta_L-\beta_L$ , le groupe amino est libre et mène à une fréquence d'élongation  $\nu_{anti}$  est élevée ( $3545 \text{ cm}^{-1}$ ). À l'inverse, lorsque le groupe NH<sub>2</sub> est impliqué dans une liaison  $C_{10}$ , cette fréquence est si-

gnificativement plus basse ( $3521\text{ cm}^{-1}$ ). La protection  $\text{NH}_2$  apporte donc une information structurale supplémentaire par rapport aux peptides méthylamidés puisqu'elle permet, dans le cas présent, de distinguer aisément les coudes  $\beta$  des formes étendues  $\beta_L\text{-}\beta_L$ .

Par conséquent, l'attribution du conformère B de Ac-Val-Phe- $\text{NH}_2$  et les calculs menés sur les deux autres dipeptides nous conduisent à attribuer le conformère Ac-Val-Tyr(Me)-NHMe à un coude  $\beta$ . On voit que la bande intense la plus rouge située à  $3411\text{ cm}^{-1}$  est en effet très bien reproduite par les calculs de fréquences effectués sur les coudes  $\beta$  (Figure 4.10). Les calculs d'énergie montrent de plus que ces structures sont beaucoup plus stables que la structure  $\beta_L\text{-}\beta_L$  (+ 2.99 kcal/mol au niveau MP2). Ceci nous permet d'affirmer que ni la nature du résidu (Tyr vs. Phe), ni la nature de la protection (-NHMe vs.  $\text{-NH}_2$ ) n'affectent la conformation globale de la chaîne mais que la formation du coude  $\beta$  résulte de la propension de la chaîne à adopter spontanément la structure repliée. L'interprétation erronée de Gerhards et collaborateurs peut néanmoins s'expliquer par les arguments suivants :

- Les résultats obtenus sur Ac-Val-Tyr(Me)-NHMe ont été comparés au système Ac-Val-Phe-OMe qui adopte la forme  $\beta_L\text{-}\beta_L$  [14]. Or, la protection -OMe très rigide empêche de fait la formation d'un coude  $\beta$ , rendant le système Ac-Val-Phe-OMe pertinent pour modéliser un brin  $\beta$ .
- Les calculs de fréquences et d'énergie relative n'ont été effectués que sur des formes étendues  $\beta_L\text{-}\beta_L$ . Celles-ci résultaient d'une recherche conformationnelle concluant que le brin  $\beta$  était la structure la plus stable. Les coudes  $\beta$  qui sont pourtant parmi les formes les plus stables dans nos systèmes, ainsi que pour le système Ac-Val-Phe-NHMe, n'ont donc pas été envisagés.

Finalement, cette discussion permet de valider l'utilisation de la protection  $\text{-NH}_2$  pour nos systèmes à la place du groupement -NHMe, plus respectueux de la fonction amide CO-NH de la liaison peptidique, en soulignant l'information structurale supplémentaire apportée par couplage vibrationnel entre les composantes du groupe  $\text{NH}_2$ . De plus, elle illustre parfaitement la difficulté d'identifier la structure de petits polypeptides de façon non équivoque en raison de la complexité du paysage conformationnel qui présente des structures d'énergies voisines, même pour des systèmes de taille modeste. Le paysage énergétique fait l'objet du paragraphe

qui suit.

#### 4.4 Stabilité des coudes $\beta$

Nous avons effectué une recherche conformationnelle détaillée pour les dipeptides Ac-Phe-Gly-NH<sub>2</sub> et Ac-Gly-Phe-NH<sub>2</sub> sans toutefois chercher à explorer la surface d'énergie potentielle de façon exhaustive. Nous avons choisi de mener cette étude théorique avec le résidu Gly car c'est le système le plus petit, ce qui permet de réduire les temps de calcul. Mais de par l'absence de chaîne latérale, c'est aussi le système le plus flexible. Les énergies relatives ont été calculées sur des géométries B3LYP avec la base 6-31+G(d), par deux méthodes : MP2 afin de tenir compte des effets de dispersion, et B3LYP. Les calculs MP2 convergeant lentement avec la taille de la base, nous avons également considéré la base 6-311+G(d,p) pour certaines structures du dipeptide Ac-Gly-Phe-NH<sub>2</sub>. Les résultats sont regroupés dans les Tableaux 4.3 et 4.4 en annexe, seul un schéma énergétique des conformations trouvées dans un domaine de 2 kcal/mol au niveau MP2/6-31+G(d) est présenté dans cette section (Figure 4.11).

Les calculs montrent l'existence de deux familles conformationnelles, chacune appartenant à un bassin conformationnel distinct. La première est représentée par des structures qui correspondent à la succession de préférences conformationnelles locales autour de chaque résidu ( $\beta_L$ - $\gamma_L$ ,  $\gamma_L$ - $\beta_L$ ,  $\gamma_L$ - $\gamma_L$ ) - cette famille a été identifiée dans le chapitre précédent - la seconde est constituée des structures secondaires coude  $\beta$  et brin  $\beta$ . Au sein d'une même famille, l'ordre énergétique est différent suivant l'ordre des résidus Phe et Gly ; la discussion détaillée des résultats sera donc traitée dans des parties distinctes.

##### Énergies des structures calculées pour Ac-Gly-Phe-NH<sub>2</sub>

Les résultats sont différents selon la méthode de calcul utilisée : la forme la plus stable trouvée avec la fonctionnelle B3LYP est la forme étendue  $\beta_L$ - $\beta_L$  équivalente au brin  $\beta$ . Dans un domaine d'énergie de 1 kcal/mol sont présentes 10 conformations. Parmi celles-ci, on peut citer les coudes  $\beta$  type II' (g+) et type II' (g-), les coudes doubles  $\gamma_D$ - $\gamma_L$ (g-),  $\gamma_L$ - $\gamma_L$ (g-) et  $\gamma_L$ - $\gamma_L$  et les formes  $\beta_L$ - $\gamma_L$  avec les trois orientations du cycle phényle (Tableau 4.3). Outre la

structure secondaire brin  $\beta$ , on constate donc que la fonctionnelle B3LYP favorise les formes locales sur chaque résidu ( $\gamma$  sur Gly et  $\beta_L$  et  $\gamma_L$  pour Phe). On note toutefois que le coude  $\beta$  type II' (g+) est particulièrement stable avec une énergie de 0.57 kcal/mol au dessus du brin  $\beta$ , ce qui est en accord avec les potentiels de résidu. En revanche, le type I est moins stable (+ 1.70 kcal/mol pour le type I (g-)) alors que c'est le type le plus couramment observé dans les protéines.

Les calculs au niveau MP2 aboutissent à des résultats différents : l'ordre énergétique est modifié et la différence d'énergie entre les conformations considérées augmente. Les coudes  $\beta$  deviennent les structures les plus stables suivis de deux coudes  $\gamma$ - $\gamma$  (Figure 4.11). Le brin  $\beta$  est très haut en énergie, il est repoussé à plus de 3 kcal/mol au-dessus du coude  $\beta$  type I (g+) qui est la forme la plus stable. Les calculs d'énergie MP2 avec la base 6-311+G(d,p) sont globalement similaires à ceux effectués avec la base 6-31+G(d) ; l'ordre énergétique est notamment conservé, bien que les énergies relatives soient légèrement plus basses. L'effet de la taille de la base sur nos calculs n'est donc pas significatif, ce qui justifie l'emploi d'une base moyenne pour tous les calculs présentés dans ce travail.

Une observation attentive des énergies nous amène à quelques constatations importantes :

- Les structures les plus stables (dans un domaine d'énergie de 2 kcal/mol) sont des coudes  $\beta$  et des coudes  $\gamma$ - $\gamma$ . Ceci conforte les attributions des conformères B et A de Ac-Gly-Phe-NH<sub>2</sub>, respectivement, aux formes coude  $\beta$  (type II' ou I) et  $\gamma$ - $\gamma$  déduites de la comparaison entre les données spectroscopiques et les fréquences théoriques. La stabilité des formes  $\gamma$ - $\gamma$  n'est pas très surprenante car elles correspondent à la succession de la forme locale sur Gly et de la deuxième forme la plus stable pour Phe (voir Section 3.3 du Chapitre 3 sur les conformères A de Ac-Xxx-Phe-NH<sub>2</sub>).
- Les coudes  $\beta$  type I (g+) et type II' (g+) sont des formes particulièrement stables. Les coudes de type I sont les plus fréquents dans les protéines et, comme nous l'avons mentionné, le type II' est spécifique du résidu Gly, sa nature achirale et, de fait, sa petite taille n'entraînant aucun encombrement stérique dans la protéine.
- Seuls les coudes  $\beta$  types I et II' avec une orientation (g+) du cycle aromatique sont parmi les minima de plus basse énergie au niveau MP2, alors que les autres coudes  $\beta$

Ac-Gly-Phe-NH <sub>2</sub>	$\Delta E$ B3LYP + $\Delta ZPE$ 6-31+G(d)	$\Delta E$ MP2 + $\Delta ZPE$ 6-31+G(d)	$\Delta E$ MP2 + $\Delta ZPE$ 6-311+G(d,p)	Moyenne sur les 3 orientations du cycle MP2/6-31+G(d)
<b>0 liaison H</b>				
$\beta_L$ - $\beta_L$ (a)	0.00	3.82	4.90	
<b>1 liaison H</b>				
$\gamma_D$ - $\beta_L$ (a)	2.28	3.48	3.71	
$\gamma_L$ - $\beta_L$ (a)	2.45	3.62	en cours	
$\beta_L$ - $\gamma_L$ (g+)	1.35	4.65	4.08	
$\beta_L$ - $\gamma_L$ (g-)	1.16	5.03	4.24	
$\beta_L$ - $\gamma_L$ (a)	1.46	5.60	4.90	
<b>2 liaisons H</b>				
$\gamma_L$ - $\gamma_L$ (g-)	0.68	0.65	0.29	
$\gamma_L$ - $\gamma_L$ (g+)	1.57	2.21	1.82	2.00
$\gamma_L$ - $\gamma_L$ (a)	1.17	3.14	2.81	
$\gamma_D$ - $\gamma_D$ (g-)	2.02	1.18	0.93	
$\gamma_D$ - $\gamma_L$ (g-)	0.59	1.77	1.43	
$\gamma_D$ - $\gamma_L$ (a)	0.86	2.50		2.37
$\gamma_D$ - $\gamma_L$ (g+)	2.82	2.85		
$\gamma_L$ - $\gamma_D$ (g-)	2.25	3.11		
<b>Coudes <math>\beta</math></b>				
type I (g+)	4.86	0.00	0.00	
type I (g-)	2.27	2.73	2.61	2.91
type I (a)	4.77	6.01	5.87	
type I' (g-)	4.28	3.80	3.44	
type I' (g+)	converge vers la forme $\gamma_D$ - $\delta_L$			
type I' (a)	6.77	7.67		
type II (g+)	5.51	3.52	2.71	
type II (g-)	3.71	3.63	3.16	4.79
type II (a)	6.64	7.22	6.38	
type II' (g+)	0.57	0.49	0.28	
type II' (g-)	1.41	2.36	2.09	2.74
type II' (a)	3.80	5.39	5.02	

TAB. 4.3 – Énergies relatives calculées au niveau B3LYP et MP2 pour différentes structures du système Ac-Gly-Phe-NH<sub>2</sub>.



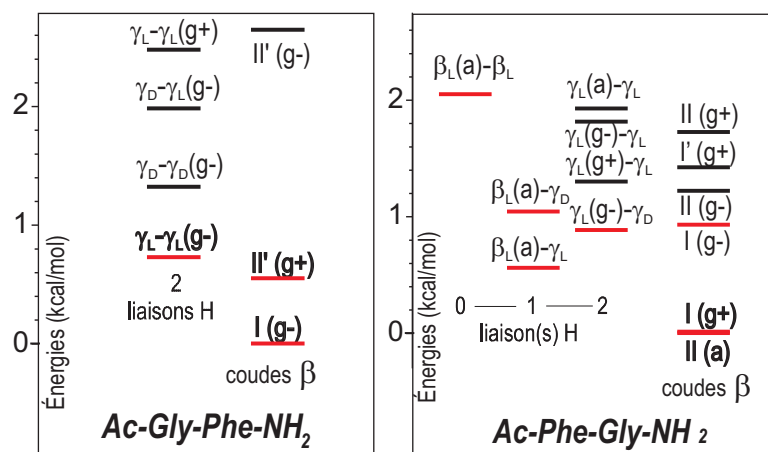


FIG. 4.11 – Schéma énergétique des conformations les plus stables de Ac-Gly-Phe-NH<sub>2</sub> et de Ac-Phe-Gly-NH<sub>2</sub> calculées au niveau MP2/6-31+G(d) à partir de géométries optimisées au niveau B3LYP/6-31+G(d).

sont défavorisés énergétiquement par rapport aux formes  $\gamma$ - $\gamma$ . La stabilité de ces deux formes est vraisemblablement due à la présence d'une interaction NH- $\pi$  permise par l'orientation (g+) de Phe et qui est absente pour les autres coudes  $\beta$  (distance NH...C $\gamma$  supérieure à 3 Å).

### Énergies des structures calculées pour Ac-Phe-Gly-NH<sub>2</sub>

Comme pour Ac-Gly-Phe-NH<sub>2</sub>, les calculs menés en DFT et au niveau MP2 ne conduisent pas au même ordre énergétique. Une dizaine de structures sont trouvées dans un domaine de 1 kcal/mol au dessus du coude  $\beta$  type II (a), qui est le minimum trouvé par les calculs B3LYP ; les formes coudes  $\gamma$  doubles,  $\beta_L$ - $\gamma$  et le brin  $\beta$  sont particulièrement stables.

Les calculs MP2 semblent en revanche défavoriser ces dernières structures, en particulier les coudes  $\gamma$ - $\gamma$  et le brin  $\beta$ . Les coudes  $\beta$  type I et type II, ainsi que les formes  $\beta_L$ - $\gamma$  sont les structures les plus stables (Figure 4.11). On peut noter que la moyenne des énergies sur les trois orientations du cycle aromatique souligne la très grande stabilité des coudes  $\beta$  type II, suivis ensuite des types I. Ces calculs, en phase gazeuse, reproduisent donc très bien ce qui est observé dans les protéines cristallisées, à savoir la propension d'un résidu à adopter un type de coude  $\beta$  donné. On peut aussi insister sur la stabilité des formes  $\beta_L$ - $\gamma$  qui conforte l'attribution proposée pour les conformères A de la série Ac-Phe-Xxx-NH<sub>2</sub>. Néanmoins, la for-

Ac-Phe-Gly-NH <sub>2</sub>	$\Delta E$ B3LYP + $\Delta ZPE$	$\Delta E$ MP2 + $\Delta ZPE$	$\Delta E$ MP2 + $\Delta ZPE$ sur les 3 orientations de Phe
<b>0 liaison H</b>			
$\beta_L(a)-\beta_L$	0.29	1.83	
<b>1 liaison H</b>			
$\beta_L(a)-\gamma_L$	0.45	0.50	
$\beta_L(a)-\gamma_D$	0.26	0.93	
$\gamma_L(g+)-\beta_L$	1.05	2.56	
$\gamma_L(g-)-\beta_L$	2.36	4.78	
$\gamma_L(a)-\beta_L$	converge vers $\beta_L(a)-\beta_L$		
<b>2 liaisons H</b>			
$\gamma_L(g-)-\gamma_D$	0.20	0.79	
$\gamma_L(g+)-\gamma_L$	0.58	1.16	
$\gamma_L(g-)-\gamma_L$	0.23	1.62	
$\gamma_D(g-)-\gamma_D$	1.35	1.65	
$\gamma_L(g+)-\gamma_D$	0.80	1.75	
<b>Coudes <math>\beta</math></b>			
type I (g+)	0.97	0.01	
type I (g-)	0.74	0.83	1.60
type I (a)	4.24	3.97	
type I' (g+)	6.10	1.27	
type I' (g-)	2.14	3.40	3.30
type I' (a)	4.83	5.22	
type II (a)	0.00	0.00	
type II (g-)	0.20	1.09	0.88
type II (g+)	2.10	1.54	
type II' (g-)	3.09	2.95	
type II' (a)	4.10	3.17	3.99
type II' (g+)	7.84	5.84	

TAB. 4.4 – Énergies relatives calculées au niveau B3LYP et MP2 (single point) avec la base 6-31+G(d) pour différentes structures du système Ac-Phe-Gly-NH<sub>2</sub>.

mation de cette structure ne résulte pas des propriétés intrinsèques de repliement du squelette peptidique comme c'est le cas pour les coudes  $\beta$ , mais proviendrait des effets de chaîne latérale, comme nous l'avons vu dans le chapitre précédent : la présence d'une interaction NH- $\pi$  entre le cycle aromatique et un groupement amide de la liaison peptidique apporte en effet une interaction stabilisante supplémentaire qui favorise la structure étendue  $\beta_L$ . En l'absence d'une telle interaction, cette structure n'est pas observée ; c'est par exemple le cas des petites chaînes de Gly ou de Ala. Enfin, on voit que les coudes doubles  $\gamma$  sont aussi très stables. Néanmoins, ils appartiennent au même bassin conformationnel que les structures  $\beta_L$ - $\gamma$ , et, parce que la relaxation conformationnelle est très efficace dans le jet, on ne peuple finalement qu'une seule espèce dans chaque bassin, c'est-à-dire la structure  $\beta_L$ - $\gamma$  qui se révèle être plus stable car elle correspond à la succession des formes locales pour chaque résidu.

Finalement, on constate que certaines structures issues de chacune des deux classes de conformations, coudes  $\beta$  et préférences conformationnelles locales, sont de stabilité très similaire et sont donc en compétition. L'observation de ces deux familles pour les dipeptides est en fait permise par le fait qu'elles appartiennent à deux bassins conformationnels différents possédant des barrières suffisamment hautes pour qu'elles ne s'interconvertissent pas.

## 4.5 Résumé et conclusions sur les coudes $\beta$

Ces travaux constituent la première observation spectroscopique en phase gazeuse des coudes  $\beta$ , l'une des structures secondaires les plus fréquentes dans les protéines. Cette étude sur les séries de dipeptides Ac-Phe-Xxx-NH<sub>2</sub> et Ac-Xxx-Phe-NH<sub>2</sub> nous a permis de montrer que les coudes  $\beta$  étaient des formes très stables, robustes en phase gazeuse, en accord avec les calculs menés en parallèle ainsi que certains travaux théoriques portant sur de petits systèmes protégés. En effet, nous avons montré que la formation des coudes  $\beta$  était spontanée dans nos conditions et dépendait peu de la nature des résidus considérés.

Nous avons également montré que le paysage conformationnel de petits peptides à deux résidus (Gly et Phe) était complexe et pouvait mener à des structures possédant des éner-

gies très proches. Ainsi, nous avons mis en évidence un phénomène de compétition entre la formation des coudes  $\beta$  et celle de structures plus locales comme les formes  $\beta_L$ - $\gamma$  et  $\gamma$ - $\gamma$ , rendant parfois difficile l'attribution de certaines espèces. Toutefois, notre double approche, expérimentale et théorique, nous a permis de distinguer la signature spectrale des coudes  $\beta$  de celle d'autres conformations, en particulier grâce à l'information structurale apportée par les modes d'élongation symétrique et antisymétrique de la protection  $\text{NH}_2$  terminale. Les coudes  $\beta$  correspondent à des structures repliées pour lesquelles les extrémités N- et C-terminales sont reliées par une liaison hydrogène  $\text{C}_{10}$ , tandis que les formes locales  $\beta_L$ - $\gamma$  et  $\gamma$ - $\gamma$  mènent à des structures « ouvertes » dont les extrémités ne sont pas reliées. Nous avons alors vu qu'il était parfois difficile de distinguer ces deux familles de structures qui ont des énergies très voisines. La signature des coudes  $\beta$  se traduit par les résultats suivants :

- une liaison hydrogène  $\text{C}_{10}$  de force modérée se traduisant par le couple de fréquences  $\nu_{\text{sym}}$ - $\nu_{\text{anti}}$  3390-3520  $\text{cm}^{-1}$ ,
- un motif de bandes CO (amide I) relativement compact correspondant également à une liaison hydrogène modérée,
- un couplage vibrationnel ( $\sim 60 \text{ cm}^{-1}$ ) entre les composantes  $\text{NH}_2$  symétrique et antisymétrique significativement plus faible que pour les coudes  $\gamma$ .

Ces résultats très précis n'auraient pu être obtenus par spectroscopie d'absorption IR en solution en raison de la faible résolution ainsi que des recouvrements spectraux de plusieurs espèces qui ne permettent pas toujours de détecter des coudes  $\beta$ . En outre, bien que très stables, ceux-ci sont des formes minoritaires et leur signature est souvent masquée par celle des formes majoritaires.

L'observation de différents types de coudes  $\beta$  en fonction de la séquence, type II pour Ac-Phe-Gly- $\text{NH}_2$ , type II' ou type I pour Ac-Gly-Phe- $\text{NH}_2$ , type VI pour Ac-Phe-Pro- $\text{NH}_2$  et probablement type I pour Ac-Phe-Val- $\text{NH}_2$  et Ac-Xxx-Phe- $\text{NH}_2$  (Xxx = Ala, Val), est en accord avec les « potentiels de résidu » tirés des spectres de diffraction X de protéines cristallisées qui sont utilisés en biochimie [15, 16, 17]. Les coudes  $\beta$  type II' et type II, qui représentent 4 et 20% des coudes  $\beta$  identifiés dans les protéines, sont presque exclusivement formés pour les résidus Gly placés respectivement en position (i+1) et (i+2) dans la séquence. Ces types

sont rarement observés pour les autres résidus en raison de l'encombrement stérique dû à la présence de chaînes latérales. Ainsi, tandis que ces potentiels renseignent sur la fréquence des résidus dans les coudes  $\beta$  en phase cristalline, les résultats que nous obtenons en phase gazeuse illustrent bien le caractère intrinsèque de la formation des coudes  $\beta$ .

Enfin, malgré leur stabilité, les coudes  $\beta$  apparaissent comme des formes minoritaires dans nos travaux ; ce sont les conformères  $\beta_L$ - $\gamma$  et  $\gamma$ - $\gamma$  qui sont les espèces les plus peuplées. Néanmoins, on observe un changement dans l'abondance des différents conformères lorsque les effets de chaîne latérale interviennent. Ainsi, le coude  $\beta$  du dipeptide Ac-Val-Phe-NH<sub>2</sub> devient-il le conformère majoritaire (sa bande UV est la plus intense dans le spectre électronique, Figure 4.1, p.120). Afin de caractériser ces effets de chaîne latérale, il serait intéressant de voir l'effet d'un résidu possédant une chaîne latérale très volumineuse telle que la leucine ou l'isoleucine. Les interactions entre chaînes latérales participent par exemple à l'effet hydrophobe qui permet de stabiliser certaines structures secondaires comme les hélices  $\alpha$  en minimisant l'accès des molécules d'eau à la chaîne principale.

**Références bibliographiques du Chapitre 4**

- [1] C. M. Venkatachalam, *Biopolymers* **6**, 1425 (1968).
- [2] G. Rose, L. M. Gierasch and J. A. Smith, *Adv. Prot. Chem.* **37**, 1 (1985).
- [3] H. J. Böhm, *J. Am. Chem. Soc.* **115**, 6152 (1993).
- [4] C.-H. Yu, L. Schafer and M. Ramek, *J. Phys. Chem. A* **103**, 8337 (1999).
- [5] M. Elstner, K. J. Jalkanen, M. Knapp-Mohammady, T. Frauenheim and S. Suhai, *Chem. Phys.* **263**, 203 (2001).
- [6] K. Möhle, M. Gussman, A. Rost, R. Cimiraglia and H.-J. Hofmann, *J. Phys. Chem. A* **101**, 8571 (1997).
- [7] P. W. Baures, W. H. Ojala, W. B. Gleason and R. L. Johnson, *J. Pept. Res.* **50** (1997).
- [8] T. Shi, S. M. Spain. and D. L. Rabenstein, *J. Am. Chem. Soc.* **126**, 790 (2004).
- [9] H. Fricke, A. Gerlach, C. Unterberg, P. Rzepecki, T. Schrader and M. Gerhards, *Phys. Chem. Chem. Phys.* **6**, 4636 (2004).
- [10] C. Unterberg, A. Gerlach, T. Schrader and M. Gerhards, *Eur. Phys. J. D* **20**, 543 (2002).
- [11] M. Gerhards and C. Unterberg, *Phys. Chem. Chem. Phys.* **4**, 1760 (2002).
- [12] M. Gerhards, C. Unterberg and A. Gerlach, *Phys. Chem. Chem. Phys.* **4**, 5563 (2002).
- [13] A. Gerlach, C. Unterberg, H. Fricke and M. Gerhards, *Molec. Phys.* **103**, 1521 (2005).
- [14] C. Unterberg, A. Gerlach, T. Schrader and M. Gerhards, *J. Chem. Phys.* **118**, 8296 (2003), et réf. citées.
- [15] P. Chou and G. Fasman, *J. Mol. Biol.* **115**, 135 (1977).
- [16] E. G. Hutchinson and J. M. Thornton, *Protein Sci.* **3**, 2209 (1994), et ref. citées.
- [17] K. Guruprasad and S. Rajkumar, *J. Biosc.* **25**, 143 (2000).



## Chapitre 5

# À la recherche des hélices

La complexité du paysage conformationnel d'un peptide augmente fortement avec sa taille. Ainsi, en étudiant des tripeptides, nous avons cherché à observer de nouvelles structures. Notamment, les tripeptides sont les plus petites structures capables de former des hélices  $3_{10}$ , c'est à dire des structures hélicoïdales périodiques présentant une liaison hydrogène  $C_{10}$ . Toutefois, on peut se demander si les structures observées pour les dipeptides n'entreront pas en compétition avec les hélices  $3_{10}$  ou si des effets de longueur de la chaîne, voire des *effets coopératifs* permettront de stabiliser de nouvelles formes. En effet, d'après la Figure 5.1 qui schématise les liaisons hydrogène possibles pour un tripeptide  $\text{Ac-R1-R2-R3-NH}_2$ , plusieurs structures peuvent être formées : i) la liaison  $C_{13}$ , retrouvée dans les structures hélice  $\alpha$ , pourrait s'établir entre les groupements  $\text{CO}_{\text{Ace}}$  et  $\text{NH}_2$ , ii) une combinaison de deux coudes  $\beta$  ( $C_{10}$ - $C_{10}$ ) pourrait exister, iii) une structure coude  $\gamma$  triple ( $C_7$ - $C_7$ - $C_7$ ) serait possible, iv) la forme  $\beta_L$ - $\beta_L$ - $\beta_L$  ( $C_5$ - $C_5$ - $C_5$ ) équivalente au brin  $\beta$  pourrait en principe être formée, v) et enfin des structures mixtes de type  $C_5$ - $C_7$ - $C_5$ ,  $C_7$ - $C_7$ - $C_5$ ,  $C_7$ - $C_{10}$ , etc. Le point important se résume en fait à savoir à partir de quelle longueur de chaîne les hélices  $3_{10}$  sont stables et se forment spontanément. De très nombreux calculs ont en effet étudié la compétition entre ces différentes structures, cependant, selon la méthode employée, les conclusions divergent. Ainsi, Barone et collaborateurs trouvent que la conformation hélice  $\alpha$  est la plus stable pour des chaînes infinies d'homopeptides (c'est-à-dire constituées soit uniquement de résidus Gly, soit de résidus Ala, soit de résidus Aib, acide aminoisobutyrique qui est un dérivé du résidu Ala) [1]. D'autres



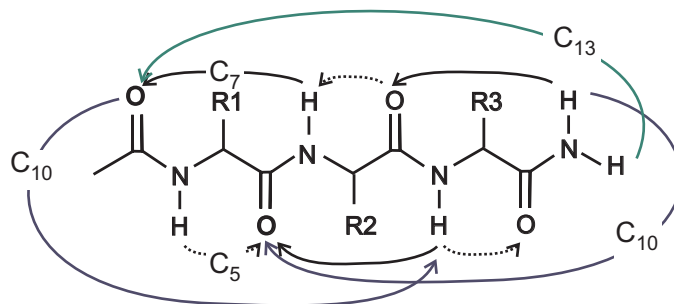


FIG. 5.1 – Schéma des liaisons hydrogène possibles pour le tripeptide Ac-R1-R2-R3-NH<sub>2</sub> : C<sub>5</sub> se réfère à la structure  $\beta_L$  (en pointillés), C<sub>7</sub> au coude  $\gamma$ , C<sub>10</sub> au coude  $\beta$  et C<sub>13</sub> à l'hélice  $\alpha$ .

travaux montrent en revanche que les hélices sont formées à partir d'une séquence composée de huit résidus [2, 3, 4, 5]. Les approches expérimentale et théorique employées dans l'étude des dipeptides ont également été appliquées à la caractérisation des tripeptides. Les résultats obtenus sur les systèmes précédents, Ac-Phe-NH<sub>2</sub> et dipeptides, seront largement exploités pour interpréter les spectres des tripeptides.

## 5.1 Ac-Phe-Gly-Gly-NH<sub>2</sub>

L'acide aminé glycine est très fréquemment employé pour caractériser le comportement de la chaîne peptidique. L'absence de chaîne latérale ainsi que sa petite taille permettent en effet de réduire les temps de calculs par rapport aux autres acides aminés. Nous avons donc commencé notre investigation sur les tripeptides avec une séquence contenant deux résidus Gly et un résidu Phe comme chromophore UV. Parce que les dipeptides Ac-Phe-Xxx-NH<sub>2</sub> présentent une spectroscopie UV plus simple que les dipeptides Ac-Xxx-Phe-NH<sub>2</sub>, nous avons choisi d'étudier le système Ac-Phe-Gly-Gly-NH<sub>2</sub> dont la séquence correspond au dipeptide Ac-Phe-Gly-NH<sub>2</sub> auquel a été ajouté un résidu Gly en bout de séquence, en nous appuyant sur la spectroscopie électronique de la série Ac-Phe-Xxx-NH<sub>2</sub> pour l'interprétation.

### Spectre électronique

Le spectre d'excitation de Ac-Phe-Gly-Gly-NH<sub>2</sub> est présenté sur la Figure 5.2. Il montre un spectre congestionné avec de nombreuses bandes ; il est plus riche que les spectres mesurés pour la série Ac-Phe-Xxx-NH<sub>2</sub>. Le spectre présente deux bandes larges entre 37550 et 37680 cm<sup>-1</sup>

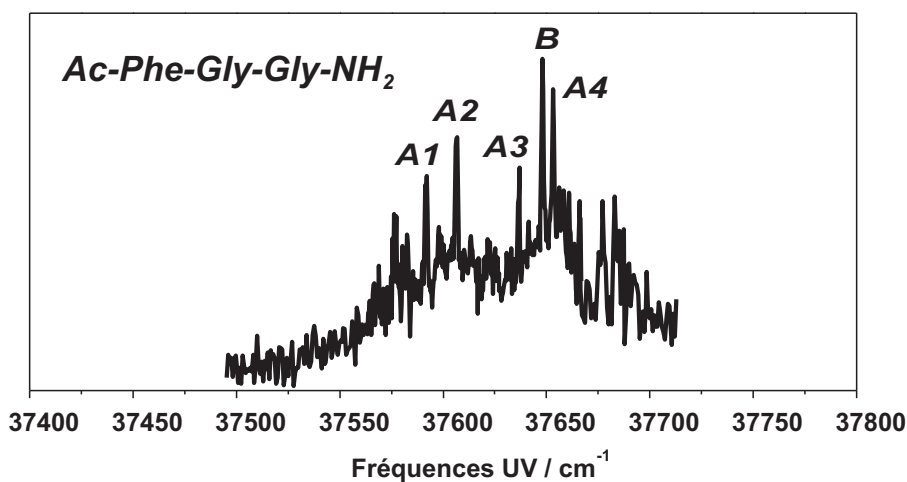


FIG. 5.2 – Spectre d'excitation du tripeptide  $\text{Ac-Phe-Gly-Gly-NH}_2$ . Les origines des cinq conformères sont identifiées par les lettres A1-A4 et B.

auxquelles se superposent quelques bandes intenses qui ont été identifiées comme appartenant à cinq conformères notés A1-A4 et B. L'espèce B présente la bande UV la plus intense. En raison du faible rapport signal/bruit, on ne peut exclure l'existence d'autres espèces. La présence de plusieurs conformères par rapport aux dipeptides pour lesquels seules deux formes majoritaires ont été mises en évidence n'est pas surprenante : plus le peptide est grand, plus les rotations autour des liaisons du squelette sont nombreuses conduisant à un paysage conformationnel très complexe. Un grand nombre de conformations est donc en principe accessible au résidu Gly puisque sa petite taille n'induit pas d'encombrement stérique. De plus, l'absence de chaîne latérale peut permettre à différentes formes d'avoir des énergies voisines, c'est par exemple le cas pour les formes énantiomères  $\gamma_L$  et  $\gamma_D$ .

Le spectre UV, bien que riche, ne semble pas montrer de progression vibrationnelle claire comme c'est le cas pour la série  $\text{Ac-Phe-Xxx-NH}_2$ . En outre, le décalage du spectre vers les hautes fréquences par rapport à celui de Phe-Gly, et des autres dipeptides, suggère qu'aucun des conformères n'adopte de structure étendue  $\beta_L$  sur le résidu Phe.

### 5.1.1 Diversité du paysage conformationnel

Les spectres IR dans la région amide A ont été enregistrés pour les cinq conformères, seul celui de B a été mesuré dans la région amide I-II (Figure 5.3 et Tableau 5.1). Les spectres

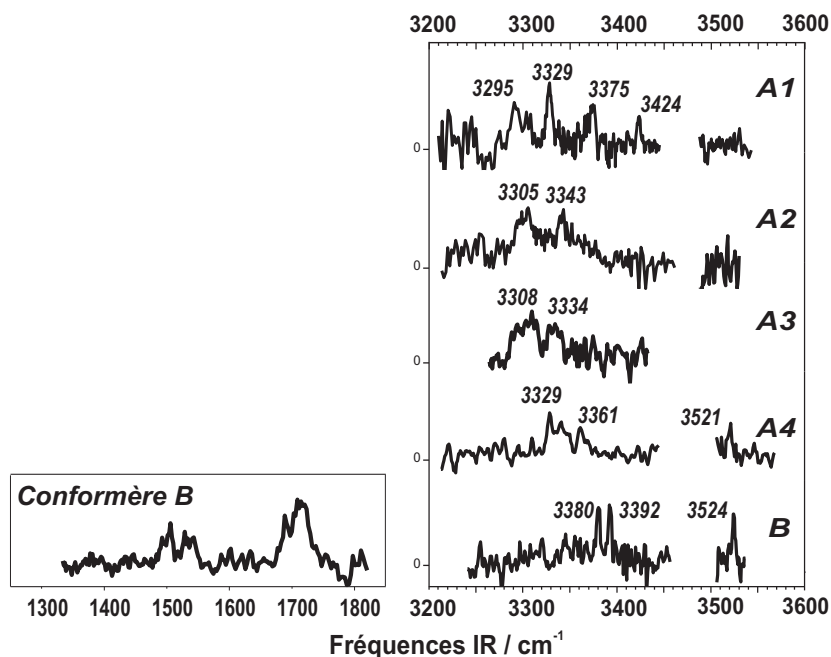


FIG. 5.3 – Spectres IR des conformères A1-A4 et B de Ac-Phe-Gly-Gly-NH<sub>2</sub> dans la région amide A, et spectre IR de B dans la région amide I-II. Les bandes IR manquantes dans la région amide A sont situées dans le domaine non couvert par le laser OPO.

des espèces A1-A4 présentent des bandes larges très décalées vers le rouge ( $\nu < 3400 \text{ cm}^{-1}$ ) associées à des liaisons hydrogène fortes, dont le profil de bande rappelle les liaisons C<sub>7</sub> du conformère A  $\gamma$ - $\gamma$  de Ac-Gly-Phe-NH<sub>2</sub>. La bande antisymétrique est absente des spectres de A1, A2 et A3, ce qui confirme une très forte perturbation sur les groupes NH<sub>2</sub>. Les structures présentant des coudes  $\gamma$  (simples, doubles ou triples) sont en effet susceptibles de reproduire les motifs de liaisons hydrogène observés mais des structures de type globulaire - c'est-à-dire possédant des liaisons hydrogène sont « non standard » car différentes des liaisons C<sub>7</sub> et C<sub>10</sub> - très compactes et présentant de multiples liaisons hydrogène, pourraient également reproduire le spectre IR. Les bandes IR manquantes dans chaque spectre - on devrait en effet mesurer cinq bandes - sont situées dans le domaine spectral non couvert par le cristal de l'OPO. Elles révèlent que les amides correspondants sont libres. Les bandes antisymétriques de A4 et B apparaissent à des fréquences proches de  $3520 \text{ cm}^{-1}$ , indiquant que les liaisons hydrogène associées sont plus faibles que pour les conformères A1-A3, et traduisent le fait que la perturbation sur NH<sub>2</sub> est similaire à tous les peptides étudiés jusqu'à présent. Le spectre du conformère A1 suggère que celui-ci possède trois liaisons hydrogène pouvant être associées à

Ac-Phe-Gly-Gly-NH <sub>2</sub>	Fréq. UV (cm <sup>-1</sup> )	Fréquences IR (cm <sup>-1</sup> )			
		Amide A		Amide I	
A1	37592	3295	3329	3375	3424 NH- $\pi$
A2	37606	3305	3343		
A3	37637	3308	3334		
A4	37653	3329	3361		3521
B	37648	3380	3392		3524. 1690-1720

TAB. 5.1 – Fréquences UV et IR des conformères A1-A4 et B identifiés pour le tripeptide Ac-Phe-Gly-Gly-NH<sub>2</sub>.

une structure  $\gamma$ - $\gamma$ - $\gamma$ , et présentant une interaction NH- $\pi$  (bande à 3424 cm<sup>-1</sup>). Les conformères A2-A4 peuvent correspondre à des structures possédant plusieurs coudes  $\gamma$  ou à des structures globulaires.

Le spectre IR de Phe-Gly-Gly B diffère de façon significative des précédents. Les bandes NH sont plus fines, leur position (3380 et 3392 cm<sup>-1</sup>) indique qu'elles sont associées à des liaisons hydrogène plus faibles que des liaisons C<sub>7</sub>. Les fréquences d'élongation sont en fait comparables à celles des coudes  $\beta$  mesurées pour les dipeptides dont les valeurs sont proches de 3390 cm<sup>-1</sup> (voir Chapitre 4). L'absence de bandes dans la région des amides perturbés (3420-3470 cm<sup>-1</sup>) révèle deux amides libres et de fait l'absence d'interaction NH- $\pi$ . La signature IR de B étant relativement singulière, nous consacrerons le paragraphe qui suit à son interprétation.

### 5.1.2 Conformère B

En plaçant les fréquences d'élongation NH de B sur le modèle de couplage vibrationnel, on constate que les deux bandes à 3380 et 3392 cm<sup>-1</sup> peuvent être correspondre la bande symétrique (Annexe A). En effet, le couple 3392/3520 cm<sup>-1</sup> correspond à un coude  $\beta$  sur la partie -Gly-Gly- de la séquence et impliquant le groupement NH<sub>2</sub>. Le couple 3380/3520 cm<sup>-1</sup> peut, lui, correspondre soit à un coude  $\beta$  comme dans le cas précédent, soit un coude  $\gamma$  faible

impliquant l'amide terminal. Il est à ce stade difficile d'associer l'une ou l'autre fréquence au mode d'élongation  $\text{NH}_2$  symétrique.

Deux grandes classes de structures peuvent rendre compte de la formation des deux liaisons hydrogène modérées : une structure présentant deux coudes  $\beta$  successifs ( $\text{C}_{10}\text{-C}_{10}$ ) et une structure « mixte » constituée d'une structure secondaire coude  $\beta$  et d'une forme locale coude  $\gamma$  ( $\text{C}_{10}\text{-C}_7$  ou  $\text{C}_7\text{-C}_{10}$ ). Nous avons alors mené une recherche conformationnelle détaillée en sélectionnant les structures répondant à ces critères (Figure 5.4). Pour les structures  $\text{C}_{10}\text{-C}_{10}$ , nous avons choisi des coudes  $\beta$  type I et type II pour la partie -Phe-Gly- par analogie avec l'étude théorique menée sur le dipeptide Ac-Phe-Gly- $\text{NH}_2$ . Les autres types de coudes  $\beta$  n'ont pas été pris en compte en raison de leur énergie plus élevée (les énergies relatives des structures calculées pour le dipeptide sont regroupées dans le Tableau 4.4 en Annexe). Seule la structure hélice  $3_{10}$  peut alors présenter deux coudes  $\beta$  successifs lorsque le premier coude  $\beta$  est de type I. Cette structure est en fait constituée de deux coudes  $\beta$  type I, dont le premier coude est légèrement déformé en raison de la géométrie périodique de l'hélice, on parle alors de type III (ou encore de type III(I)) pour faire référence à cette structure d'hélice  $3_{10}$ . Par ailleurs, nous avons constaté que la combinaison type I-I et l'hélice  $3_{10}$  convergeaient vers la même géométrie type III-I. Dans le cas d'un coude  $\beta$  type II en première position, seule la combinaison type II-I' mène à la formation d'une structure avec deux coudes  $\beta$  successifs. Les trois orientations de Phe ont été calculées pour l'hélice. Pour la combinaison type II-I', seules les orientations (a) et (g-), faisant partie des structures les plus stables pour le dipeptide Phe-Gly, ont été prises en compte. D'autres formes possédant deux liaisons hydrogène ont également été considérées :

- Une structure *globulaire* «  $\gamma_L(a)$  -  $\text{C}_{11}$  » possédant une liaison  $\text{C}_7$  centrale et une liaison hydrogène  $\text{C}_{11} \text{NH}_{\text{Phe}} \rightarrow \text{CO}_{\text{Gly}2}$ .
- Une structure «  $\gamma_L(g+)$  -  $\text{C}_{13}$  » avec un coude  $\gamma_L$  central et une liaison  $\text{C}_{13}$  reliant l'amide terminal au groupe  $\text{CO}_{\text{Ace}}$ , cette liaison hydrogène est spécifique à l'hélice  $\alpha$ .
- La structure  $\beta_L(a)\text{-}\gamma_L\text{-}\gamma_L$ , par analogie avec le conformère A de Phe-Gly, bien que le spectre IR de B ne montre pas de bande dans le domaine des amides perturbés, permettant ainsi de rejeter une géométrie étendue  $\beta_L$  sur le chromophore aromatique Phe. L'orientation (a) de Phe permet de stabiliser la forme  $\beta_L$  sur Phe.

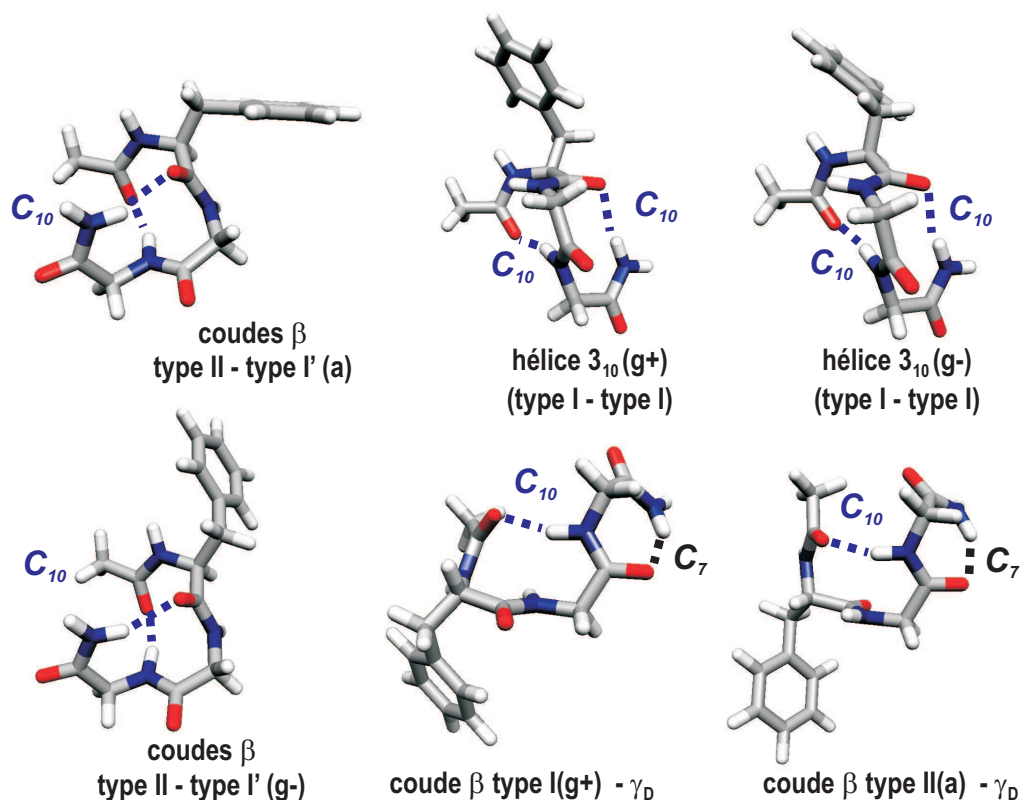


FIG. 5.4 – Échantillons de structures (B3LYP/6-31+G(d)) calculées pour Ac-Phe-Gly-Gly-NH<sub>2</sub>.

Enfin, nous avons aussi retenu les formes coude  $\gamma$  triple  $\gamma$ - $\gamma$ - $\gamma$  et  $\beta_L$ - $\beta_L$ - $\beta_L$  équivalente au brin  $\beta$  afin d'évaluer leur stabilité. Les calculs sont résumés dans le Tableau 5.2.

### Énergie des conformations

La relaxation conformationnelle, on l'a vu jusqu'à présent, est très efficace dans nos conditions expérimentales grâce au refroidissement par collision dans la détente. En conséquence, on s'attend à ne peupler que les conformations les plus stables. On peut raisonnablement éliminer les conformations situées à plus de 3-4 kcal/mol au dessus de la forme II(a) - I' qui est le minimum le plus stable parmi toutes les structures calculées. Ainsi on peut remarquer que le brin  $\beta$  est très défavorisé énergétiquement. Pareillement, les deux formes  $\beta_L$ - $\gamma_L$ - $\gamma_L$  et  $\gamma_L$ - $\gamma_D$ - $\gamma_D$  topologiquement proches du brin  $\beta$  sont également très hautes en énergie, montrant que les liaisons hydrogène, même fortes, ne sont pas les seuls facteurs responsables de la stabilité d'une structure. On constate en effet que les formes  $\gamma_L$ (g+) - C<sub>13</sub> et  $\gamma_L$ (a) - C<sub>11</sub> ont

des énergies élevées malgré la présence de deux liaisons hydrogène dont une liaison C<sub>7</sub>. Il faut cependant souligner que les liaisons C<sub>11</sub> et C<sub>13</sub> sont relativement faibles avec des fréquences de 3399 et 3439 cm<sup>-1</sup>, cette dernière étant d'ailleurs comparable aux interactions faibles NH- $\pi$  et C<sub>5</sub>. Parmi les formes mixtes coude  $\gamma$  - coude  $\beta$ , le coude  $\gamma_D$  sur Gly mène à des structures légèrement plus stables que le coude  $\gamma_L$ , la forme II(a) -  $\gamma_D$  notamment possède une énergie relativement basse. Parmi les successions de coudes  $\beta$ , l'hélice 3<sub>10</sub> (a) est significativement plus haute en énergie par rapport aux orientations (g+) et (g-) de Phe. L'énergie plus basse pour ces dernières est due à la présence d'une interaction NH- $\pi$  qui apporte une stabilisation supplémentaire. Enfin, conformément aux calculs effectués pour le système Ac-Phe-Gly-NH<sub>2</sub>, le coude  $\beta$  type II conduit à des formes particulièrement stables. Sur un critère énergétique, sont finalement retenues les structures suivantes : les combinaisons II(a) - I' et II(g-) - I', les hélices 3<sub>10</sub> (g-) et (g+), la forme II(a) -  $\gamma_D$  (et I(g+) -  $\gamma_D$ , et la forme  $\gamma_L$ (g+) - C<sub>13</sub>.)

### Amide A

La comparaison entre le spectre IR de B et le spectre théorique des structures citées ci-dessus dans le domaine amide A permet d'écarter les formes mixtes C<sub>10</sub>-C<sub>7</sub> dont les fréquences calculées sont beaucoup plus basses que celles mesurées. En particulier, comme dans le cas des calculs effectués pour Ac-Phe-Gly-NH<sub>2</sub> et Ac-Phe-Val-NH<sub>2</sub>, on peut noter que les fréquences C<sub>10</sub> sont plus rouges pour le coude  $\beta$  type II que pour le type I - l'hélice 3<sub>10</sub> ou coude  $\beta$  type III possède des angles dièdres  $\phi$  et  $\psi$  proches du type I. La forme  $\gamma_L$  - C<sub>13</sub> ne reproduit pas le spectre IR de B. La conformation C<sub>13</sub> correspond à une liaison hydrogène très faible (3439 cm<sup>-1</sup>), de force similaire aux interactions C<sub>5</sub> et NH- $\pi$ , ce que confirme aussi la fréquence d'élongation NH<sub>2</sub> antisymétrique (3552 cm<sup>-1</sup>) proche de la valeur d'un groupe NH<sub>2</sub> libre [6]. Parmi les structures possédant deux coudes  $\beta$ , les formes II-I' donnent un meilleur accord que l'hélice 3<sub>10</sub>. Les fréquences d'élongation NH trouvées pour l'hélice indiquent que les liaisons C<sub>10</sub> sont plus faibles que pour le type II, et en particulier la fréquence du mode NH<sub>2</sub> antisymétrique qui est significativement plus élevée ( $\sim 3537$  cm<sup>-1</sup>) que la valeur mesurée (3524 cm<sup>-1</sup>), alors que celle-ci a très bien été reproduite par les calculs jusqu'à présent. On peut donc supposer que le conformère B a pour structure une combinaison de deux coudes  $\beta$  successifs type II-I'.

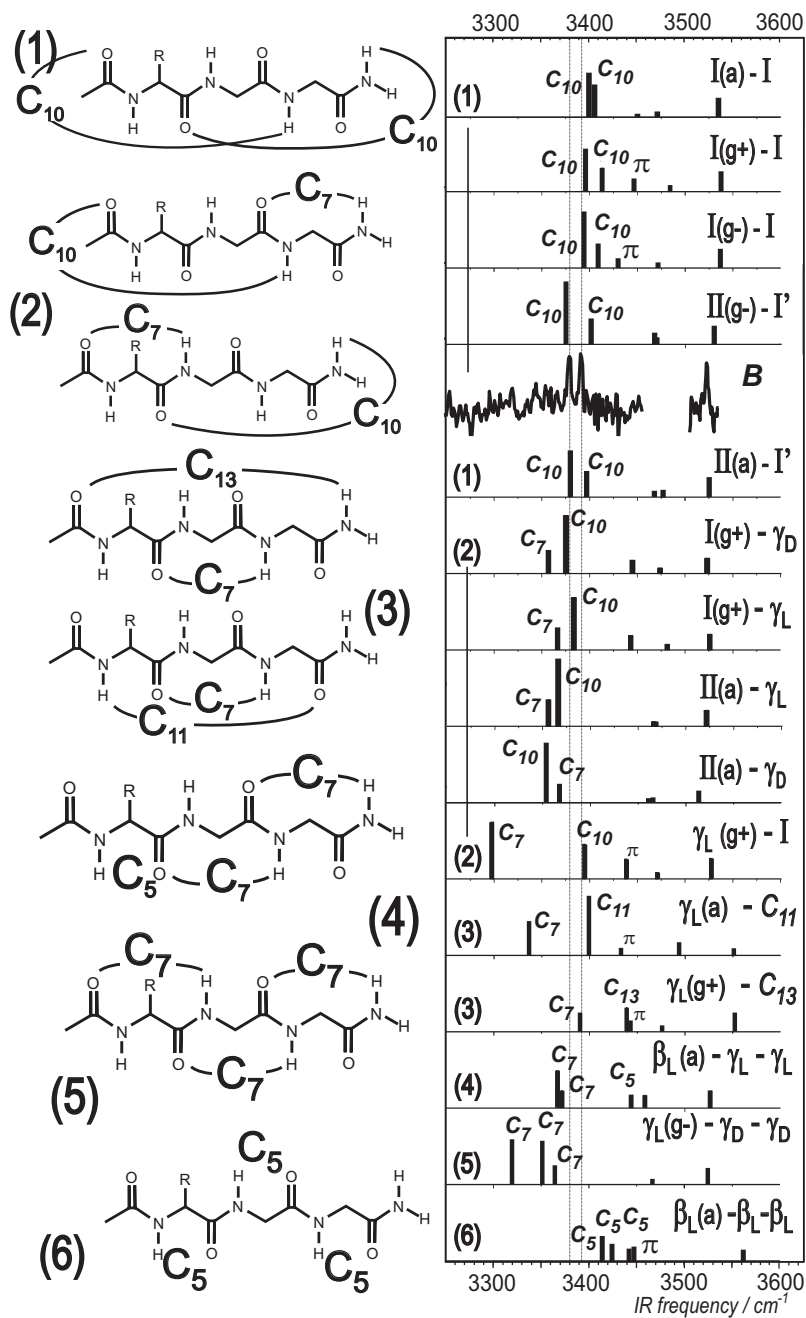


FIG. 5.5 – Spectre IR du conformère B de Ac-Phe-Gly-Gly-NH<sub>2</sub> dans la région amide A, comparé aux fréquences calculées pour une sélection de structures. Les fréquences sont calculées avec la méthode B3LYP/6-31+G(d) et sont corrigées avec le facteur 0.96 pour tenir compte de l'anharmonicité et du défaut de la base.



## Amide I

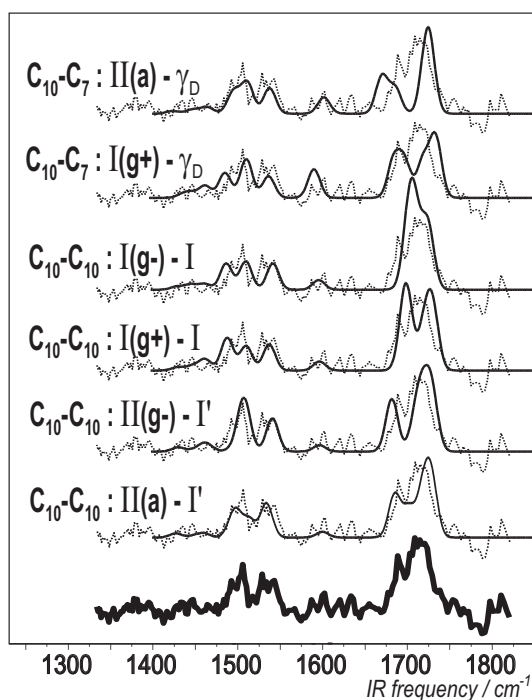


FIG. 5.6 – Spectre IR du conformère B de Ac-Phe-Gly-Gly-NH<sub>2</sub> dans les régions amide I et II, comparé aux fréquences calculées pour une sélection de structures. Les fréquences sont calculées avec la méthode B3LYP/6-31+G(d) et sont respectivement corrigées avec le facteur 0.98 et 0.97 dans les domaines amide I et II pour tenir compte de l'anharmonicité.

Le spectre IR de B est peu résolu dans le domaine amide I et II - qui correspondent respectivement aux modes d'élongation CO couplés aux modes d'élongation CN et aux modes de déformation NH, et aux modes de déformation NH couplés aux modes d'élongation CO et de déformation CH. Il présente des bandes larges centrées à 1715, 1530 et 1500 cm<sup>-1</sup> (Figure 5.6). Les spectres théoriques présentés ont été obtenus en convoluant les fréquences avec une fonction gaussienne d'une largeur à mi-hauteur de 15 cm<sup>-1</sup> afin de tenir compte de la largeur spectrale du faisceau IR FELIX qui est de 15 cm<sup>-1</sup>. Les bandes IR expérimentales dans la région amide II sont peu intenses en raison du faible rapport signal/bruit, mais la présence de deux bandes permet d'écarter les formes type I(g-)- $\gamma_D$  et les hélices dont le spectre théorique indique trois bandes. Le spectre du conformère B montre des bandes plus intenses dans la région amide I. Les spectres théoriques montrent que les structures possédant des coudes  $\gamma$  ont des bandes CO très décalées vers le rouge et très espacées les unes des autres, rappelant les

résultats obtenus sur le système Ac-Phe-NH<sub>2</sub> et sur les dipeptides. L'écart théorique entre les bandes extrêmes est par exemple de 50-60 cm<sup>-1</sup> pour les formes mixtes C<sub>10</sub>-C<sub>7</sub>. À l'inverse, les géométries avec deux coudes  $\beta$  dont les liaisons hydrogène sont plus faibles que les liaisons C<sub>7</sub> forment des motifs CO plus compacts avec un écart de 30-40 cm<sup>-1</sup> entre les bandes CO extrêmes. Comme dans la région amide A, on observe un décalage vers le rouge plus marqué pour les coudes  $\beta$  type II que pour le type I ou l'hélice  $3_{10}$ . La combinaison type II(a)-I' ainsi que l'hélice  $3_{10}$  (g+) reproduisent le mieux le spectre de B.

Les signatures IR expérimentales et théoriques dans les régions amide A et amide I et II nous permettent en conséquence d'attribuer le conformère B à la combinaison de deux coudes  $\beta$  successifs type II(a)-I'. L'attribution est de plus confortée par les calculs d'énergie qui indiquent que la structure type II(a)-I' est la forme la plus stable.

### 5.1.3 Conclusion

Nous avons montré dans cette partie que le tripeptide Ac-Phe-Gly-Gly-NH<sub>2</sub> possédait au moins cinq conformères dont les signatures IR indiquent qu'il s'agit de structures différentes. Les conformères A1-A4 peuvent être attribués à des formes repliées et très compactes possédant des liaisons hydrogène très fortes telles que des structures composées de coudes  $\gamma$  multiples ou bien des structures globulaires. Celles-ci présentent des liaisons hydrogène différentes des liaisons standard C<sub>7</sub>, C<sub>10</sub> ou C<sub>5</sub> mises en évidence pour le système Ac-Phe-NH<sub>2</sub> et pour les dipeptides.

La signature IR du cinquième conformère (B) étant très différente, nous avons axé notre étude sur celui-ci. Parce que plusieurs structures étaient susceptibles de correspondre à ce conformère, nous avons mené une recherche conformationnelle détaillée. Parmi toutes les structures considérées, nous avons vu que celles constituées de deux coudes  $\beta$  successifs étaient les meilleures candidates. En combinant les spectres IR expérimentaux et théoriques dans trois régions d'absorption IR, amide A, amide I et amide II, nous avons attribué le conformère B à une combinaison de coude  $\beta$  type II(a)-I' qui s'est en outre révélée être la structure la plus stable. Finalement, le fait que cette structure soit plus stable que l'hélice  $3_{10}$  est dû à la présence du résidu Gly. Le coude  $\beta$  type II est en effet spécifique au résidu Gly placé en

Phe-Gly-Gly		Fréquences expérimentales du conformère B								
		<i>a</i>	<i>a</i>			1690-1720				
Conformations calculées	$\Delta E$ (kcal/mol)	Fréquences théoriques								
		Amide A			Amide I					
		NH Phe	NH Gly1	NH Gly2	NH <sub>2</sub> sym.	NH <sub>2</sub> anti.	CO Ace	CO Phe	CO Gly1	CO Gly2
<b>coude <math>\beta</math> -</b>										
<b>coude <math>\beta</math></b>										
C <sub>10</sub> -C <sub>10</sub>										
II(a)-I'	0	3468 l.	3477 l.	3380 C <sub>10</sub>	3397 C <sub>10</sub>	3525	1686	1702	1719	1728
II(g <sup>-</sup> )-I'	1.50	3468 l.	3470 l.	3375 C <sub>10</sub>	3401 C <sub>10</sub>	3530	1681	1709	1718	1728
3 <sub>10</sub> (g <sup>+</sup> )	2.19	3430 $\pi$	3472 l.	3394 C <sub>10</sub>	3409 C <sub>10</sub>	3537	1700	1708	1718	1727
3 <sub>10</sub> (g <sup>-</sup> )	2.83	3446 $\pi$	3484 l.	3396 C <sub>10</sub>	3413 C <sub>10</sub>	3538	1696	1699	1720	1729
3 <sub>10</sub> (a)	5.65	3451 l.	3472 lib.	3400 C <sub>10</sub>	3406 C <sub>10</sub>	3536	1706	1713	1719	1728
<b>coude <math>\gamma</math> -</b>										
<b>coude <math>\beta</math></b>										
C <sub>7</sub> -C <sub>10</sub>										
II(a) - $\gamma_D$	1.61	3461 l.	3466 l.	3354 C <sub>10</sub>	3368 C <sub>7</sub>	3517	1671	1722	1687	1727
II(a) - $\gamma_L$	4.05	3466 l.	3469 l.	3366 C <sub>10</sub>	3356 C <sub>7</sub>	3522	1683	1691	1724	1733
I(g <sup>+</sup> ) - $\gamma_D$	3.88	3443 $\pi$	3481 l.	3384 C <sub>10</sub>	3366 C <sub>7</sub>	3526	1683	1698	1717	1735
I(g <sup>+</sup> ) - $\gamma_L$	4.43	3444 $\pi$	3474 l.	3375 C <sub>10</sub>	3357 C <sub>7</sub>	3523	1684	1697	1718	1733
C <sub>10</sub> -C <sub>7</sub>										
$\gamma_L$ (g <sup>+</sup> ) - I	5.19	3438 $\pi$	3298 C <sub>7</sub>	3471 l.	3394 C <sub>10</sub>	3527	1675	1697	1727	1733
<b>2 liaisons H</b>										
$\gamma_L$ (g <sup>+</sup> ) - C <sub>13</sub>	3.89	3443 $\pi$	3476 l.	3390 C <sub>7</sub>	3439 C <sub>13</sub>	3526	1683	1698	1717	1735
C <sub>7</sub> // C <sub>13</sub>										
$\gamma_L$ (a) - C <sub>11</sub>	7.79	3399 C <sub>11</sub>	3494 l.	3337 C <sub>7</sub>	3433 l.	3551	1675	1702	1720	1729
C <sub>7</sub> // C <sub>11</sub>										
$\beta_L$ (a)- $\gamma_L$ - $\gamma_L$	5.31	3444 C <sub>5</sub>	3458 $\pi$	3366 C <sub>7</sub>	3371 C <sub>7</sub>	3526	1672	1698	1702	1737
C <sub>5</sub> -C <sub>7</sub> -C <sub>7</sub>										
<b>3 liaisons H</b>										
$\gamma_L$ (g <sup>-</sup> )- $\gamma_D$ - $\gamma_D$	5.92	3467 l.	3337 C <sub>7</sub>	3352 C <sub>7</sub>	3365 C <sub>7</sub>	3525	1677	1687	1699	1735
$\gamma_L$ (g <sup>-</sup> )- $\gamma_L$ - $\gamma_L$	6.31	3467 l.	3326 C <sub>7</sub>	3349 C <sub>7</sub>	3364 C <sub>7</sub>	3525	1678	1688	170	1735
C <sub>7</sub> -C <sub>7</sub> -C <sub>7</sub>										
<b>0 liaison H</b>										
$\beta_L$ (a)- $\beta_L$ - $\beta_L$	8.43	3446 C <sub>5</sub>	3423 C <sub>5</sub>	3413 C <sub>5</sub>	3441 l.	3561	1700	1685	1712	1740
C <sub>5</sub> -C <sub>5</sub> -C <sub>5</sub>										

TAB. 5.2 – Énergies (MP2) et fréquences B3LYP/6-31+G(d) d'élongation NH (amide A) et CO (amide I) d'une sélection de structures calculées pour Ac-Phe-Gly-Gly-NH<sub>2</sub>. Les fréquences sont respectivement corrigées avec les facteurs 0.96 et 0.98 pour tenir compte de l'anharmonicité. <sup>a</sup> Les bandes manquantes sont situées dans le domaine spectral non couvert par le cristal de l'OPO.

deuxième position dans le coude comme nous l'avons observé pour le dipeptide Ac-Phe-Gly-NH<sub>2</sub> dans le chapitre précédent. À cela s'ajoute le fait que la petite taille du résidu ainsi que l'absence de chaîne latérale sur Gly rendent la partie -Gly-Gly- du tripeptide très flexible et permettent à la chaîne peptidique d'avoir accès à de nombreuses conformations d'énergies voisines, pouvant expliquer la formation d'au moins cinq conformères significativement peuplés pour Ac-Phe-Gly-Gly-NH<sub>2</sub>.

Enfin, soulignons que les calculs d'énergie relative indiquent que de nouvelles conformations sont accessibles, en particulier les hélices  $3_{10}$  et la structure possédant un motif C<sub>13</sub> qui ont des énergies plus élevées que la structure type II(a)-I' de seulement 3-4 kcal/mol. Ainsi, bien que ces formes ne soient pas observées dans le cas de Ac-Phe-Gly-Gly-NH<sub>2</sub> en tant qu'espèces significativement peuplées, l'étude de systèmes différents peut peut-être nous permettre de les former.

## 5.2 Effet de la séquence : tripeptides -Ala-Ala-Ala-

En continuité à l'étude sur le tripeptide Ac-Phe-Gly-Gly-NH<sub>2</sub>, nous avons intégré l'acide aminé Ala dans la séquence car c'est le résidu qui possède la plus grande propension à former des hélices en solution. Nous avons donc mené une étude théorique sur le tripeptide Ac-Ala-Ala-Ala-NH<sub>2</sub>. Néanmoins, l'objectif de cette étude n'est pas d'explorer de façon exhaustive le paysage conformationnel du système mais d'évaluer la stabilité de plusieurs familles conformationnelles, en particulier les formes hélicoïdales. Aussi nous sommes-nous appuyés sur les résultats obtenus sur Ac-Phe-Gly-Gly-NH<sub>2</sub> dans le choix des conformations à calculer, à savoir des conformations présentant les interactions et liaisons hydrogène « classiques » de la biologie (C<sub>5</sub>, C<sub>7</sub>, C<sub>10</sub>). La présence du résidu Phe nous est indispensable pour la partie expérimentale puisqu'il joue le rôle de chromophore UV. Il y a alors trois possibilités de placer le résidu dans le système de référence Ac-Ala-Ala-Ala-NH<sub>2</sub> : Ac-**Phe**-Ala-Ala-NH<sub>2</sub>, Ac-Ala-**Phe**-Ala-NH<sub>2</sub> et Ac-Ala-Ala-**Phe**-NH<sub>2</sub>. Nous verrons dans la discussion que, bien que ces systèmes soient voisins - ils ne diffèrent que par la position du résidu Phe - ils présentent des signatures spectroscopiques spécifiques. Le changement de position de Phe nous permettra alors de dé-

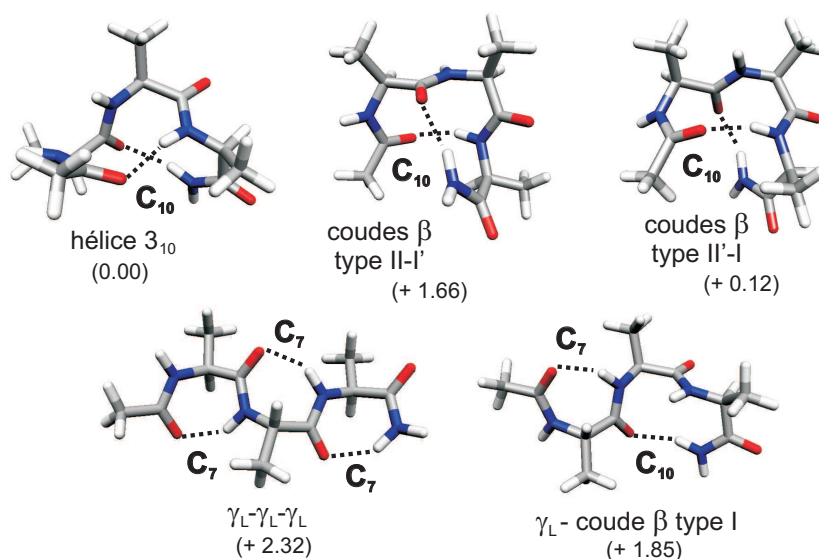


FIG. 5.7 – Géométries B3LYP/6-31+G(d) de l'hélice  $3_{10}$ , des combinaisons type II-I' et type II'I, et des formes  $\gamma_L$ - type I et  $\gamma_L$ - $\gamma_L$ - $\gamma_L$  de Ac-Ala-Ala-Ala-NH<sub>2</sub>.

terminer l'influence du cycle aromatique sur la conformation globale de la chaîne peptidique. L'attribution s'appuiera également sur la comparaison avec le système Ac-Ala-Ala-Ala-NH<sub>2</sub> ainsi que sur des calculs de structures, de fréquences et d'énergie relative pour chaque système.

Nous exposerons dans la section qui suit les calculs réalisés sur le tripeptide Ac-Ala-Ala-Ala-NH<sub>2</sub>. Nous discuterons ensuite des résultats expérimentaux et théoriques obtenus sur les trois systèmes Ac-Phe-Ala-Ala-NH<sub>2</sub>, Ac-Ala-Phe-Ala-NH<sub>2</sub> et Ac-Ala-Ala-Phe-NH<sub>2</sub>. Leur spectroscopie électronique sera présentée dans une même partie, puis nous examinerons séparément leur spectroscopie IR.

### 5.2.1 Étude théorique du tripeptide Ac-Ala-Ala-Ala-NH<sub>2</sub>

Nous avons sélectionné des structures présentant deux coudes  $\beta$  successifs, une structure coude  $\gamma$  triple correspondant à la succession de formes locales  $\gamma_L$  pour le résidu Ala - la conformation  $\beta_L$  n'a pas été prise en compte car elle est déstabilisée par rapport au coude  $\gamma_L$  pour les résidus non aromatiques [7, 8] - et une structure « mixte » composée d'un coude  $\gamma$  et d'un coude  $\beta$ ,  $\gamma_L$  - type I (Figure 5.7). Celle-ci est une combinaison entre une forme locale centrée sur un résidu, un coude  $\gamma_L$  sur Ala, et une structure secondaire embrassant deux

Ac-Ala-Ala-Ala-NH <sub>2</sub>	$\Delta E$ MP2	$\Delta E$ B3LYP
3 <sub>10</sub>	0.00	1.42
type II-I	0.12	2.93
type II-I'	1.66	4.94
$\gamma_L$ - type I	1.85	0.92
$\gamma_L$ - $\gamma_L$ - $\gamma_L$	2.32	0.00

TAB. 5.3 – Énergies relatives (kcal/mol) calculées au niveau MP2 et B3LYP avec la base 6-31+G(d). Les énergies tiennent compte de la correction d'énergie de point zéro évaluée au niveau B3LYP.

résidus, un coude  $\beta$  sur la deuxième partie de la séquence. L'énergie de cette conformation nous permettra d'évaluer s'il y a compétition entre les formes locales et les structures secondaires des différentes sous-parties de la séquence. Les combinaisons de coudes  $\beta$  type II-I' et type III-I, cette dernière étant équivalente à l'hélice 3<sub>10</sub>, ont été considérées car elles sont très stables pour le système Ac-Phe-Gly-Gly-NH<sub>2</sub> (Phe-Gly-Gly). Nous avons de plus calculé la combinaison type II-I car elle est géométriquement très proche de l'hélice 3<sub>10</sub>, elle ne diffère que par l'orientation du groupement acétyle. La structure en hélice  $\alpha$  qui, ici, ne présenterait qu'une seule liaison C<sub>13</sub>, n'a pas été prise en compte car de nombreux travaux théoriques indiquent que huit résidus sont nécessaires à sa formation [3, 4, 5, 9]. Les calculs de structures ont été effectués avec la méthode B3LYP/6-31+G(d) et les calculs d'énergie relative au niveau MP2 avec la même base, les résultats sont regroupés dans le Tableau 5.3.

L'ordre énergétique est différent de celui trouvé pour Phe-Gly-Gly ; on peut notamment signaler que l'ordre est inversé entre la combinaison type II-I' et l'hélice 3<sub>10</sub>, au niveau MP2. Celle-ci est la forme la plus stable avec la combinaison type II-I, considérée comme isoénergétique. Les formes  $\gamma_L$  - type I et  $\gamma_L$ - $\gamma_L$ - $\gamma_L$  sont stabilisées pour le système Ala-Ala-Ala par rapport à Phe-Gly-Gly. Une énergie plus élevée pour la combinaison type II-I' n'est en fait pas surprenante car les coudes  $\beta$  type II sont presque exclusivement formés pour le résidu Gly placé en deuxième position du coude dans les protéines, comme c'est le cas pour les systèmes Phe-Gly et Phe-Gly-Gly. La présence de chaîne latérale pour les autres résidus peut conduire à des effets d'encombrement stérique. L'hélice 3<sub>10</sub>, en revanche, est topologiquement proche

d'un coude  $\beta$  type I qui est le type le plus fréquemment rencontré dans les protéines. Par ailleurs, on peut noter que l'hélice  $3_{10}$  calculée est en fait légèrement déformée en fin de séquence et converge vers un type I comme dans le cas Phe-Gly-Gly. La stabilité particulière de la combinaison type II'-I est peut-être moins attendue car le type II' est également spécifique à Gly lorsque celui-ci est placé en première position du coude. Cependant, la présence d'un coude  $\beta$  type I à la suite du coude  $\beta$  type II' rend cette conformation topologiquement proche de l'hélice  $3_{10}$ , les deux formes ne diffèrent que par l'angle  $\phi$  du premier résidu Ala (cf paramètres géométriques du Tableau C.2 en Annexe). Nos calculs sont en accord avec les résultats de Barone et collaborateurs qui trouvent, par des calculs en DFT réalisés avec la fonctionnelle PBE, que l'hélice  $3_{10}$  est très stable après l'hélice  $\alpha$  pour des chaînes de polygly, polyAla et polyAib [1]. On doit toutefois souligner que ces calculs sont réalisés sur des chaînes de polypeptides infinies utilisant des conditions périodiques, expliquant le fait que l'hélice  $\alpha$  soit plus stable même pour de petites chaînes, alors que de nombreux calculs indiquent qu'il faut au minimum huit résidus pour stabiliser cette conformation [9, 10]. À l'inverse, d'autres calculs effectués avec la méthode Hartree-Fock, ainsi qu'en DFT avec la fonctionnelle B3LYP sur des tripeptides trouvent que les formes les plus stables sont les coudes  $\gamma$  multiples et les structures étendues de type brin  $\beta$  pour lequel les résidus adoptent tous une conformation  $\beta_L$  [11]. Cette différence avec nos calculs peut s'expliquer par la méthode employée. En effet, nous avons vu dans le chapitre précédent que la prise en compte des effets de dispersion pouvait modifier l'ordre énergétique. Ainsi, les calculs d'énergie avec la fonctionnelle B3LYP a-t-elle tendance à favoriser les préférences conformationnelles locales autour de chaque résidu (Tableau 5.3), ainsi que le brin  $\beta$  alors que ces structures sont déstabilisées au niveau MP2 (voir aussi les Tableaux 4.4 et 4.3 en Annexe).

Dans notre étude sur les systèmes Ac-Phe-Ala-Ala-NH<sub>2</sub>, Ac-Ala-Phe-Ala-NH<sub>2</sub> et Ac-Ala-Ala-Phe-NH<sub>2</sub>, on peut s'attendre à observer les formes les plus stables trouvées dans l'étude théorique de Ac-Ala-Ala-Ala-NH<sub>2</sub> (au niveau MP2), c'est-à-dire l'hélice  $3_{10}$  et les combinaisons de coudes  $\beta$  type II'-I et type II-I', ainsi que la forme mixte  $\gamma_L$  - type I qui sont situées dans un domaine d'énergie inférieure à 2 kcal/mol au-dessus de l'hélice  $3_{10}$ .

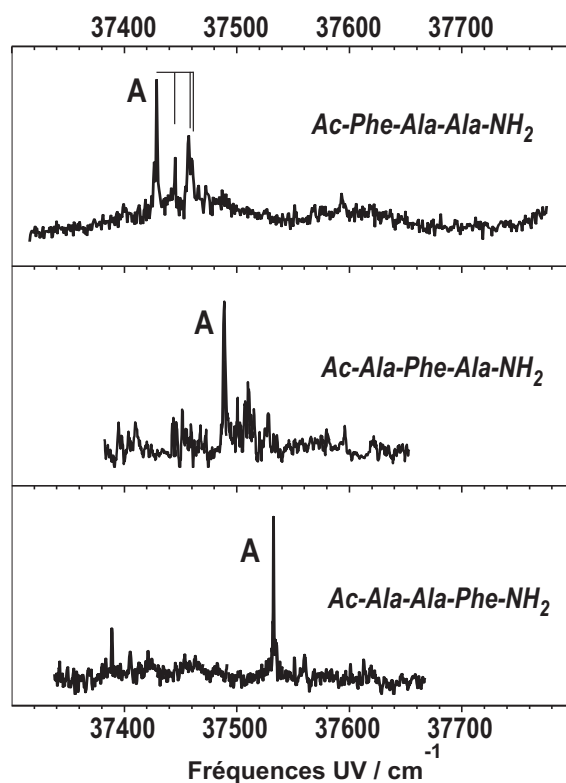


FIG. 5.8 – Spectre d’excitation des tripeptides  $\text{Ac-Phe-Ala-Ala-NH}_2$ ,  $\text{Ac-Ala-Phe-Ala-NH}_2$  et  $\text{Ac-Ala-Ala-Phe-NH}_2$ . Les bandes origine sont identifiées par la lettre A.

### 5.3 Étude expérimentale des tripeptides

Les spectres d’excitation des tripeptides  $\text{Ac-Phe-Ala-Ala-NH}_2$ ,  $\text{Ac-Ala-Phe-Ala-NH}_2$  et  $\text{Ac-Ala-Ala-Phe-NH}_2$  présentent des bandes fines, intenses et peu nombreuses suggérant que peu de conformères sont peuplés dans le jet, contrairement au tripeptide  $\text{Ac-Phe-Gly-Gly-NH}_2$  dont le spectre congestionné contient la contribution de plusieurs espèces (Figure 5.8). Ainsi, malgré la complexité attendue du paysage conformationnel pour un système de cette taille, un seul conformère principal, noté A, a été identifié par les mesures de double résonance IR-UV. On peut également remarquer des bandes d’absorption plus faibles à divers endroits des spectres mais aucune mesure en IR n’a pu être effectuée en raison de la faible intensité du signal. On ne peut donc pas exclure la présence d’autres conformères moins peuplés.

Le spectre électronique de  $\text{Phe-Ala-Ala}$  présente une bande fine et intense située à  $37429 \text{ cm}^{-1}$  suivie d’une progression vibrationnelle de  $16 \text{ cm}^{-1}$  qui rappelle celles observées pour les conformères  $\beta_L(a)$ - $\gamma$  des dipeptides  $\text{Ac-Phe-Xxx-NH}_2$  en termes de domaine spectral et de séparation



entre chaque bande. Cette signature UV suggère fortement que le résidu Phe, ici placé en première position dans la séquence, adopte une conformation  $\beta_L$  pour laquelle les groupements  $\text{NH}_{Phe}$  et  $\text{CO}_{Phe}$  sont reliés par une interaction  $\text{C}_5$ . On peut remarquer que l'origine de la progression vibrationnelle est décalée de  $35 \text{ cm}^{-1}$  vers le rouge par rapport au dipeptide Ac-Phe-Ala-NH<sub>2</sub>. Une bande fine et très intense, notée A, située à  $37489 \text{ cm}^{-1}$  domine le spectre de Ala-Phe-Ala. Les structures situées dans le bleu de cette bande pourraient constituer une progression vibrationnelle.

Le spectre de Ala-Ala-Phe présente essentiellement une bande fine et intense, attribuée au conformère principal A, située à  $37532 \text{ cm}^{-1}$ . L'absence de progression vibrationnelle semble indiquer que le résidu Phe n'adopte pas la même conformation que le système Phe-Ala-Ala.

Le profil des bandes d'absorption ainsi que le domaine spectral de ces trois tripeptides sont très différents de ceux de Phe-Gly-Gly, montrant que le cycle aromatique interagit de façon différente avec le squelette. Cela signifie que les espèces formées sont vraisemblablement différentes de celles observées pour Phe-Gly-Gly. Enfin, on constate que les bandes origine des conformères A sont décalées vers le bleu à mesure que le résidu Phe est placé en première, deuxième puis troisième position dans la séquence. Les spectres de dépopulation IR-UV ont été enregistrés dans la région amide A ; ils montrent tous des signatures différentes et seront détaillés dans les paragraphes suivants. Comme dans le cas de l'étude réalisée sur le système Ac-Phe-Gly-Gly-NH<sub>2</sub> où le tripeptide était comparé aux résultats expérimentaux et théoriques de Ac-Phe-Gly-NH<sub>2</sub>, les spectres IR des tripeptides constitués avec les résidus Ala seront aussi comparés à des dipeptides, en l'occurrence Ac-Ala-Phe-NH<sub>2</sub> et Ac-Phe-Ala-NH<sub>2</sub>, afin de raisonner par analogie (Tableau 5.4).

### 5.3.1 Conservation des formes locales

Le spectre IR de Phe-Ala-Ala a été enregistré en fixant le laser de sonde UV sur la bande origine  $0_0^0$  qui est la plus intense. Il indique cinq bandes NH conformément au nombre de groupements amide, à savoir trois liaisons peptidiques et le groupement NH<sub>2</sub> terminal (Figure 5.9). Les bandes UV de la progression vibrationnelle ont été identifiées comme appar-

Tripeptides	Fréq. UV ( $\text{cm}^{-1}$ )	Fréquences IR expérimentales ( $\text{cm}^{-1}$ )				
		NH (i+1)	NH (i+2)	NH (i+3)	NH <sub>2</sub> sym	NH <sub>2</sub> anti
Phe-Ala-Ala	37429	3423 C <sub>5</sub>	3447 NH- $\pi$	3338 C <sub>7</sub>	3363 C <sub>7</sub>	3522
Ala-Phe-Ala	37489	<sup>a</sup>	3429 NH- $\pi$	3374 C <sub>10</sub>	3408 C <sub>10</sub>	3534
Ala-Ala-Phe	37532	<sup>a</sup>	3303 C <sub>7</sub>	3440 NH- $\pi$	3389 C <sub>7</sub>	3524
<b>Dipeptides</b>						
Ala-Phe						
$\gamma$ - $\gamma$	37477	<sup>a</sup>	3360 C <sub>7</sub>		3303 C <sub>7</sub>	3520
coude $\beta$	37560	<sup>a</sup>	3440 NH- $\pi$		3389 C <sub>10</sub>	3520
Phe-Ala	37464	3422 C <sub>5</sub>		3445 NH- $\pi$		3370 C <sub>7</sub> 3521
$\beta_L(a)$ - $\gamma_L$						

TAB. 5.4 – Fréquences UV et fréquences IR dans la région amide A (modes d'élongation NH) des tripeptides Ac-Phe-Ala-Ala-NH<sub>2</sub>, Ac-Ala-Phe-Ala-NH<sub>2</sub> et Ac-Ala-Ala-Phe-NH<sub>2</sub>, comparées aux données expérimentales des dipeptides Ac-Ala-Phe-NH<sub>2</sub> et Ac-Phe-Ala-NH<sub>2</sub>. <sup>a</sup> Les bandes manquantes sont situées dans la zone non accessible au laser OPO.

tenant au conformère A en vérifiant de façon « statique » les déplétions IR mesurées sur la bande origine, c'est-à-dire en fixant la pompe IR sur une absorption du conformère A et la sonde sur chaque bande UV de la progression. Les deux bandes larges situées à des fréquences inférieures à  $3400 \text{ cm}^{-1}$  révèlent la présence de deux liaisons hydrogène relativement fortes. La fréquence du mode d'élongation antisymétrique du groupe NH<sub>2</sub> ( $3522 \text{ cm}^{-1}$ ) indique que celui-ci est impliqué dans une des deux liaisons hydrogène. Deux bandes plus fines sont observées dans le domaine des amides légèrement perturbés ( $3420\text{-}3470 \text{ cm}^{-1}$ ). Or, la spectroscopie électronique suggère que le résidu Phe adopte une conformation  $\beta_L$ , l'une des deux bandes est donc attribuée à une interaction C<sub>5</sub>, la bande restante peut être associée à une interaction C<sub>5</sub> ou NH- $\pi$ .

Deux structures sont susceptibles de reproduire les deux liaisons hydrogène, une structure constituée de deux coudes  $\gamma$  et une structure possédant deux coudes  $\beta$  (voir schéma de la Figure 5.9). La structure coude  $\gamma$  double (C<sub>7</sub>-C<sub>7</sub>) est compatible avec la conformation  $\beta_L$  sur Phe et conduit à une forme  $\beta_L$ - $\gamma$ - $\gamma$  ou encore C<sub>5</sub>-C<sub>7</sub>-C<sub>7</sub> en termes d'interaction. En revanche, la formation de l'interaction C<sub>5</sub> sur le résidu Phe n'est pas compatible avec une structure coude  $\beta$  double mais avec un coude  $\beta$  simple, ce qui conduirait à la formation d'une seule liaison hydrogène. On peut donc écarter les structures coudes  $\beta$  double et simple.

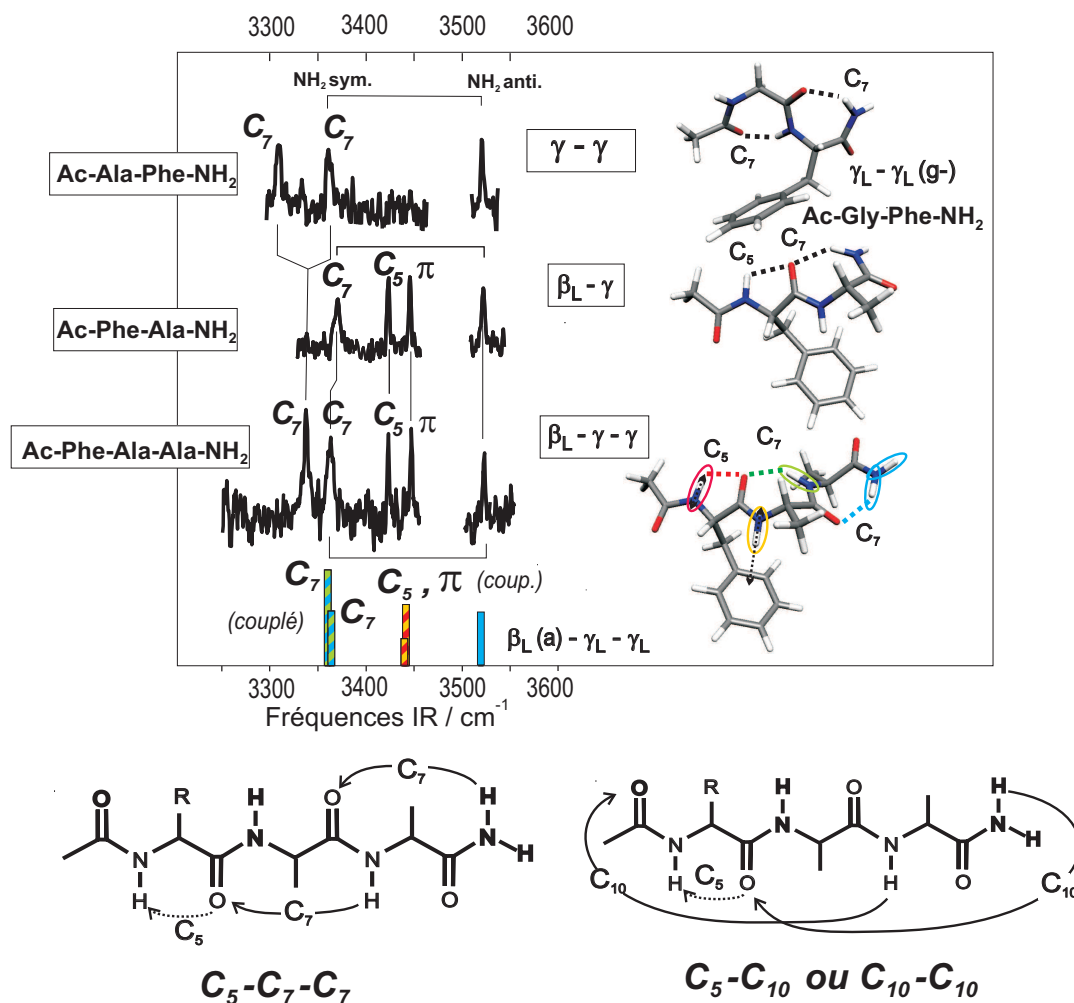


FIG. 5.9 – Spectre IR de Ac-Phe-Ala-Ala-NH<sub>2</sub> dans la région amide A comparé aux spectres expérimentaux des conformères  $\gamma$ - $\gamma$  de Ac-Ala-Phe-NH<sub>2</sub> et  $\beta_L(a)$ - $\gamma_L$  de Ac-Phe-Ala-NH<sub>2</sub> (haut de la figure) et aux fréquences théoriques de la forme  $\beta_L(a)$ - $\gamma_L$ - $\gamma_L$  calculée avec la méthode B3LYP/6-31+G(d) (bas de la figure). Les fréquences théoriques sont corrigées avec le facteur 0.96 pour tenir compte de l'anharmonicité.

**Comparaison avec les systèmes Ac-Ala-Phe-NH<sub>2</sub> et Ac-Phe-Ala-NH<sub>2</sub>** La comparaison du spectre IR de Phe-Ala-Ala avec le conformère  $\beta_L(a)$ - $\gamma_L$  de Phe-Ala est frappante (Figure 5.9) : hormis la bande la plus rouge de Phe-Ala-Ala, les spectres sont quasi identiques. On retrouve les bandes d'élongation NH<sub>2</sub> antisymétrique ainsi que les bandes C<sub>5</sub> et NH- $\pi$  à des fréquences quasi inchangées. Cette comparaison permet de plus d'attribuer le mode symétrique à la bande située à 3368 cm<sup>-1</sup>. La fréquence de celle-ci, légèrement plus basse que celle de Phe-Ala, indique que la liaison C<sub>7</sub> qui correspond au résidu Ala en dernière position est un peu plus forte. Ce décalage en fréquence traduit en fait la perturbation induite par l'ajout

d'un acide aminé Ala dans la séquence. La chaîne principale semble donc adopter la préférence conformationnelle locale autour de chaque résidu, à savoir  $\beta_L$  pour Phe et coude  $\gamma_L$  pour Ala. La bande la plus rouge de Phe-Ala-Ala peut vraisemblablement être attribuée à une liaison hydrogène  $C_7$  menant alors à une conformation  $\beta_L$ - $\gamma_L$ - $\gamma_L$ . La liaison  $C_7$  du coude  $\gamma_L$  sur le premier résidu Ala (de la partie  $\gamma$ - $\gamma$  du tripeptide) mène cependant à une liaison hydrogène plus faible que celle de la conformation coude  $\gamma$  double du dipeptide Ala-Phe qui génère des bandes NH très décalées vers le rouge ( $3308\text{ cm}^{-1}$ ). Ceci est dû au fait que les deux coudes  $\gamma$  ont une liaison peptidique en commun, entraînant la polarisation de celle-ci. Ainsi, on peut s'attendre à ce que cet effet de polarisation soit plus élevé dans le cas de Ala-Phe que dans le cas de la sous-partie Ala-Ala du tripeptide, ce qui conduit à une liaison  $C_7$  plus forte sur le résidu Ala du dipeptide Ala-Phe.

**Calculs** Nous avons alors comparé le spectre IR de Phe-Ala-Ala aux fréquences calculées pour la forme  $\beta_L$ (a)- $\gamma_L$ - $\gamma_L$ . Il est important de rappeler qu'en l'absence de cycle aromatique (benzène, phénol ou indole), la forme  $\beta_L$  ne serait pas observée car l'interaction  $C_5$  ne suffit pas à stabiliser la forme étendue. Nous avons vu en revanche que des effets de chaîne latérale permettaient de favoriser cette forme pour le résidu Phe en apportant une interaction NH- $\pi$  supplémentaire via l'orientation (a) du cycle aromatique. Les fréquences théoriques donnent un accord relativement correct avec le spectre expérimental bien que les fréquences des liaisons  $C_7$  ne soient pas très bien reproduites (bas de la Figure 5.9). En effet, comme dans le cas des calculs effectués sur le coude  $\gamma$ - $\gamma$  du dipeptide Ac-Gly-Phe-NH<sub>2</sub> (voir la section 3.3 du Chapitre 3, p. 101), la polarisation de la liaison peptidique n'est pas prise en compte par les calculs. Ce défaut est imputé à la méthode de calcul employée. De plus, les modes normaux sont des mélanges des modes locaux d'élongation NH<sub>Ala2</sub> et NH<sub>2</sub>. Ces mélanges résultent de la quasi-dégénérescence, accidentelle, des modes locaux ( $\nu_{th} \sim 3360\text{ cm}^{-1}$ ) et conduisent à des combinaisons correspondant à des élongations en phase et en opposition de phase des deux modes, la bande la plus intense étant l'élongation en opposition de phase. Un tel mélange est aussi retrouvé pour les modes locaux NH<sub>Phe</sub> et NH<sub>Ala1</sub> correspondant aux amides impliqués dans les interactions  $C_5$  et NH- $\pi$  et qui apparaissent quasi-dégénérés ( $\nu_{th} \sim 3440\text{ cm}^{-1}$ ).

Les ressemblances entre les signatures IR de Phe-Ala-Ala et des dipeptides Phe-Ala/Ala-Phe laissent supposer que l'ajout d'un troisième résidu Ala dans la séquence affecte finalement peu le comportement de la chaîne. Plus précisément, on a l'impression que la conformation de la chaîne correspond à la somme des préférences conformationnelles locales de chaque brique élémentaire, à savoir les formes locales  $\beta_L$  pour Phe et  $\gamma_L$  pour Ala, sans interaction entre elles comme dans le cas des dipeptides. On peut souligner que la conformation  $\beta_L\text{-}\gamma_L\text{-}\gamma_L$  serait très défavorisée dans le cas du système Ala-Ala-Ala car la forme  $\beta_L$  est plus haute de l'ordre de 2 kcal/mol au-dessus du minimum  $\gamma_L$  pour le résidu Ala isolé. De plus, l'étude théorique menée sur Ala-Ala-Ala montre que la conformation  $\gamma_L\text{-}\gamma_L\text{-}\gamma_L$ , qui est topologiquement proche de  $\beta_L\text{-}\gamma_L\text{-}\gamma_L$ , est également déstabilisée de 2.32 kcal/mol par rapport au minimum (Tableau 5.3). On peut dès lors attribuer la formation de la structure  $\beta_L\text{-}\gamma_L\text{-}\gamma_L$  aux effets de chaîne latérale de Phe qui apportent une interaction NH- $\pi$  entre le cycle aromatique et la chaîne principale permettant ainsi de stabiliser les formes locales autour de chaque résidu.

### 5.3.2 Vers l'hélice $3_{10}$

Le système Ala-Ala-Phe présente la séquence inversée par rapport à Phe-Ala-Ala ; la séquence est comparable au système Ala-Phe auquel on a ajouté un résidu Ala en début de séquence. On peut alors se demander si on observera des structures constituées d'une juxtaposition de conformations locales sur chaque résidu comme pour le conformère A  $\gamma\text{-}\gamma$  du dipeptide Ala-Phe ou si l'on observera le coude  $\beta$  trouvé pour l'espèce B du dipeptide.

Le spectre IR dans le domaine amide A présente quatre bandes (Figure 5.10). Il manque une bande NH qui est vraisemblablement située dans la région non accessible au laser OPO révélant qu'un groupement amide est libre. Les deux bandes larges et intenses dans la partie rouge du spectre sont associées à une liaison hydrogène forte ( $3303\text{ cm}^{-1}$ ) et une modérée ( $3389\text{ cm}^{-1}$ ). L'une des deux liaisons implique le groupe  $\text{NH}_2$ , comme l'indique la fréquence de la bande antisymétrique ( $3524\text{ cm}^{-1}$ ) légèrement décalée par rapport à un groupe  $\text{NH}_2$  libre [12]. La bande fine à  $3440\text{ cm}^{-1}$  peut appartenir à une interaction  $\text{C}_5$  sur le résidu Phe ou bien une interaction NH- $\pi$ .

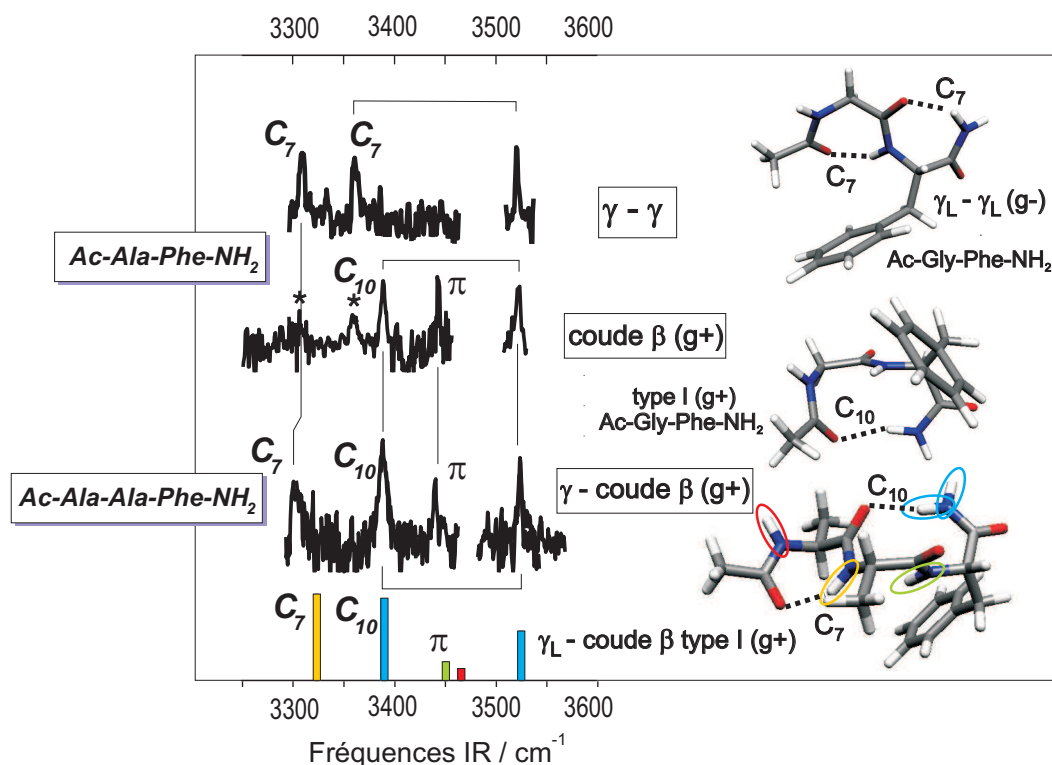


FIG. 5.10 – Spectres IR dans la région amide A du tripeptide  $\text{Ac-Ala-Ala-Phe-NH}_2$  et des conformères  $\gamma$ - $\gamma$  et coude  $\beta$  du dipeptide  $\text{Ac-Ala-Phe-NH}_2$ , comparés aux fréquences théoriques de la structures  $\gamma_L$ -type I (g+) calculée pour le tripeptide. Les fréquences (B3LYP/6-31+G(d)) sont corrigées avec le facteur 0.96 pour tenir compte de l'anharmonicité. Aucun calcul n'ayant été mené pour  $\text{Ac-Ala-Phe-NH}_2$ , les formes  $\gamma_L$ - $\gamma_L$ (g-) et coude  $\beta$  type I (g+) de  $\text{Ac-Gly-Phe-NH}_2$  sont montrées à titre d'illustration.

**Comparaison avec le système  $\text{Ac-Ala-Phe-NH}_2$**  Le spectre IR de  $\text{Ala-Ala-Phe}$  est très similaire à celui du conformère coude  $\beta$  du dipeptide  $\text{Ala-Phe}$ . Excepté la bande la plus rouge, les bandes apparaissent à des fréquences quasi identiques (Figure 5.10). Le modèle de couplage vibrationnel indique que la bande d'élongation symétrique correspond à la bande intense située à  $3389 \text{ cm}^{-1}$  (Annexe A). Or, ces deux fréquences apparaissent dans la région des coudes  $\beta$ . La bande à  $3389 \text{ cm}^{-1}$  correspond alors à l'amide  $\text{NH}_{Ala2}$  impliqué dans la liaison  $C_{10}$  du coude  $\beta$  sur la partie  $-\text{Ala-Phe-}$  du tripeptide, et la bande située à  $3440 \text{ cm}^{-1}$  est compatible avec le groupement  $\text{NH}_{Phe}$  qui pointe vers le phényle via une interaction  $\text{NH}-\pi$ . La fréquence, très basse, de la bande la plus rouge ( $3303 \text{ cm}^{-1}$ ) est, elle, analogue à celle de la structure  $\gamma$ - $\gamma$  de  $\text{Ala-Phe}$  ( $3308 \text{ cm}^{-1}$ ) (Figure 5.10). On peut l'attribuer au groupe amide  $\text{NH}_{Ala1}$  lié dans une liaison hydrogène  $C_7$ . Comme précédemment, un tel décalage vers le rouge est dû à

la polarisation de la liaison peptidique entre les résidus Ala1 et Ala2 jouant à la fois le rôle de donneur de proton vis à vis du coude  $\gamma$  et d'accepteur vis à vis du coude  $\beta$ . En outre, la position de la bande UV ( $37532 \text{ cm}^{-1}$ ) est proche de l'origine du coude  $\beta$  de Ala-Phe ( $37560 \text{ cm}^{-1}$ ) (voir Figure 4.1 du Chapitre 4, p. 120).

On peut par conséquent proposer que le conformère A de Ala-Ala-Phe adopte la structure mixte  $C_7$ - $C_{10}$ , à savoir un coude  $\gamma$  sur le premier résidu Ala suivi d'un coude  $\beta$  sur la sous-partie Ala-Phe.

**Calculs et attribution** Nous avons comparé le spectre expérimental de Ala-Ala-Phe avec les spectres théoriques de forme mixte coude  $\gamma$  - coude  $\beta$ . Nous avons sélectionné la forme  $\gamma_L$  pour le coude  $\gamma$  sur le résidu Ala1 et un coude  $\beta$  de type I pour la partie Ala-Phe car c'est le type le plus couramment observé dans les protéines [13]. Dans cette configuration, le cycle aromatique doit avoir une orientation ( $g+$ ) pour interagir avec la chaîne principale via une interaction  $\text{NH}-\pi$ , et ainsi reproduire la bande située à  $3440 \text{ cm}^{-1}$ . Les résultats sont présentés dans le bas de la Figure 5.10. On note un très bon accord entre le spectre expérimental et le spectre calculé, bien que la force de la liaison  $C_7$  soit légèrement sous-estimée par les calculs. Ceci peut être dû au fait que les coudes  $\gamma_L$  et  $\beta$  partagent une même liaison amide, polarisant ainsi la liaison peptidique, et, comme dans le cas des coudes  $\gamma$ - $\gamma$  et des interactions de dispersion  $C_5$  et  $\text{NH}-\pi$ , cet effet n'est pas toujours convenablement pris en compte par la DFT.

On peut donc attribuer au conformère A de Ala-Ala-Phe la forme  $\gamma_L$  - coude  $\beta$  type I ( $g+$ ). Contrairement au conformère  $\beta_L$ - $\gamma_L$ - $\gamma_L$  de Phe-Ala-Ala, cette structure n'est pas constituée de la somme des préférences conformationnelles locales autour de chaque résidu mais présente une compétition entre une forme locale  $\gamma_L$ , et une structure secondaire coude  $\beta$  au sein de la même séquence.

### 5.3.3 Émergence de l'hélice $3_{10}$

La séquence du tripeptide Ala-Phe-Ala est peut-être la plus intéressante des trois tripeptides. Suite aux résultats obtenus sur les systèmes précédents, on peut en effet se demander

si on observera uniquement les formes locales pour chaque résidu comme c'est le cas pour Phe-Ala-Ala, ou s'il y aura un phénomène de compétition entre les sous-parties -Ala-Phe- et -Phe-Ala- de la chaîne conduisant à la formation d'un coude  $\beta$ , ou encore si de nouvelles structures telles que l'hélice  $3_{10}$  seront accessibles.

Le spectre IR du conformère A de Ala-Phe-Ala présente quatre bandes fines et intenses (Figure 5.11). La bande manquante, située dans le domaine non accessible par le laser OPO est attribuée à un groupement amide libre. Les bandes situées dans la partie rouge du spectre appartiennent à deux liaisons hydrogène de force modérée, dont l'une implique le groupe  $\text{NH}_2$ . Leur position ( $\sim 3380 \text{ cm}^{-1}$ ) rappelle la signature IR des coudes  $\beta$  dont les fréquences apparaissent à des valeurs proches  $3390 \text{ cm}^{-1}$ . La bande  $\text{NH}_2$  antisymétrique ( $3534 \text{ cm}^{-1}$ ) est significativement décalée vers le bleu par rapport aux tripeptides et en général par rapport à **tous** les systèmes étudiés jusqu'à présent, c'est-à-dire les coudes  $\beta$ , les coudes  $\gamma$  simples et doubles, ainsi que le système Ac-Phe- $\text{NH}_2$ , qui montraient des fréquences d'élongation  $\text{NH}_2$  antisymétrique dans le domaine  $3515\text{-}3525 \text{ cm}^{-1}$ . Ce décalage en fréquence révèle que la liaison hydrogène sur le groupement  $\text{NH}_2$  est faible, bien que compatible avec l'hypothèse des coudes  $\beta$ , et ne peut donc correspondre qu'à la bande située à  $3408 \text{ cm}^{-1}$ . Les combinaisons de deux coudes  $\beta$  ( $\text{C}_{10}\text{-C}_{10}$ ) et les structures mixtes coude  $\gamma$  - coude  $\beta$  ( $\text{C}_7\text{-C}_{10}$  ou  $\text{C}_{10}\text{-C}_7$ ) sont donc des candidats potentiels. La bande fine située à  $3429 \text{ cm}^{-1}$  peut provenir soit d'un amide impliqué dans une interaction  $\text{NH}\text{-}\pi$ , soit du groupe  $\text{NH}_{Phe}$  impliqué dans une interaction  $\text{C}_5$  sur le résidu Phe.

Le modèle de couplage vibrationnel indique que la composante symétrique doit être associée à la bande située à  $3408 \text{ cm}^{-1}$ , les fréquences  $\nu_{sym}$  et  $\nu_{anti}$  apparaissant dans la région des coudes  $\beta$  à liaison  $\text{C}_{10}$  faible, en fait la plus faible parmi tous les coudes  $\beta$  observés (Annexe A). Les bandes correspondent donc à une liaison  $\text{C}_{10}$  sur la sous-partie -Phe-Ala- reliant les groupes  $\text{CO}_{Ala1}$  et  $\text{NH}_2$ . La bande fine à  $3429 \text{ cm}^{-1}$  située dans la région des interactions faibles ( $\text{C}_5$  ou  $\text{NH}\text{-}\pi$ ) est alors attribuée à une interaction  $\text{NH}\text{-}\pi$  entre le cycle et le groupe  $\text{NH}_{Phe}$  ou  $\text{NH}_{Ala1}$  suivant le type du coude  $\beta$ . La bande située à  $3374 \text{ cm}^{-1}$  peut, elle, correspondre à une liaison  $\text{C}_7$  ou à une liaison  $\text{C}_{10}$ .



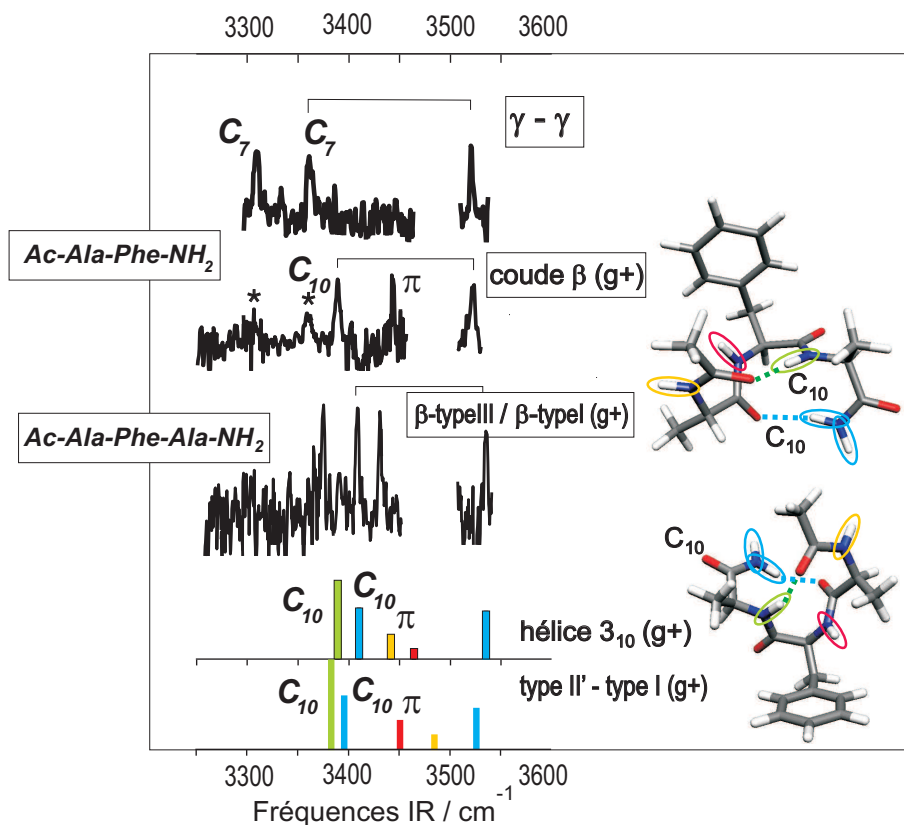


FIG. 5.11 – Spectres IR dans la région amide A du tripeptide  $\text{Ac-Ala-Phe-Ala-NH}_2$  et des conformères  $\gamma\text{-}\gamma$  et coude  $\beta$  de  $\text{Ac-Ala-Phe-NH}_2$  et du conformère  $\beta_L(a)\text{-}\gamma_L$  de  $\text{Ac-Phe-Ala-NH}_2$ . En bas de la figure, fréquences et calculées pour l'hélice  $3_{10}$  (g+) avec la méthode B3LYP/6-31+G(d). Les fréquences théoriques sont corrigées avec le facteur 0.96 pour tenir compte de l'anharmonicité. La structure coude  $\beta$  type I (g+) de  $\text{Ac-Gly-Phe-NH}_2$  est montrée à titre d'illustration.

**Comparaison avec le dipeptide  $\text{Ac-Ala-Phe-NH}_2$  et calculs** Les conclusions déduites de la discussion menée sur les systèmes  $\text{Ala-Ala-Phe}$  et  $\text{Ala-Phe}$  nous permettent d'affirmer qu'un coude  $\gamma$  sur le résidu  $\text{Ala}$  en première position dans la séquence mène à une liaison  $\text{C}_7$  forte dont la fréquence d'élongation  $\text{NH}$  est d'environ  $3303\text{ cm}^{-1}$ . La bande située à  $3374\text{ cm}^{-1}$  dans le spectre IR de  $\text{Ala-Phe-Ala}$  n'est donc pas compatible avec un coude  $\gamma$  sur le premier résidu  $\text{Ala}$ . Cette bande correspond donc plus vraisemblablement à la liaison  $\text{C}_{10}$  d'un coude  $\beta$  reliant les groupes  $\text{CO}_{\text{Ace}}$  et  $\text{NH}_{\text{Ala}2}$  (Figure 5.11). On peut de plus souligner que cette bande est relativement rouge pour un groupement amide impliqué dans une liaison  $\text{C}_{10}$ .

Afin de reproduire les deux liaisons hydrogène modérées observées dans le spectre de  $\text{Ala-Phe-Ala}$ , nous avons sélectionné l'hélice  $3_{10}$  et la combinaison de coude  $\beta$  type II'-I qui sont les structures les plus stables trouvées par les calculs pour le tripeptide de référence  $\text{Ala-Ala-Ala}$ .

Le coude  $\beta$  type II' correspond au réarrangement de l'extrémité N-terminale par rapport au type I, la combinaison type II'-I ne diffère donc de l'hélice  $3_{10}$  que par l'orientation de la première liaison peptidique par rapport au reste de la molécule. La combinaison type II'-I a été trouvée plus haute en énergie pour Ala-Ala-Ala (cf Tableau 5.3, p. 161), elle n'a donc pas été considérée. L'orientation (g+) de la chaîne latérale de Phe est choisie afin d'établir une interaction NH- $\pi$  entre la chaîne principale et le cycle aromatique. Les deux structures donnent un bon accord avec le spectre expérimental de Ala-Phe-Ala, cet accord est meilleur dans le cas de l'hélice  $3_{10}$ . La fréquence de la bande NH<sub>2</sub> antisymétrique mesurée à 3534 cm<sup>-1</sup> est notamment **très bien reproduite** par les calculs menés sur l'hélice qui trouvent une valeur de 3535 cm<sup>-1</sup>. On peut en effet insister sur le fait que la combinaison type II'-I(g+) montre une fréquence d'élongation antisymétrique significativement plus basse, d'une valeur de 3525 cm<sup>-1</sup>, semblable aux fréquences de tous les coudes  $\beta$  trouvés jusqu'à présent, expérimentalement et par les calculs. On peut également remarquer que l'écart expérimental entre les deux bandes C<sub>10</sub> (34 cm<sup>-1</sup>) n'est pas retrouvé par les calculs effectués pour la combinaison type II'(g+)-I qui trouvent écart trop faible (13 cm<sup>-1</sup>). De plus, les calculs d'énergie au niveau MP2 sur les géométries optimisées en DFT indiquent que cette structure est déstabilisée de 1.39 kcal/mol par rapport à l'hélice  $3_{10}$  qui est la forme la plus stable parmi celles calculées. Le conformère A de Ala-Phe-Ala adopte donc très probablement une structure hélice  $3_{10}$ .

## 5.4 Discussion : effet du résidu aromatique

Nous avons présenté dans la partie qui précède des calculs de fréquences pour une sélection de conformations afin de conforter l'interprétation déduite des spectres IR et du modèle de couplage vibrationnel pour chaque tripeptide, mais nous avons également considéré les structures qui ont été calculées pour le système de référence Ac-Ala-Ala-Ala-NH<sub>2</sub>. Dans cette section, nous traiterons donc des calculs d'un point de vue énergétique afin de déterminer si la présence du résidu aromatique, et de fait sa position dans la séquence, a une influence sur la stabilité des différentes structures. Nous avons donc considéré les familles conformationnelles représentées par des combinaisons de coudes  $\beta$ , des coudes  $\gamma$  triples et les formes mixtes coude  $\gamma$  - coude  $\beta$

Tripeptides	$\Delta E$ MP2 (kcal/mol)
<b>Phe-Ala-Ala</b>	
$3_{10}$ (g+)	0.00
$3_{10}$ (g-)	0.85
$3_{10}$ (a)	3.71
type II-I' (a)	1.80
type II-I' (g-)	3.94
$\beta_L$ (a)- $\gamma_L$ - $\gamma_L$	1.92
$\gamma_L$ (g+)- $\gamma_L$ - $\gamma_L$	3.74
$\gamma_L$ (g-)- $\gamma_L$ - $\gamma_L$	4.26
<b>Ala-Ala-Phe</b>	
$3_{10}$ (g-)	0.00
$\gamma_L$ - type I (g+)	0.11
$\gamma_L$ - type I (g-)	2.94
$\gamma_L$ - $\gamma_L$ - $\gamma_L$ (g-)	2.50
<b>Ala-Phe-Ala</b>	
$3_{10}$ (g+)	0.00
$3_{10}$ (g-)	3.24
type II'-I (g+)	1.39
type II'-I (g-)	3.66
$\gamma_L$ - type I (g+)	3.31

TAB. 5.5 – Énergies relatives calculées au niveau MP2/6-31+G(d) à partir des géométries optimisées au niveau B3LYP/6-31+G(d) de trois types de combinaisons de coudes  $\beta$  (hélice  $3_{10}$ , type II'-I, type II-I'), de la conformation  $\gamma_L$ - $\gamma_L$ - $\gamma_L$  et de la forme  $\gamma_L$  - type I, pour les tripeptides Ac-Phe-Ala-Ala-NH<sub>2</sub>, Ac-Ala-Ala-Phe-NH<sub>2</sub> et Ac-Ala-Phe-Ala-NH<sub>2</sub>. Les énergies sont corrigées de la ZPE calculée au niveau B3LYP/6-31+G(d).

(Tableau 5.5). Le choix des structures s'appuyant toutefois sur les interprétations déduites des données expérimentales, toutes les structures calculées pour Ac-Ala-Ala-Ala-NH<sub>2</sub> n'ont pas été prises en compte pour les trois autres tripeptides. Rappelons enfin que les formes les plus stables (parmi les structures calculées) pour le système Ac-Ala-Ala-Ala-NH<sub>2</sub> sont l'hélice  $3_{10}$  et la combinaison de coudes  $\beta$  type II'-I qui sont isoénergétiques, suivies de la conformation  $\gamma_L$  - type I (C<sub>7</sub>-C<sub>10</sub>) et de la combinaison type II-I'. La conformation  $\gamma_L$ - $\gamma_L$ - $\gamma_L$  qui correspond à la somme des formes locales sur le résidu Ala est la forme la plus haute en énergie.

#### 5.4.1 Phe-Ala-Ala

Les calculs sont les plus nombreux pour la séquence Phe-Ala-Ala qui est le premier de la série des tripeptides à avoir été étudié. L'hélice  $3_{10}$  a été calculée avec les trois orientations

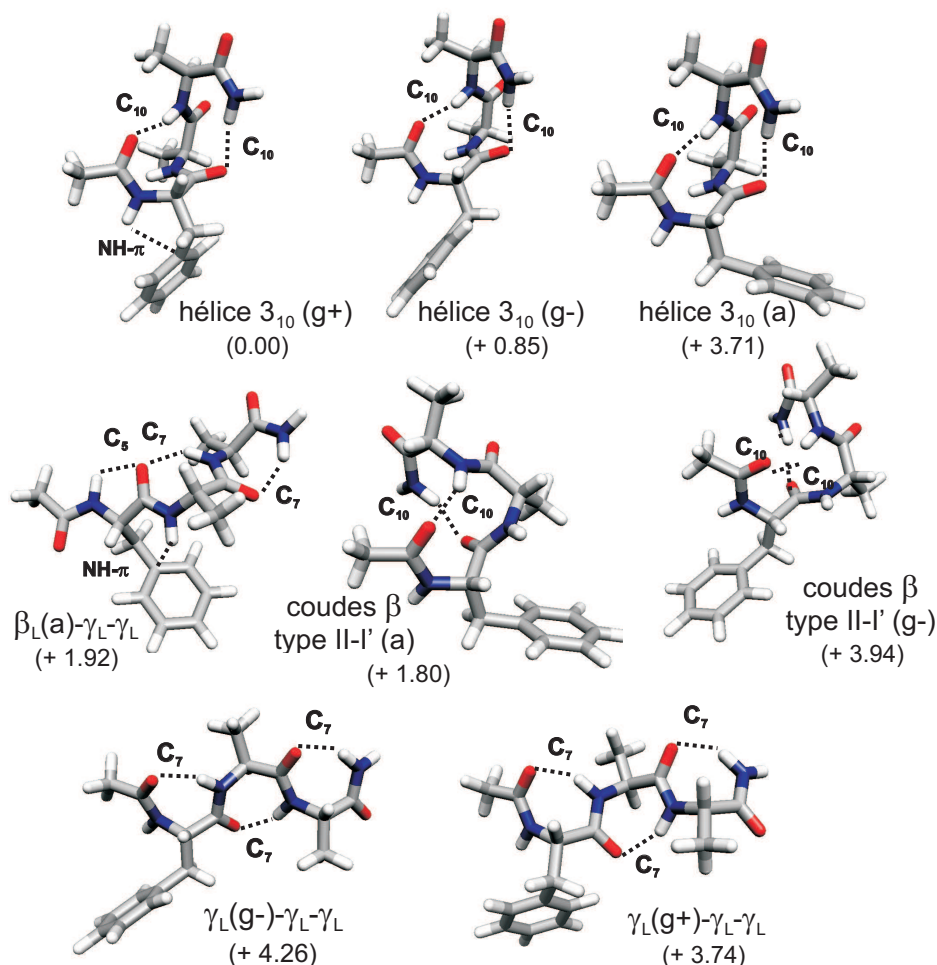


FIG. 5.12 – Structures calculées pour Ac-Phe-Ala-Ala-NH<sub>2</sub> avec la méthode B3LYP/6-31+G(d). Les énergies relatives sont calculées au niveau single point MP2 avec la même base.

du cycle aromatique. La combinaison type II-I' avec les orientations (a) et (g+) du phényle, par analogie avec le système Phe-Gly-Gly (voir Tableau 5.2, p. 158), a aussi été calculée bien qu'elle soit plus haute en énergie pour Ala-Ala-Ala. Les coudes triples  $\gamma$  (g+) et (g-) ont aussi été considérés (Figure 5.12).

Conformément aux résultats trouvés pour le système Ala-Ala-Ala, les structures de plus basse énergie sont les hélices  $3_{10}$  (g+) et (g-); elles sont séparées de 0.85 kcal/mol et peuvent être considérées comme isoénergétiques. Suivent la combinaison de coudes  $\beta$  type II(a)-I' et la forme  $\beta_L(a)$ - $\gamma_L$ - $\gamma_L$  à des énergies d'environ 1.80 kcal/mol plus élevées; les autres structures sont situées à plus de 3 kcal/mol au-dessus du minimum (Tableau 5.5). Quelques observations

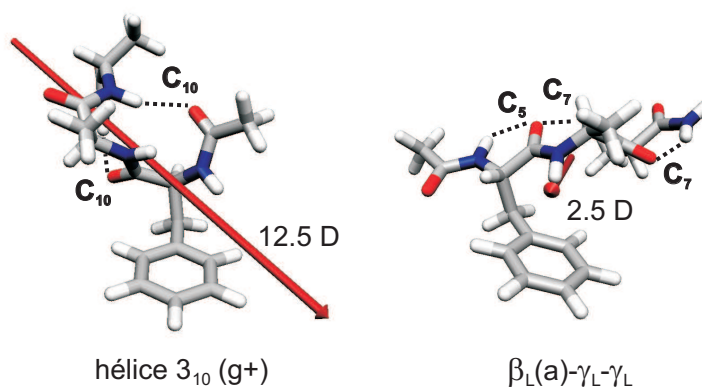


FIG. 5.13 – Structures hélice  $3_{10}$  (g+) et  $\beta_L(a)\text{-}\gamma_L\text{-}\gamma_L$  calculées pour Ac-Phe-Ala-Ala-NH<sub>2</sub> avec la méthode B3LYP/6-31+G(d). La direction des moments dipolaires (en debye) est indiquée par la flèche rouge.

essentielles peuvent être faites : la présence du résidu Phe ne modifie pas l'ordre énergétique global ; en revanche, l'orientation de la chaîne latérale a une grande influence sur les énergies relatives. On remarque en effet que les orientations menant à la formation d'une interaction NH- $\pi$  entraînent une stabilisation des structures concernées. Ainsi, l'hélice  $3_{10}$  (a) qui ne conduit à aucune interaction supplémentaire est très déstabilisée par rapport aux deux autres orientations. L'interaction NH- $\pi$  entre l'amide NH<sub>Phe</sub> et le cycle aromatique permet aussi de stabiliser la forme  $\beta_L(a)\text{-}\gamma_L\text{-}\gamma_L$  par rapport aux coudes triples  $\gamma$ . La structure  $\beta_L(a)\text{-}\gamma_L\text{-}\gamma_L$  correspond alors à la juxtaposition des formes locales pour chaque résidu, ce qui n'est pas le cas de la structure  $\gamma_L\text{-}\gamma_L\text{-}\gamma_L$  puisque le coude  $\gamma_L$  n'est que la deuxième forme locale pour le résidu Phe isolé (cf Chapitre 3). On peut aussi signaler que les énergies relatives sont davantage comparables au tripeptide Ala-Ala-Ala qu'à Phe-Gly-Gly puisque l'ordre énergétique est également inversé par rapport à Phe-Gly-Gly où la combinaison type II(a)-I' était le minimum et l'hélice  $3_{10}$  plus haute de 2.19 kcal/mol. Ceci est dû au fait que la présence d'une chaîne latérale pour le résidu Ala entraîne des effets d'encombrement stérique qui défavorisent alors la formation d'un coude  $\beta$  type II.

Les calculs d'énergie relative corroborent donc l'attribution du conformère de Phe-Ala-Ala à une structure  $\beta_L(a)\text{-}\gamma_L\text{-}\gamma_L$ . Cette conformation est plus haute en énergie que l'hélice  $3_{10}$  qui est la forme la plus stable, mais elle peut néanmoins être formée dans le jet. En revanche, l'absence de l'hélice  $3_{10}$  est beaucoup plus surprenante. L'hélice et la forme  $\beta_L(a)\text{-}\gamma_L\text{-}\gamma_L$  ap-

partiennent à deux bassins conformationnels assez différents et sont a priori séparées par une barrière suffisamment haute pour ne pas s'interconvertir. Cependant, la forme  $\beta_L(a)-\gamma_L-\gamma_L$  est topologiquement très proche de la structure étendue brin  $\beta$  qui est favorisée entropiquement à haute température [14], alors que les facteurs enthalpiques favorisent l'hélice [4]. En conséquence, la formation de  $\beta_L-\gamma_L-\gamma_L$  peut être favorisée par les facteurs entropiques. Une autre explication possible peut provenir de problèmes de photoionisation dus à l'interaction entre le dipôle total de l'hélice et le cycle aromatique. Dans une structure hélicoïdale, les liaisons peptidiques se placent de sorte que les moments dipolaires - le moment dipolaire de la liaison peptidique est de 3.7 Debye - soient alignés selon l'axe de l'hélice, les groupements carbonyles pointant en direction de l'extrémité N-terminale, ce qui conduit à un dipôle total élevé. C'est le cas de l'hélice  $3_{10}$  bien que la légère déviation des moments dipolaires par rapport à l'axe de l'hélice mène à une valeur du dipôle plus petite que pour l'hélice  $\alpha$ . Le dipôle total calculé pour l'hélice  $3_{10}$  (g+) est de 12.5 D, ce qui est relativement élevé. Il se trouve que le dipôle pointe directement vers le phényle, ce qui entraîne une interaction déstabilisante entre celui-ci et le dipôle dans l'état ionique : en effet, le processus de photoionisation consiste à éjecter un électron de l'orbitale  $\pi$  liante du cycle aromatique. L'ion est alors chargé positivement, la charge étant délocalisée sur tout le cycle. Or, l'interaction électrostatique entre le moment dipolaire et cette charge est déstabilisante et repousse l'ion vers les hautes énergies, réduisant ainsi la zone Franck-Condon accessible. La valeur de la section efficace d'ionisation est alors très faible, ce qui pourrait nous empêcher d'ioniser efficacement l'hélice  $3_{10}$ . Ce problème n'interviendrait pas dans le cas de la conformation  $\beta_L(a)-\gamma_L-\gamma_L$  dont le dipôle est plus faible (2.5 D) et donc n'interagirait pas avec la charge portée par le cycle dans l'ion (Figure 5.13). Ceci pourrait expliquer pourquoi nous observons la structure  $\beta_L(a)-\gamma_L-\gamma_L$  pour le tripeptide Ac-Phe-Ala-Ala-NH<sub>2</sub> bien que ce ne soit pas la forme la plus stable.

#### 5.4.2 Ala-Ala-Phe

Outre la forme mixte coude  $\gamma$  - coude  $\beta$  (C<sub>7</sub>-C<sub>10</sub>)  $\gamma_L$ -type I (g+) déjà présentée pour Ala-Ala-Phe, nous avons calculé la forme  $\gamma_L$ -type I (g-), le coude  $\gamma$  triple  $\gamma_L-\gamma_L-\gamma_L$ (g-) et la structure hélice  $3_{10}$  (g-) (Figure 5.14). L'hélice  $3_{10}$  avec une orientation (g+) n'a pas été cal-

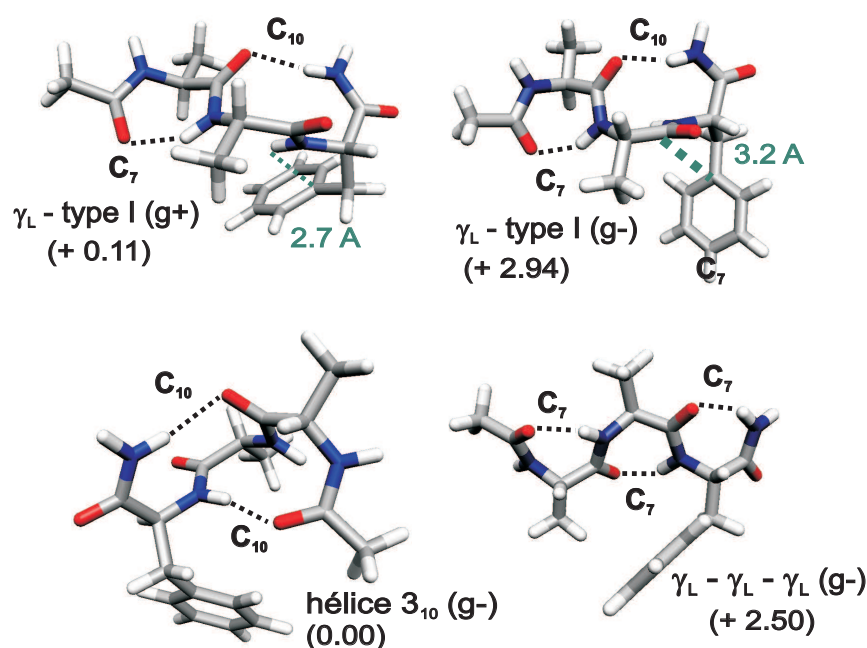


FIG. 5.14 – Structures calculées pour Ac-Ala-Ala-Phe-NH<sub>2</sub> (B3LYP/6-31+G(d)). Les énergies relatives sont données en kcal/mol. Les distances NH<sub>Phe</sub>-C<sub>γ</sub> sont indiquées en Angströms.

culée car elle ne pourrait reproduire la bande IR attribuée à l'interaction NH- $\pi$  (3440 cm<sup>-1</sup>); dans une telle conformation, ce groupement amide est en fait impliqué dans une liaison C<sub>10</sub>. L'hélice 3<sub>10</sub> (g-) et la forme  $\gamma_L$ -type I (g+) sont les formes les plus stables et sont isoénergétiques, elles sont trouvées séparées de 0.11 kcal/mol, ce qui n'est pas significatif. La structure  $\gamma_L$ - $\gamma_L$ - $\gamma_L$  constituée d'une succession de formes locales sur chaque résidu - le coude  $\gamma_L$  est la deuxième forme la plus stable pour le résidu Phe (cf section 3.1 du Chapitre 3) - est moins stable de 2.50 kcal/mol. L'orientation (g-) du phényle de la forme mixte  $\gamma_L$ -type I conduit à une énergie plus élevée que l'orientation (g+), ceci est vraisemblablement dû à l'absence d'interaction NH- $\pi$ . Au regard des données énergétiques et des calculs de fréquences, on peut donc attribuer le conformère Ala-Ala-Phe à une structure mixte  $\gamma_L$ -type I (g+).

On doit à présent se demander pourquoi nous n'observons pas l'hélice 3<sub>10</sub> qui est pourtant isoénergétique de la forme  $\gamma_L$ -type I (g+). Les calculs menés sur Ala-Ala-Ala ainsi que sur Phe-Ala-Ala montrent qu'il s'agit de la forme la plus stable dans les deux cas. Nous avons

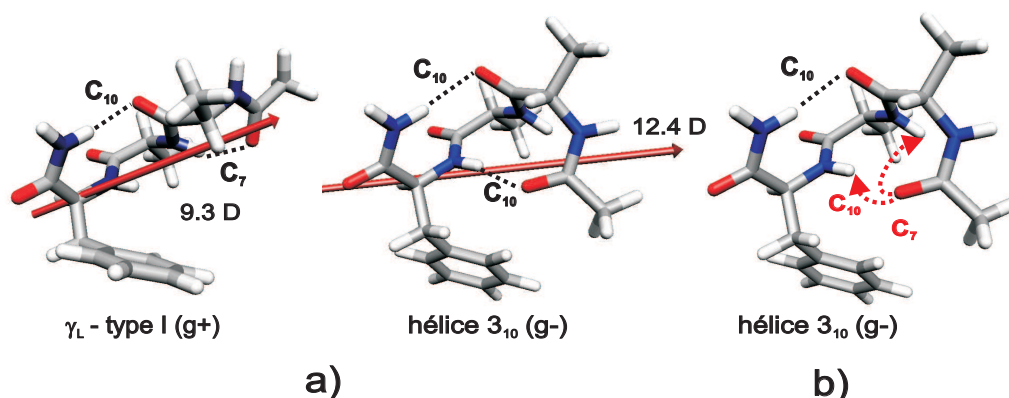


FIG. 5.15 – a) Structures  $\gamma_L$ -type I (g+) et hélice  $3_{10}$  (g-) calculées pour Ac-Ala-Ala-Phe-NH<sub>2</sub> avec la méthode B3LYP/6-31+G(d). Les moments dipolaires totaux, exprimés en Debye, sont représentés par une flèche rouge. b) Les flèches montrent qu’une torsion de l’angle dièdre  $\psi(\text{Ala1})$  permet de passer aisément de la conformation C<sub>10</sub> à la conformation C<sub>7</sub>.

mentionné le cas de Phe-Ala-Ala où l’interaction entre le moment dipolaire total interagissait avec le cycle aromatique dans l’ion pouvant entraîner des difficultés lors de la photoionisation de l’hélice. Or, dans le cas de Ala-Ala-Phe, la valeur calculée du moment dipolaire total est en effet élevée (12.4 D) mais le moment dipolaire ne pointe pas directement vers le cycle comme l’illustre la Figure 5.15a. En conséquence, l’absence de l’hélice  $3_{10}$  ne résulte pas de problème de photoionisation. En revanche, on observe bien la formation d’un coude  $\beta$  sur la sous-partie Ala-Phe du tripeptide - l’hélice  $3_{10}$  est constituée de deux coudes  $\beta$  de type III(I) - mais la présence de l’interaction NH<sub>Phe</sub>- $\pi$  empêche la formation d’une liaison C<sub>10</sub> (CO<sub>Ace</sub>  $\leftarrow$  NH<sub>Phe</sub>) sur la partie Ala-Ala, de sorte que le premier résidu Ala adopte sa préférence conformationnelle locale  $\gamma_L$ . Pourtant, les valeurs de  $\phi$  et  $\psi$  du Tableau C.3 en Annexe (p. 216) révèlent que l’hélice  $3_{10}$  est topologiquement proche de la forme  $\gamma_L$ -type I. Seules les valeurs de  $\psi$  du résidu Ala1 indiquent des différences significatives, montrant qu’une torsion de la liaison C $\alpha$ -N du résidu permet de former soit une interaction NH- $\pi$ , soit une liaison hydrogène C<sub>10</sub>.

Finalement, la formation du conformère mixte  $\gamma_L$ -type I (g+) pour le tripeptide Ac-Ala-Ala-Phe-NH<sub>2</sub> illustre bien le caractère très flexible de la chaîne principale qui permet aux effets de chaîne latérale de favoriser la structure mixte coude  $\gamma$  - coude  $\beta$ . La présence d’une interaction NH- $\pi$  entraîne donc un phénomène de compétition entre la formation d’une liaison C<sub>7</sub> et celle d’une liaison C<sub>10</sub> qui n’est pas compatible avec l’interaction NH- $\pi$ . On peut ce-



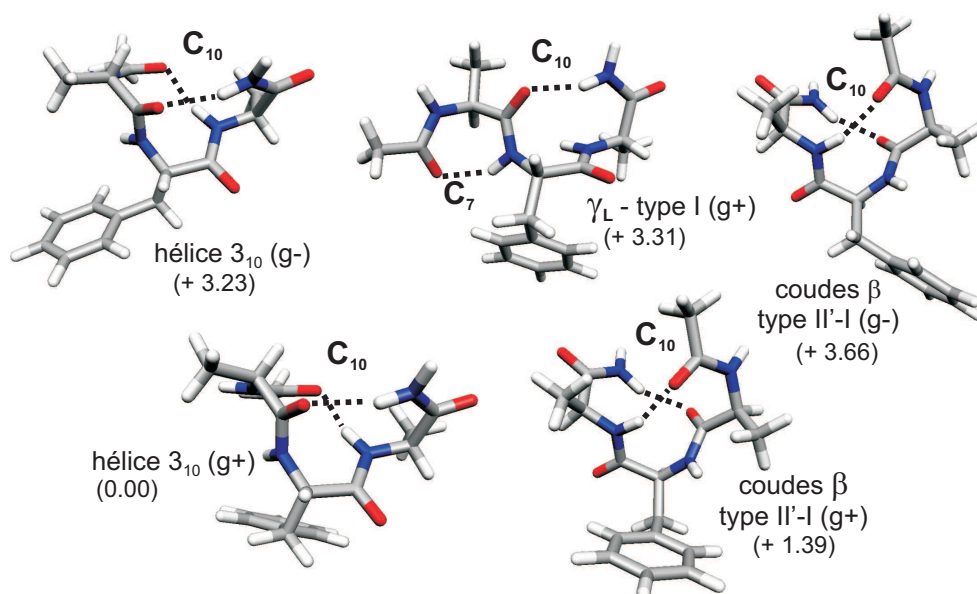


FIG. 5.16 – Structures calculées pour Ac-Ala-Phe-Ala-NH<sub>2</sub> (B3LYP/6-31+G(d)).

pendant supposer qu'en l'absence d'une telle interaction, l'hélice  $3_{10}$  serait très probablement formée puisque d'une part, elle est isoénergétique de la structure  $\gamma_L$  - type I et d'autre part, la structure  $\gamma_L$  - type I est plus haute de 1.85 kcal/mol en l'absence de résidu aromatique Phe comme c'est le cas dans le tripeptide Ac-Ala-Ala-Ala-NH<sub>2</sub>.

### 5.4.3 Ala-Phe-Ala

Outre l'hélice  $3_{10}$  (g+) et la combinaison de coudes  $\beta$  type II'-I (g+) présentées dans la section précédente, nous avons considéré l'orientation (g-) pour le phényle pour chacune des deux conformations. Nous avons également calculé la forme mixte coude  $\gamma_L$  - coude  $\beta$  type I. L'hélice  $3_{10}$  (g+) est le minimum le plus stable parmi les formes calculées, elle est suivie de la combinaison type II'-I (g+) plus haute de 1.39 kcal/mol ; les autres structures ont des énergies supérieures à 3 kcal/mol (Tableau 5.5). Comme pour les autres tripeptides, la présence d'une interaction NH- $\pi$  apporte une stabilisation supplémentaire, c'est le cas des formes avec une orientation (g+) du phényle qui sont de loin les plus stables. En particulier, l'hélice  $3_{10}$  (g-) est déstabilisée de 3.24 kcal/mol par rapport à l'hélice  $3_{10}$  (g+). On peut aussi noter que la combinaison type II'-I est significativement plus élevée que le minimum, contrairement au

système Ala-Ala-Ala où elle est isoénergétique de l'hélice  $3_{10}$ . Les paramètres géométriques regroupés dans le Tableau C.2 en Annexe indiquent que les différences principales entre les deux tripeptides Ala-Ala-Ala et Ala-Phe-Ala résultent de la présence du résidu Phe : seules les valeurs des angles dièdres  $\phi$  et  $\psi$  du résidu en deuxième position, respectivement Ala et Phe pour les systèmes Ala-Ala-Ala et Ala-Phe-Ala, montrent une légère variation.

Ces résultats théoriques corroborent donc l'attribution du conformère A de Ac-Ala-Phe-Ala-NH<sub>2</sub> à une hélice  $3_{10}$ , en particulier, nous pouvons affirmer que l'hélice adopte une orientation (g+) pour le cycle aromatique. Il s'agit de la première observation spectroscopique d'une hélice  $3_{10}$  en phase gazeuse.

#### 5.4.4 Signature de l'hélice $3_{10}$

Les calculs menés sur **tous** les tripeptides, le système Ac-Phe-Gly-Gly-NH<sub>2</sub> compris, montrent une signature IR particulière pour l'hélice  $3_{10}$ . Cette signature réside en fait dans les valeurs **significativement plus élevées** des fréquences d'élongation symétrique et antisymétrique du groupe NH<sub>2</sub> : la fréquence symétrique est de l'ordre de 3405 cm<sup>-1</sup> et la fréquence antisymétrique de l'ordre de 3535 cm<sup>-1</sup> (Tableau 5.6). Le conformère A de Ac-Ala-Phe-Ala-NH<sub>2</sub> que nous avons attribué à la structure hélice  $3_{10}$  (g+) présente ainsi cette signature. Les liaisons C<sub>10</sub> des coudes  $\beta$  observés expérimentalement pour les deux séries de dipeptides Ac-Phe-Xxx-NH<sub>2</sub> et Ac-Xxx-Phe-NH<sub>2</sub>, ainsi que pour le tripeptide Ac-Ala-Ala-Phe-NH<sub>2</sub>, correspondent, elles, à des fréquences  $\nu_{sym}$  et  $\nu_{anti}$  proches du couple 3390/3520 cm<sup>-1</sup>. Tous les calculs réalisés sur les coudes  $\beta$  comprenant des types différents montrent aussi cette tendance, excepté pour le coude type VI de Ac-Phe-Pro-NH<sub>2</sub> dont la fréquence très basse (< 3350 cm<sup>-1</sup>) est due à la configuration *cis* de la liaison peptidique centrale du coude.

L'hélice  $3_{10}$  est une variante périodique de deux coudes  $\beta$  type I successifs. Néanmoins, ses paramètres géométriques sont tout de même sensiblement différents. En effet, comme nous l'avons mentionné dans la section sur le tripeptide Ac-Phe-Gly-Gly-NH<sub>2</sub>, la combinaison type I-I et la conformation de l'hélice  $3_{10}$  idéale convergent vers la même géométrie finale type III-I. Les calculs trouvent donc que l'hélice  $3_{10}$  est légèrement déformée dans la partie C-terminale et converge vers un coude  $\beta$  type I. La combinaison type II'-I présente elle aussi un coude type I

Tripeptides	Angles dièdres (°)			Fréquences (cm <sup>-1</sup> )	
	$\phi/\psi$ (i)	$\phi/\psi$ (i+1)	$\phi/\psi$ (i+2)	NH <sub>2</sub> anti.	NH <sub>2</sub> sym.
Phe-Gly-Gly	-74/-10	-69/-10	-97/6	3538	3413
Phe-Ala-Ala	-75/-9	-68/-13	-97/3	3534	3411
Ala-Phe-Ala	-68/-21	-74/-1	-100/3	3535	3410
Ala-Ala-Phe	-69/-19	-68/-12	-95/1	3531	3404

TAB. 5.6 – Paramètres géométriques de l'hélice  $3_{10}$  calculés pour les tripeptides protégés Ac-...-NH<sub>2</sub> : Phe-Gly-Gly(g+), Phe-Ala-Ala(g+), Ala-Phe-Ala(g+) et Ala-Ala-Phe(g-) et fréquences IR théoriques (B3LYP/6-31+G(d)) des modes d'élongation NH<sub>2</sub> symétrique et antisymétrique, corrigées avec le facteur 0.96 pour tenir compte de l'anharmonicité et du défaut de la base.

sur la partie C-terminale ; cependant, les calculs ne trouvent pas de décalage en fréquences vers le bleu comme c'est le cas de l'hélice mais indiquent des fréquences plus basses de l'ordre de 3385/3525 cm<sup>-1</sup>. Par conséquent, le couple de fréquences 3405/3535 cm<sup>-1</sup> est caractéristique de la structure périodique hélice  $3_{10}$ .

## 5.5 Conclusion du Chapitre 5

Nous avons étudié deux types de tripeptides dans ce chapitre, le système Ac-Phe-Gly-Gly-NH<sub>2</sub> et trois tripeptides dérivés du système Ac-Ala-Ala-Ala-NH<sub>2</sub>, pour lesquels le résidu aromatique Phe est successivement placé dans les trois positions de la séquence, Ac-Phe-Ala-Ala-NH<sub>2</sub>, Ac-Ala-Phe-Ala-NH<sub>2</sub> et Ac-Ala-Ala-Phe-NH<sub>2</sub>.

Nous avons identifié au moins cinq conformères pour le tripeptide Ac-Phe-Gly-Gly-NH<sub>2</sub>, dont quatre sont des structures relativement compactes et repliées présentant des liaisons hydrogène multiples. Ces liaisons peuvent correspondre à des coudes  $\gamma$  mais également à des formes globulaires dont les liaisons hydrogène ne sont pas les liaisons classiques C<sub>7</sub>, C<sub>10</sub>, ou encore C<sub>5</sub> des structures remarquables de la biologie. Un des conformères de Ac-Phe-Gly-Gly-NH<sub>2</sub> présente à l'inverse une signature IR indiquant des liaisons hydrogène plus faibles, typiques des coudes  $\beta$ . L'étude théorique que nous avons menée en parallèle nous a permis d'attribuer ce conformère (noté B) à la structure type II-I'(a), composée de deux coudes  $\beta$  successifs type II et type I', qui s'avère être la forme la plus stable parmi toutes les structures

calculées. L'observation de plusieurs conformères témoigne de la complexité du paysage conformationnel de Ac-Phe-Gly-Gly-NH<sub>2</sub>. Une telle complexité est liée non seulement à la longueur du système mais aussi à l'absence de chaîne latérale pour le résidu Gly et à sa petite taille, lui permettant d'accéder à un espace conformationnel très large par rapport aux 19 autres résidus naturels qui possèdent une chaîne latérale. La formation d'un coude  $\beta$  type II est ainsi presque uniquement observée pour les résidus Gly placé en deuxième position du coude  $\beta$  dans les protéines. Notre attribution pour le conformère B est donc conforme aux résultats observés en biologie par diffraction des rayons X. Néanmoins, nos résultats, expérimentaux et théoriques, ne nous ont pas permis d'observer de structures hélicoïdales.

Nous avons ensuite étudié des tripeptides contenant le résidu Ala afin d'observer la structure en hélice  $3_{10}$  qui peut en principe être formée. Le choix du résidu s'est porté sur Ala car il présente la plus grande propension à former des hélices en solution. L'utilisation de sources lasers UV impose la présence d'un chromophore aromatique. Nous avons donc placé le résidu Phe dans les trois positions possibles dans le système de référence Ac-Ala-Ala-Ala-NH<sub>2</sub>. Pour chacun des trois tripeptides, un seul conformère principal, suffisamment peuplé, a été identifié. Cet unique conformère démontre bien que la présence de chaîne latérale dans la séquence réduit sensiblement le nombre de conformations accessibles à la chaîne peptidique. Les études expérimentales ont été accompagnées d'une contrepartie théorique qui s'appuie sur les résultats de l'étude théorique menée sur Ac-Ala-Ala-Ala-NH<sub>2</sub>. En accord avec les travaux réalisés en solution et avec les calculs de chimie quantique effectués sur divers systèmes en phase gazeuse, nous avons trouvé que l'hélice  $3_{10}$  était le minimum de plus basse énergie parmi les formes considérées. Toutefois, dans les tripeptides contenant le résidu Phe, suivant la position de ce résidu dans la séquence, d'autres structures se sont révélées capables de concurrencer l'hélice  $3_{10}$ . Nous avons ainsi pu montrer que les tripeptides adoptaient tous des conformations différentes, à savoir la structure  $\beta_L(a)-\gamma_L-\gamma_L$  pour Ac-Phe-Ala-Ala-NH<sub>2</sub>, correspondant à la juxtaposition des préférences conformationnelles locales sur chaque résidu, la structure coude  $\gamma_L$  - coude  $\beta$  pour Ac-Ala-Ala-Phe-NH<sub>2</sub> qui est une structure mixte composée d'une forme locale coude  $\gamma$  et d'une structure secondaire coude  $\beta$ , et la structure périodique hélice  $3_{10}$  pour Ac-Ala-Phe-Ala-NH<sub>2</sub>. Ces travaux constituent la première observation expérimentale en

phase gazeuse d'une hélice  $3_{10}$ .

Nous avons donc montré que l'observation de structures différentes pour les tripeptides résultaient des effets de chaîne latérale qui pouvaient mener à un phénomène de compétition, ou plutôt d'un équilibre, entre la formation de liaisons hydrogène et celle d'interactions plus faibles telles que les interactions NH- $\pi$ , pouvant favoriser une structure au dépend d'une autre. Cet effet est également retrouvé par les calculs qui montrent que les formes les plus stables présentent une interaction NH- $\pi$ .

**Références bibliographiques du Chapitre 5**

- [1] R. Improta, V. Barone, K. N. Kudin and G. E. Scuseria, *J. Am. Chem. Soc.* **123**, 3311 (2001).
- [2] M. Elstner, K. J. Jalkanen, M. Knapp-Mohammady, T. Frauenheim and S. Suhai, *Chem. Phys.* **256**, 15 (2000), et réf. citées.
- [3] I. A. Topol, S. K. Burt, E. Deretey, T.-H. Tang, A. Perczel, A. Rashin and I. G. Csizmadia, *J. Am. Chem. Soc.* **123**, 6054 (2001).
- [4] R. Wieczorek and J. J. Dannenberg, *J. Am. Chem. Soc.* **125**, 14065 (2003), et réf. citées.
- [5] R. Wieczorek and J. J. Dannenberg, *J. Am. Chem. Soc.* **126**, 14198 (2004).
- [6] E. G. Robertson, M. R. Hockridge, P. D. Jelfs and J. P. Simons, *Phys. Chem. Chem. Phys.* **3**, 786 (2001).
- [7] R. F. Frey, J. Coffin, S. Q. Newton, M. Ramek, V. K. W. Cheng, F. A. Momany and L. Schafer, *J. Am. Chem. Soc.* **114**, 5369 (1992).
- [8] R. Vargas, J. Garza, B. P. Hay and D. Dixon, *J. Phys. Chem. A* **106**, 3213 (2002).
- [9] K. J. Jalkanen, M. Elstner and S. Suhai, *J. Mol. Struct. (Theochem)* **675**, 61 (2004), et réf. citées.
- [10] A. Perczel, I. Jákli and I. G. Csizmadia, *Chem. Eur. J.* **9**, 5332 (2004).
- [11] Y.-D. Wu and Y.-L. Zhao, *J. Am. Chem. Soc.* **123**, 5313 (2001).
- [12] E. G. Robertson and J. P. Simons, *Phys. Chem. Chem. Phys.* **3**, 1 (2001).
- [13] G. Rose, L. M. Gierasch and J. A. Smith, *Adv. Prot. Chem.* **37**, 1 (1985).
- [14] P. Dugourd, R. Antoine, G. Breaux, M. Broyer and M. F. Jarrold, *J. Am. Chem. Soc.* **127**, 4675 (2005), et réf. citées.



# Conclusion générale

L'objectif de cette thèse était de caractériser les propriétés intrinsèques de repliement de la chaîne peptidique. Cette thématique nous a conduit à déterminer la structure qu'adoptent spontanément de petits peptides protégés (Ac-...-NH<sub>2</sub>) en phase gazeuse. L'observation de certaines formes typiques de la biologie nous a permis de montrer que ces systèmes pouvaient reproduire le comportement de petits segments de chaîne peptidique. Nous avons notamment observé des structures secondaires de la biologie : des coudes  $\beta$  et une hélice  $3_{10}$ . Leur formation en phase gazeuse est spontanée. Ces structures figurent toujours parmi les plus stables. Cette étude constitue la première observation spectroscopique de structures secondaires en phase gazeuse.

Les outils expérimentaux et théoriques utilisés lors de ces travaux ont été particulièrement efficaces pour caractériser la structure de petites molécules biologiques. L'emploi des techniques de spectroscopie de double résonance IR-UV a permis de mesurer le spectre IR d'une espèce sélectionnée par son spectre UV. Une telle méthode peut donc être très sélective, à la fois en terme de masse par couplage à la spectrométrie de masse, et en terme de conformation de l'espèce étudiée.

La simplification spectrale apportée par le refroidissement en détente supersonique combinée à la finesse spectrale des sources laser permet alors d'obtenir des informations très précises. Les modes de vibration amide A, amide I et II donnent ainsi accès aux réseaux de liaisons hydrogène, et donc à la structure des molécules.

Les résultats issus de notre étude préliminaire sur le petit système Ac-Phe-NH<sub>2</sub> ont montré que seul un petit nombre de conformères était significativement peuplé dans le jet malgré la



complexité du paysage conformationnel <sup>1</sup>. Nous avons ainsi vu que deux types de structures étaient particulièrement stables pour un résidu aromatique : une forme étendue  $\beta_L$ , élément de base du brin  $\beta$ , et une forme repliée équivalente au motif central d'un coude  $\gamma$ . La stabilité remarquable de la structure  $\beta_L$  (qui n'est pas observée dans les résidus non aromatiques) résulte en fait de la présence d'une interaction NH- $\pi$  entre la chaîne peptidique et le cycle aromatique. Ces deux structures illustrent les préférences conformationnelles locales de la chaîne peptidique autour du résidu Phe.

L'étude détaillée de ce système nous a permis d'élaborer un modèle décrivant le couplage vibrationnel entre les oscillateurs NH du groupement amino. Le choix d'un groupement NH<sub>2</sub> plutôt qu'un groupement NH-Me en position C-terminale constitue en ce sens une « astuce » spectroscopique puisque le couplage vibrationnel apporte une information structurale supplémentaire. Il permet en particulier de localiser une des liaisons hydrogène en indiquant si le groupement amino est impliqué dans le réseau de liaisons hydrogène : le décalage vers le rouge des fréquences d'élongation NH<sub>2</sub> symétrique et antisymétrique ainsi que l'écart entre les deux modes sont ainsi des indicateurs de la force de la liaison hydrogène sur le groupement NH<sub>2</sub>.

Nous avons ensuite cherché à caractériser l'effet de la longueur de la chaîne peptidique en étudiant deux séries de dipeptides, Ac-Phe-Xxx-NH<sub>2</sub> et Ac-Xxx-Phe-NH<sub>2</sub>, pour lesquels nous avons fait varier de façon systématique la nature du résidu hydrophobe Xxx (Xxx = Gly, Ala, Val, Pro). Nous avons alors montré que deux grandes classes de structures appartenant à des bassins conformationnels différents mais d'énergie proche étaient en compétition. La première classe correspond à la juxtaposition des préférences conformationnelles locales autour de chaque résidu : les structures  $\beta_L$ - $\gamma$  sont favorisées lorsque les résidus au sein de la chaîne interagissent peu <sup>2</sup>, tandis que des effets coopératifs résultant d'interactions entre résidus mènent à la formation de structures de type  $\gamma$ - $\gamma$  <sup>3</sup>. La deuxième classe de structures est constituée de coudes  $\beta$  <sup>4</sup>, omniprésents pour toutes les chaînes étudiées, sauf deux. Ceci souligne donc que la conformation de la chaîne peptidique dépend peu de la nature du résidu

<sup>1</sup>Chin *et al.*, J. Phys. Chem. A **109**, 5281 (2005); Chass *et al.*, J. Phys. Chem. A **109**, 5289 (2005); Chin *et al.*, Phys. Chem. Chem. Phys. **6**, 2700 (2004).

<sup>2</sup>Chin *et al.*, J. Chem. Phys. **123**, 084301 (2005).

<sup>3</sup>Chin *et al.*, J. Chem. Phys. **122**, 054317 (2005).

<sup>4</sup>Chin *et al.*, J. Am. Chem. Soc. **127**, 707 (2005).

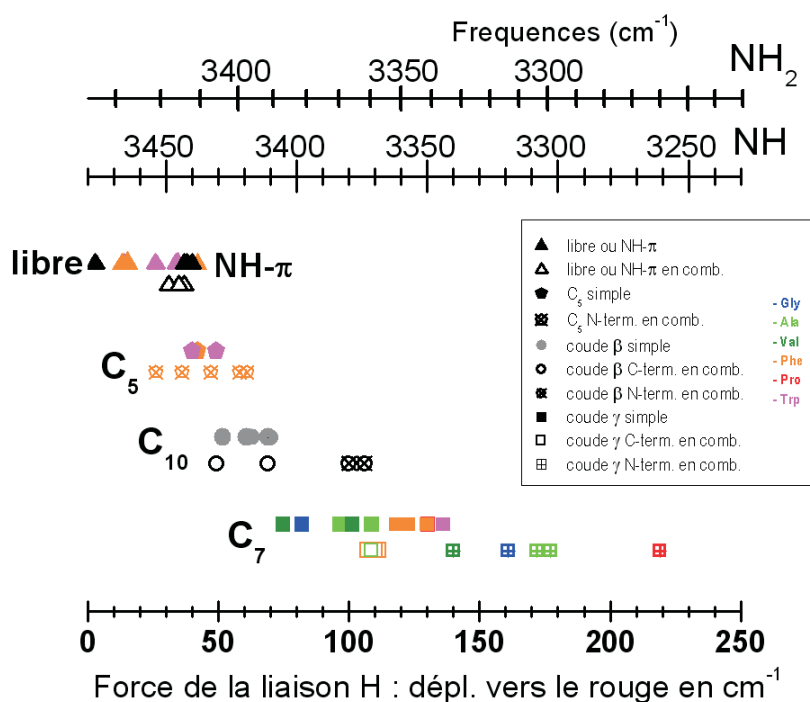


FIG. 5.17 – Échelle des déplacements spectraux de la vibration d'élongation des groupes  $\text{NH}$  et  $\text{NH}_2$  dans les chaînes peptidiques en fonction du type de liaison hydrogène dans lesquels ils sont engagés. *Bas de la figure* : Force de la liaison estimée à partir du déplacement vers le rouge pour un groupement  $\text{NH}$  peptidique (en supposant une fréquence de vibration à  $3480 \text{ cm}^{-1}$  pour un oscillateur libre), et à partir de l'écart entre les fréquences d'élongation symétrique et antisymétrique pour un groupement  $\text{NH}_2$  terminal (analysé dans le cadre du modèle de couplage vibrationnel décrit dans l'Annexe A). Sur la figure sont portées les différentes fréquences mesurées en fonction du type de liaison H auquel elles sont attribuées. On a fait la différence entre des structures simples, en faible interaction avec leurs voisins, et celles observées en combinaison avec une autre structure (par exemple,  $\beta\text{-L-}\gamma$ ,  $\gamma\text{-}\gamma$  ou double coude  $\beta$ ). Les coudes  $\beta$  et  $\gamma$ , pour lesquels des effets coopératifs sont observés, sont distingués par leur position N- ou C-terminale dans la chaîne. *Haut de la figure* : Échelles de fréquences indiquant les fréquences de vibration d'élongation  $\text{NH}$  et  $\text{NH}_2$  antisymétrique correspondantes. Cette compilation regroupe l'ensemble des systèmes étudiés dans le présent travail ainsi que les travaux de M. Gerhards et collaborateurs (Ac-Phe-NHMe et Ac-Val-(Me)Tyr-NHMe) et T.S. Zwier et collaborateurs (Ac-Trp-NHMe et Ac-Trp-NH2).

hydrophobe Xxx.

Ce travail s'est poursuivi par l'étude de tripeptides <sup>5</sup>, systèmes dans lesquels un précurseur de l'hélice  $3_{10}$  a été observé <sup>6</sup>.

La synergie entre méthodes expérimentale et théorique a permis de caractériser précisément les conformères observés, permettant une compréhension très fine des interactions responsables de la conformation de la chaîne peptidique. Grâce à la précision et à la sélectivité apportées par les techniques de double résonance IR-UV, nous avons identifié la signature spectrale de chacune des deux familles. Nous avons ainsi établi une échelle de déplacements spectraux en fonction du type de liaison hydrogène adopté par la chaîne (Figure 5.17) : liaisons hydrogène C<sub>7</sub>, C<sub>10</sub> ou bien interactions C<sub>5</sub> et NH- $\pi$ , plus faibles. Nous avons également souligné l'importance des effets de chaîne latérale qui peuvent influencer la compétition entre les différentes conformations. Ces effets peuvent se manifester par l'établissement d'interactions NH- $\pi$ , par des interactions de contact entre le squelette peptidique et les chaînes latérales ou encore par des interactions entre chaînes latérales.

Enfin, l'étude théorique menée sur nos systèmes a permis de valider l'emploi de la fonctionnelle B3LYP en DFT pour les calculs de fréquences. En revanche, elle nous a aussi permis de montrer que cette approximation n'était pas adaptée aux calculs d'énergie relative mais que la prise en compte des effets de corrélation et de dispersion (méthodes coupled-cluster, théorie des perturbations, etc.) était indispensable pour rendre compte du paysage énergétique des systèmes étudiés.

## Perspectives

Les résultats très concluants obtenus sur nos petites chaînes de peptides nous encouragent à étendre nos travaux à des systèmes plus grands, et notamment à la recherche de l'hélice  $\alpha$ . L'allongement par étape de la chaîne peptidique pourrait, en premier lieu, nous permettre d'observer des effets coopératifs au sein de la chaîne peptidique. L'étude de peptides à cinq résidus adoptant vraisemblablement une conformation en hélice  $3_{10}$  permettrait de caractéri-

---

<sup>5</sup>Chin *et al.*, J. Am. Chem. Soc. **127**, 1388 (2005).

<sup>6</sup>Chin *et al.*, J. Am. Chem. Soc. **127**, 11900 (2005).

ser ces effets puisqu'un tel système est la plus petite chaîne pouvant présenter deux chaînes parallèles de liaisons hydrogène coopératives.

Les hélices  $3_{10}$ , fréquemment retrouvées aux extrémités des hélices  $\alpha$ , sont parfois considérées comme des intermédiaires de celles-ci. Certains travaux théoriques estiment ainsi que l'hélice  $\alpha$  est la structure la plus stable à partir de huit résidus<sup>7</sup>. L'étude de peptides à sept et neuf résidus est donc envisagée afin de déterminer la taille critique pour observer la transition entre hélice  $3_{10}$  et hélice  $\alpha$ .

De telles investigations pourraient néanmoins entraîner quelques difficultés. Les méthodes de chimie quantique traditionnelles seraient en effet difficilement applicables à des systèmes de cette taille en raison de temps de calcul trop élevés. Afin de réduire ces temps de calcul, on devra alors recourir à des méthodes de chimie quantique faisant appel à d'autres approximations telles que les méthodes RI-DFT. De plus, des difficultés expérimentales peuvent aussi survenir. Le nombre croissant de liaisons peptidiques pourrait poser le problème de recouvrements spectraux dans les spectres IR. Par ailleurs, l'alignement des dipôles de chaque liaison peptidique dans le cas d'une structure hélicoïdale pourrait aussi entraîner des phénomènes de couplage entre les modes de vibration, pouvant rendre les spectres plus difficiles à interpréter. En outre, nous avons pu voir, dans le cas de l'hélice  $3_{10}$ , que l'interaction entre le dipôle total et la charge portée par le cycle aromatique dans l'ion pouvait conduire à des difficultés de détection par photoionisation. Or, la présence de chromophore UV est indispensable dans nos expériences. Une technique élégante pour contourner ce problème serait de recourir à la technique de transfert d'électron de Rydberg couplée à des techniques de spectroscopie IR. Cette méthode est en effet particulièrement adaptée aux molécules à fort moment dipolaire et pourrait donc permettre de caractériser la signature IR des hélices.

L'étude des hydrates de nos systèmes est également envisageable et permettrait d'identifier les sites d'hydratation et de caractériser les changements induits par les molécules d'eau, solvant biologique naturel, dans la conformation de la chaîne peptidique. En particulier, l'étude de l'effet de l'hydratation sur les structures secondaires, telles que les coudes  $\beta$ , permettrait de tester la robustesse des liaisons hydrogène sous jacentes.

---

<sup>7</sup>Jalkanen *et al.*, J. Mol. Struct. (Theochem) **675**, 61 (2004), et réf. citées.

Enfin, il serait très utile d'appliquer la technique d'isomérisation ou « hole-filling spectroscopy<sup>8</sup> » sur nos systèmes, afin d'évaluer les hauteurs de barrière séparant les différents conformères, et par comparaison, de déterminer leur énergie relative, donnée jusqu'à présent inaccessible expérimentalement.

---

<sup>8</sup>Dian *et al*, Science **303**, 11969 (2004).

## Annexe A

# Modèle de couplage vibrationnel

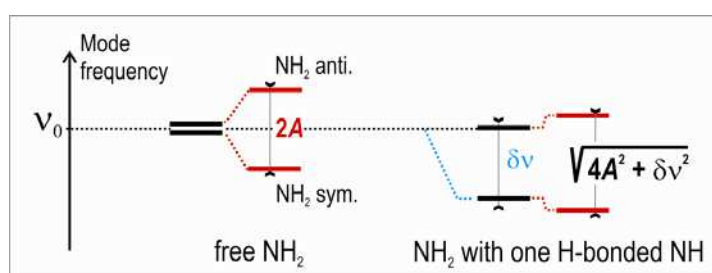
Nous avons établi un modèle phénoménologique décrivant les modes normaux d'élongation d'un groupement NH<sub>2</sub>, en particulier impliqué dans une liaison hydrogène. Cette annexe en expose le principe <sup>1</sup>.

---

<sup>1</sup>tiré de *Gas-Phase Formation of a  $3_{10}$ -Helix in a Three-Residue Peptide Chain : Role of Side Chain-Backbone Interactions as Evidenced by IR-UV Double Resonance Experiments*, W. Chin, F. Piuze, J.-P. Dognon, I. Dimicoli, B. Tardivel and Michel Mons, *J. Am. Chem. Soc.* (2005), sous presse

## Phenomenological model describing the NH stretch normal modes of a H-bonded NH<sub>2</sub> group.

In the absence of an external perturbation, the vibrational stretch motion of the NH<sub>2</sub> group can be viewed, in a zero approximation, as identical local NH stretches (mode of frequency  $\nu_0$ ), undergoing a vibrational coupling due to the fact that they share the same nitrogen atom (Figure 1).<sup>1</sup> This coupling (red dots; coupling constant  $A$ ) splits these local modes into a doublet of normal modes (splitting  $2A$ , according to first order perturbation theory), composed of a symmetric and an antisymmetric component; the latter being the higher frequency mode.



**Figure 1** : Phenomenological model accounting for the vibrational coupling between zero approximation NH stretching local modes (black levels) of a NH<sub>2</sub> group, in the absence (left) or the presence (right) of an external perturbation (see text). The normal modes resulting from the vibrational coupling (red dots) are shown as red levels. The effect of the external perturbation on one of the local modes (red-shift in the case of H-bonding) is indicated by blue dots.

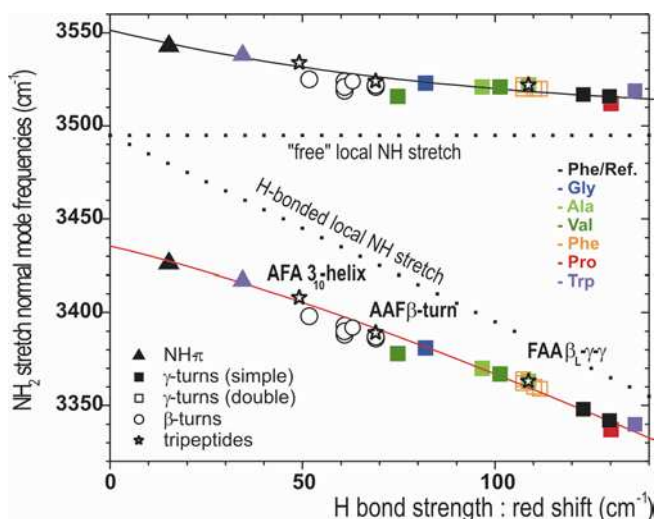
The normal modes resulting from the vibrational coupling (red dots) are shown as red levels. The effect of the external perturbation on one of the local modes (red-shift in the case of H-bonding) is indicated by blue dots.

When one of the NH moieties is perturbed, for instance by an H-bond, the zero approximation local modes are no longer degenerate: one of them is red shifted (red-shift  $\delta\nu$ ; relative blue dashes). However the vibrational coupling, which caused the  $2A$  splitting in the free amide regime, is still acting on this two-level system. Its effect is characterised by the same coupling constant  $A$  since the physical reason for the vibrational coupling remains unchanged. The resulting normal modes, eigenstates of the Hamiltonian, are split by the quantity  $\sqrt{\delta\nu^2 + 4A^2}$ , which is larger than both  $2A$  and  $\delta\nu$ .<sup>1-3</sup> Their frequency (figure 2) as well as their physical nature evolves when  $\delta\nu$  increases, with two remarkable limiting cases:

- When  $\delta\nu / A \gg 1$ , the effect of vibrational coupling is negligible: the two modes are essentially local: one being very red-shifted because of the H-bond perturbation, the other one converging towards the frequency of a free NH stretch local mode.
- When  $\delta\nu / A \ll 1$ , the normal modes are a mixing of local NH stretches which strongly resemble the antisymmetric / symmetric combinations of the free scheme. For this reason the

high and low frequency component of the doublet are often still referred to as “antisymmetric” and “symmetric” respectively, despite the NH<sub>2</sub> system is no longer symmetric.

More generally, starting from the free NH<sub>2</sub> regime, the antisymmetric and symmetric normal modes correlate in the strong H-bond scheme to free and bonded local modes respectively. From Figure 2, one observes that when  $\delta\nu$  increases, the frequency of the high frequency normal mode decreases from  $(\nu_0+A)$  to  $\nu_0$ ; the frequency of the other mode also decreases and the splitting between them increases. This observation suggests therefore using the frequency of the blue component of the NH<sub>2</sub> stretch doublet as an indicator of H-bonding of the NH<sub>2</sub> moiety: frequencies close to  $\nu_0$  will suggest strong H-bonds, whereas frequencies closer to  $(\nu_0+A)$  will suggest weak H-bonds. In addition measuring the splitting ( $\sqrt{\delta\nu^2 + 4A^2}$ ) will provide an estimate of the red-shift  $\delta\nu$  and therefore of the H-bond strength.



**Figure 2:** Evolution of the vibrational frequencies of the normal modes as a function of the H-bond strength, expressed in terms of red-shift (lines). The model parameters ( $A = 116 \text{ cm}^{-1}$ ,  $\nu_0 = 3493 \text{ cm}^{-1}$ ) have been adjusted to fit spectroscopic data of the three conformers of Ac-Phe-NH<sub>2</sub>.<sup>3</sup> Also given are the IR experimental NH<sub>2</sub> stretch frequencies measured for the several conformers observed for two series of dipeptides, Ac-Phe-Xxx-NH<sub>2</sub> and Ac-Xxx-

Phe-NH<sub>2</sub>, Xxx = Gly, Ala, Val and Pro,<sup>3-6</sup> for Ac-Phe-Xxx-NH<sub>2</sub><sup>7</sup> as well as for the three protected tripeptides studied in the present work. For each conformer observed, the IR experimental frequencies of both NH<sub>2</sub> components are plotted as a function of the H bond-induced red-shift  $\delta\nu$ , deduced from the experimental splitting.

Such a phenomenological model (Fig. 2, solid lines), adjusted on the three conformers observed for the Ac-Phe-NH<sub>2</sub> molecule (black filled symbols), fits well all the NH<sub>2</sub> stretch doublet frequencies measured so far on protected dipeptides possessing a -CO-NH<sub>2</sub> C-terminus amide group, as well as on the three present tripeptides (stars). The plot clearly provides evidence for the weaker effect of C<sub>10</sub> interactions ( $\beta$ -turns, hollow circles), compared to C<sub>7</sub> H-bonds (squares). It also illustrates the unusual weakness of this C<sub>10</sub> interaction in the case of main conformer of Ac-Ala-Phe-Ala-NH<sub>2</sub> (labelled as AFA), which is assigned to 3<sub>10</sub> helix conformation.



## References

- (1) Herzberg, G. In *Molecular Spectra and Molecular Structure, Vol II, Chapter II.5*; Van Nostrand Company, Inc: New York, 1945; pp 216-217 (5<sup>th</sup> printing).
- (2) Nakanaga, T.; Ito, F. *J. Phys. Chem. A* **1999**, *103*, 5440.
- (3) Chin, W.; Mons, M.; Dognon, J.-P.; Piuzzi, F.; Tardivel, B.; Dimicoli, I. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2700.
- (4) Chin, W.; Dognon, J. P.; Piuzzi, F.; Tardivel, B.; Dimicoli, I.; Mons, M. *Journal of the American Chemical Society* **2005**, *127*, 707.
- (5) Chin, W.; Dognon, J.-P.; Canuel, C.; Piuzzi, F.; Dimicoli, I.; Mons, M.; Compagnon, I.; von Helden, G.; Meijer, G. *J. Chem. Phys.* **2005**, *122*, 054317.
- (6) Chin, W.; Dognon, J.-P.; Dimicoli, I.; Piuzzi, F.; Mons, M. *J. Chem. Phys.* **2005**, *accepted*.
- (7) Dian, B. C.; Longarte, A.; Mercier, S.; Evans, D. A.; Wales, D. J.; Zwiernik, T. S. *J. Chem. Phys.* **2002**, *117*, 10688.

## Annexe B

### Article

Characterization of the Conformational Probability of N-Acetyl-Phenylalanyl-NH<sub>2</sub> by RHF, DFT, and MP2 Computation and AIM Analyses, Confirmed by Jet-Cooled Infrared Data.

## Characterization of the Conformational Probability of N-Acetyl-Phenylalanyl-NH<sub>2</sub> by RHF, DFT, and MP2 Computation and AIM Analyses, Confirmed by Jet-Cooled Infrared Data

Gregory A. Chass,<sup>\*†‡.§.⊥.☆.◇</sup> Reinard S. Mirasol,<sup>†,‡,||</sup> David H. Setiadi,<sup>†,‡</sup> Ting-Hua Tang,<sup>†,‡</sup> Wutharath Chin,<sup>§</sup> Michel Mons,<sup>§</sup> Iliana Dimicoli,<sup>§</sup> Jean-Pierre Dognon,<sup>§</sup> Bela Viskolcz,<sup>||</sup> Sándor Lovas,<sup>⊥</sup> Botond Penke,<sup>#,★</sup> and Imre G. Csizmadia<sup>†,‡,#</sup>

Global Institute of COmputational Molecular and Materials Science (GIOCOMMS), 2-158 Major Street, Toronto, Ontario M5S 2K2, Canada, Lash Miller Chemical Laboratories, Department of Chemistry, University of Toronto, 80 St. George Street, Toronto, Ontario M5G 2E8, Canada, Laboratoire Francis Perrin (URA CNRS 2453), Service des Photons, Atomes et Molécules, and CEA / DSM / DRECAM, Centre d'Etudes de Saclay, Bât. 522, 91191 Gif-sur-Yvette Cedex, France, Department of Chemistry, Juhasz Gyula Teaching College, University of Szeged, Szeged 6725, Hungary, Department of Biomedical Sciences, School of Medicine, Creighton University, 2500 California Plaza, Omaha, Nebraska 68178, Department of Medical Chemistry, University of Szeged, Dóm tér 8, H-6720 Szeged, Hungary, Protein Chemistry Research Group, Hungarian Academy of Sciences, University of Szeged, Dóm tér 8, H-6720 Szeged, Hungary, Instituto de Biocomputación y Física de Sistemas Complejos (BIFI), Corona de Aragón 42, 50009 Zaragoza, Spain, and Zymeworks Inc., Suite 201, 1401 West Broadway, Vancouver, British Columbia V6H 1H6, Canada

Received: November 15, 2004

Computational and experimental determinations were carried out in parallel on the conformational probability of N-Acetyl-Phenylalanine-NH<sub>2</sub> (NAPA). Ab initio computations were completed at the BLYP/6-311G(df,p), B3LYP/6-31G(d), B3LYP/6-31G(d,p), and B3LYP/6-31+G(d) levels of theory, labeled L/61fp, B/6, B/6p, and B/6+, respectively. Three experimentally identified conformers were compared with theoretical data, confirming their identities as the  $\beta_L^{anti}$ ,  $\gamma_L^{gauche^+}$ , and  $\gamma_L^{gauche^-}$  (BACKBONE<sup>SIDECHAIN</sup>) conformers. Evidence comes from matching experimental and theoretical data for all three constituent N–H stretches of NAPA, with a  $\Delta_{\text{Experimental-Theoretical}} = \sim 1-3 \text{ cm}^{-1}$ ,  $\sim 0-5 \text{ cm}^{-1}$ , and  $\sim 1-6 \text{ cm}^{-1}$ , at the L/61fp and B/6+ levels, respectively. Corrected-ZPE relative energies were computed to be 0.14, 0.00, 0.26 and 0.00, 0.67, 0.57 (kcal\*mol<sup>-1</sup>) for the  $\beta_L^{anti}$ ,  $\gamma_L^{gauche^+}$ , and  $\gamma_L^{gauche^-}$  conformers, respectively, at the L/61fp and B/6+ levels, respectively. The MP2/6-31+G(d) level of theory was subsequently found to give similar relative energies. Characterization of the intramolecular interactions responsible for red and blue shifting of the N–H stretches showed the existence of the following intramolecular interactions: C=O<sub>[i]</sub>- -HN<sub>[i]</sub>, (Ar<sub>[i]</sub>)-C<sub>γ</sub>- -HN<sub>[i+1]</sub>, (Ar<sub>[i]</sub>)-C<sub>δ</sub>-H- -O=C<sub>[i-1]</sub> for  $\beta_L^{anti}$ ; C=O<sub>[i-1]</sub>- -HN<sub>[i+1]</sub>, (Ar<sub>[i]</sub>)-C<sub>γ</sub>- -HN<sub>[i+1]</sub>, (Ar<sub>[i]</sub>)-C-H- -O=C<sub>[i]</sub> for  $\gamma_L^{gauche^+}$ ; and C=O<sub>[i-1]</sub>- -HN<sub>[i+1]</sub> for  $\gamma_L^{gauche^-}$ . Each of these interactions were further investigated and subsequently characterized by orbital population and Atoms-In-Molecules (AIM) analyses, with the identity of overlap and bond critical points (BCP) serving as ‘scoring criteria’, respectively. Experimental and theoretical carbonyl stretches were also compared and showed good agreement, adding further strength to the synergy between experiment and theory.

### 1. Introduction

Protein folding has become a central challenge to many disciplines, including the biomedical sciences, biology, biochemistry, chemistry, molecular physics, mathematics, informatics, and computer science.

The underlying mechanisms of protein folding may first be tackled through the characterization of the folding of short-peptides, with the latter as a prelude to the former.<sup>1-3</sup> Two

approaches exist, specifically a holistic global, or macromolecular, one and a reductionist localized micromolecular alternative.<sup>1,2</sup> Precise and accurate characterization, using several methods, will allow for a complete understanding at the molecular level of the conformational preferences of individual amino acids in larger peptide models.

The setting up of a standardized and numerical definition of molecular structure, nondependent upon visualization<sup>3,4</sup> allows for a central data set to be engineered, constructed, and compiled. This is able to be used by all disciplines finding usefulness in the characterization and management of ‘applicable’ data sets (i.e. data warehousing, networking, and storage).

Through the investigation of molecular conformational probability distributions, one may be able to understand the bases of stability and the reason for the ‘intermittent’ or instantaneous ‘poses’ (conformations) the molecule adopts.<sup>5-10</sup> The forces stabilizing the molecule and allowing/disallowing the change

\* Corresponding author e-mail: gchass@giocomms.org.

† Global Institute of COmputational Molecular and Materials Science (GIOCOMMS).

‡ University of Toronto.

§ Laboratoire Francis Perrin.

|| Department of Chemistry, University of Szeged.

⊥ Creighton University.

# Department of Medical Chemistry, University of Szeged.

★ Hungarian Academy of Sciences, University of Szeged.

☆ Instituto de Biocomputación y Física de Sistemas Complejos (BIFI).

◇ Zymeworks Inc.

from one conformational 'pose' to another (conformational change) are tied into the bases of the interactions themselves. Stable structural poses (conformations) represent structural manifestations of an 'energetic agreement' between all degrees of freedom.<sup>5,11-12</sup>

All degrees of freedom must, therefore, be accurately characterized to quantitatively evaluate their separate energetic contributions to the total energy of a structure; known as energetic partitioning.<sup>5,10,11,13</sup> Energetic partitioning can only be performed accurately and precisely (i.e. repeatable), if and only if all 3N-6 degrees of intramolecular freedom are included into an analytical expression. This analytic formalism may then be used to uncover and to quantitatively characterize the coupling between all of these 3N-6 degrees of freedom.

A proper and complete examination of the normal modes of molecular vibration (stretches, bends, wags, rocks, rotations, scissoring, etc.) can be used to evaluate the magnitude of coupling between the degrees of freedom and may be performed using theoretical quantum chemical computations in joint with experimental Infrared (IR) examinations.<sup>11</sup>

The degree of red and blue shifting of theoretically and experimentally determined absorption bands is attributed to structural properties and intramolecular interactions. Differing structures, isomers, enantiomers, and even conformers provide nonequivalent environs for each normal mode, effecting an activation or deactivation of nuclear vibration. Those vibrational modes' movements that are facilitated or restricted require higher or lower energy, respectively, to activate.

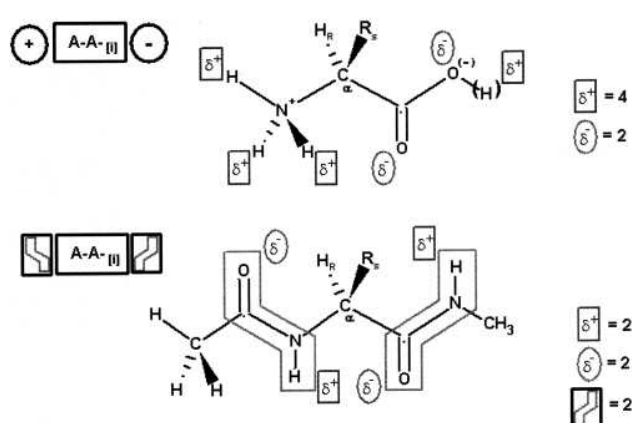
For example, the fully extended beta-backbone conformation,  $\beta_L$  ( $\varphi_i, \psi_i \equiv anti \equiv +120^\circ \leftrightarrow +240^\circ$ ), restricts the normal modes involving the carbonyl oxygen and amidic-hydrogen forming a hydrogen bond (H-bond) (i.e. C=O and N-H stretches) effecting a higher energy to activate. The elevated energetic requirement may be observed as a blueshift in the absorption band of the affected normal modes.

These red and blue shifted absorption bands constitute the basis of conformational assignment from experimental IR data. Working with the red and blue shifting of these peaks, one can work a solution into the basis for these shifts—due to interactions with those moieties making the stretch either more or less 'energetically costly' to activate/animate. It is expected that the results from this theoretical study on Ac-Phe-NH<sub>2</sub> (NAPA) will aid in the experimental and theoretical assignment of peaks for di- and tripeptide systems containing phenylalanine.

The IR experiments must be highly deconvoluted and precise enough, where working on cooled species, with a conformational selection carried out using UV spectroscopy (double resonance IR/UV spectroscopy) to allow for structural settling into the minima on the PEHS. This affords high resolution, facilitating the end goal of characterizing each peak. Clearly the gas phase is the best choice for quantitative characterization of the qualitatively proposed experimental hypotheses and expectation.

With a joint experimental and theoretical characterization of the normal modes one may accurately quantify energetic topologies (Potential Energy Curves, Surfaces and Hypersurfaces, PEC, PES, and PEHS, respectively) and Morse potentials. From the Morse potentials, one can then reparametrize force fields not only based on molecular constitution (i.e. peptide primary structure) but also on conformation and its probability distribution. This sort of mathematical amelioration of force field codes is currently ongoing and will improve the existing and related molecular mechanics (MM) and molecular dynamics (MD) studies,<sup>14-16</sup> specific to each peptide residue, one of the long-term goals of this work.

CHART 1



Free amino acid models are not appropriate for modeling peptide folding,<sup>17</sup> since the free amino acid model allows for stabilizing or destabilizing forces to be included that are not valid for peptides.<sup>17</sup> Chart 1 shows how this is possible for an amino acid (labeled i); here the relatively acidic and highly polar C-terminal carboxyl hydroxyl group ( $-COOH$ ) allows for intramolecular interactions (whether of H-bond, van der Waals (vdW) or London and Dispersion character) which may be either stabilizing or destabilizing in nature.

The N- and C-terminally protected amino acid diamide, as a dipeptide model, is able to model the inductive (through bond) and field (through space) energetic and electronic density contributions of neighboring peptide residues. As can be seen on the right side of Chart 1, the free amino acid model neither includes these neighboring influences nor does it properly provide the electrostatic environment of the peptide group.

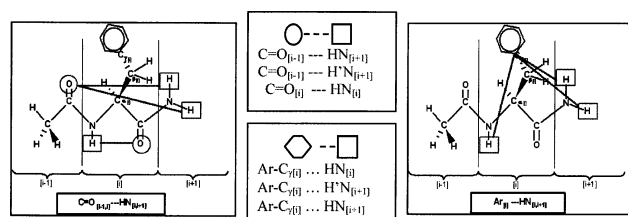
Intramolecular ring formation, via hydrogen bonding, is discussed in further detail in section 4.4.1; however, it is stated here that the  $\beta_L$ ,  $\gamma_L$ , and  $\gamma_D$  conformers form 5-, 7- and 7-membered rings, respectively. These conformers are sometimes named according to their intramolecular rings, specifically  $C_5$ ,  $C_7^{ax}$  (ax = axial), and  $C_7^{eq}$  (eq = equatorial), for the  $\beta_L$ ,  $\gamma_L$ , and  $\gamma_D$  structures, respectively; axial and equatorial are in reference to the position of the  $C_\beta$  (atom #13 in this model) relative to the intramolecular ring formed.

The  $\beta_L$  backbone conformation provides the required geometry for a  $C=O_{[i]} - -HN_{[i]}$  interaction (also known as the  $C_5$  conformation) to occur, where a gamma backbone conformation ( $\gamma_L, \gamma_D$ ) allows for a  $C=O_{[i-1]} - -HN_{[i+1]}$  stabilizing interaction ( $C_7^{ax}, C_7^{eq}$ , respectively). Free amino acid models do not allow for the latter type of intramolecular stabilization to occur, as they lack the  $C=O_{[i-1]}$  and  $HN_{[i+1]}$  groups.

The Ac-Phe-NH<sub>2</sub> system was chosen in this work to allow for both symmetric and antisymmetric C-terminal amine N-H stretches to be collected experimentally. The focus of this work favors the formulation of a theoretical-experimental synergy, therefore using the  $-NH_2$  terminus, over the use of the more structurally accurate N-methylamide ( $NH-CH_3$ ) C-terminal protecting group. Future works could undertake the characterization of the differences and similarities of Ac-Phe-NH<sub>2</sub> and Ac-Phe-NH-CH<sub>3</sub>.<sup>18</sup>

Weakly polar interactions have long been thought to be influential in both the conformational probabilities (preferences) of peptide models and the perturbation of PEHSs.<sup>6-9,19-39</sup> All interactions, particularly  $C=O - -HN$  and  $Ar - -HN$  (Chart 2), and even the possible  $C=O - -HC$  and  $N - -HN$  interactions, may in fact be directing, influential, and observable in selected experimental and theoretical undertakings.

## CHART 2



The possible C=O...HN and Ar-C<sub>γ</sub>...HN interactions are diagrammatically depicted in Chart 2.

However, the quantitative characterization of the nature, identity, and energetic contribution of these interactions is a challenge to modern experimental and theoretical techniques and remains to be established. For example, a recent work's attempt to characterize the Ar...HN interaction was biased in its hypothesis that the interaction involves the centroid of the aromatic ring.<sup>17</sup> The centroid may be defined as the center of geometry, charge, or nucleophilicity (i.e. the specific atom or orbitals) involved in the donation of density to an amidic hydrogen.

It is proposed that matching of theoretical IR band signatures of Ac-Phe-NH<sub>2</sub> with experimentally determined bands will allow for a more complete understanding of the conformational preference of Phe in peptides as well as further established synergy between theory and experiment.<sup>40</sup>

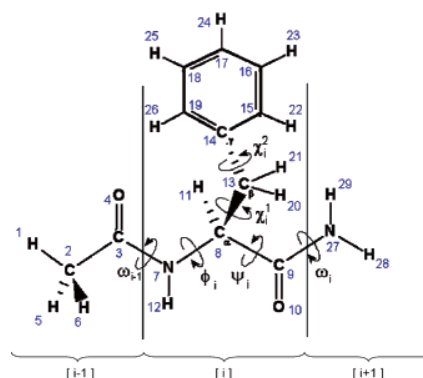
## 2. Computational Methods

The Gaussian 98 program package (G98)<sup>41</sup> was used for all computations in this work. The common convergence criteria of  $3.0 \times 10^{-4}$ ,  $4.5 \times 10^{-4}$ ,  $1.2 \times 10^{-3}$ , and  $1.8 \times 10^{-3}$  were used for the gradients of the Root Mean Square (RMS) Force, Maximum Force, RMS Displacement, and Maximum Displacement vectors, respectively.

To meet the 'design criteria' for scalable *ab initio* biological 'building block' studies, a modular construct was employed that allowed for addition and/or removal of any portion of the model, without gross perturbation to the remainder.<sup>3,40-41</sup> The systematic construction of the Ac-Phe-NH<sub>2</sub> peptide model using precomputed molecular moiety 'modules' also benefited from the use of precomputed data. In turn, the constituent and 'total' Ac-Phe-NH<sub>2</sub> module will itself be able to be used in subsequent studies of larger and more complex (di-, tri-, oligo-) peptide systems containing Phe.<sup>18</sup> Although all assemblies of precomputed 'building blocks' must still be geometry optimized, the use of preoptimized portions allows for an overall increase in computational efficiency as well as theoretical accuracy and precision.<sup>18,42</sup> Figure 1 shows the modular nature of the model.

Unix-shell and *Practical Extraction and Report Language* (PERL)<sup>43</sup> scripts were developed and employed in order to increase efficiency of data and networking management. Computations were performed on highly available and distributed algorithm-specific hardware architectures to achieve fast, efficient, and highly organized results, housed in a growing database of computed structures,<sup>3,4</sup> easily accessed, and reused for related works. This and other ongoing computational studies were designed in preparation of processing vast amounts of data and allowing On-Line Analytical Processing (OLAP) manipulation, subsequent tabulation, and analysis of results.<sup>43</sup>

All atoms were numbered (Figure 1) and input structures constructed in accordance with an established standardized, numeric, and explicit methodology.<sup>3,4</sup> Each structure was initially geometry optimized using the *ab initio*<sup>44</sup> Restricted



**Figure 1.** Representation of an established standardized numbering system of constituent atomic nuclei. Modular nature assigns a number to each atom of N-Ac-Phe-NH<sub>2</sub> (NAPA) consecutively from the N-terminus (left-most [i-1] module) through the central Phe (central [i] module) to C-terminal NH<sub>2</sub> protecting group (right-most [i+1] module). The 6 most structurally influential dihedral angles are shown (curved arrows) with their symbolic labels.

Hartree-Fock (RHF)<sup>45</sup> method, employing the split-valence 3-21G basis set.<sup>46-48</sup> Multi-Dimensional Conformational Analysis (MDCA)<sup>49</sup> was used to define the scope for the exhaustive conformational search, as the topologically possible set of conformers dictated by a grid-defined set of catchment regions. The large number of conformers is necessary to accurately characterize the topologically probable (stable) set of conformers emerging from the topologically possible set.<sup>1-4,17,18,40,42,50-54</sup> Conformational nomenclature follows the rules outlined in the literature.<sup>1-4,17,18,40,50-54</sup>

The  $\chi_1^2$  side chain dihedral angle (phenyl ring rotor) was modeled in the *gauche*<sup>+</sup> (*g*<sup>+</sup>), *anti* (*a*), and *gauche*<sup>-</sup> (*g*<sup>-</sup>) conformations. However, the *g*<sup>+</sup> and *g*<sup>-</sup> rotamers have been shown to be stable and degenerate;<sup>17</sup> only the *g*<sup>+</sup> conformers are reported in this work.

The RHF/3-21G geometry optimized structural parameters were used as input in a subsequent theoretical refinement step, achieved using the more mathematically complete 6-31G(d) basis set. The RHF/6-31G(d) results were further refined through the inclusion of electron correlation effects at the BLYP/6-311G(df,p) level, having been established as reliable for reproducing vibrational frequencies<sup>55</sup> and labeled L/6fp. The B3LYP method<sup>56-58</sup> was also used for comparison with other works, employing the split-valence 6-31G(d), 6-31G(d,p), and 6-31+G(d) levels of theory,<sup>46-48</sup> labeled B/6, B/6p, and B/6+, respectively. Any conformer leaving its MDCA-defined catchment region at the lower RHF/3-21G level, its ideal MDCA-defined structure was reattempted at the RHF/6-31G(d) level and if necessary at any of the L/6fp, B/6, B/6p, and B/6+ levels. Of the RHF and B3LYP methods, only the L/6fp and B/6+ computed results are reported in this work; the B/6 and B/6p results are reported as Supporting Information.

Each stable conformer was subjected to frequency calculations in order to confirm their identity as being at true minima. The results also provided Zero Point Energy (ZPE) values, which were scaled by using a correction factor of 0.967<sup>59</sup> and added to the total energy of each conformer to provide more accurate energetic characterization of the conformers as well as the frequency of each of the normal modes. Frequencies were scaled by factors of 0.9600 as 0.9800, for higher and lower frequencies, respectively.<sup>59</sup>

Orbital populations and wave function-outputs were generated from the B3LYP/6-31G(d,p) optimized structures. Atoms-In-Molecules (AIM)<sup>60</sup> analysis was employed on the wave func-



tion-outputs. Bond Critical Points (BCPs), Ring Critical Points (RCPs), and Cage Critical Points (CCPs) were identified, and their positions were located for the  $\beta_L^a$ ,  $\gamma_L^{g^+}$ , and  $\gamma_L^{g^-}$  conformers. In the present work, BCP properties were obtained using the AIMPAC<sup>61</sup> and AIM98PC<sup>62</sup> program packages. The molecular graphs (AIM diagrams) presented were calculated and plotted using the AIM2000<sup>63</sup> program.

Structural variables of the  $\beta_L^a$ ,  $\gamma_L^{g^+}$ , and  $\gamma_L^{g^-}$  conformers, emerging from the stable B3LYP/6-31+G(d) geometry optimized set, were used as input files for geometry optimizations using the Møller–Plesset second-order (MP2) method,<sup>64</sup> in combination with the 6-31+G(d) basis set. This level of theory is labeled M/6+. This refinement employing perturbation theory was applied to Acetyl-Glycyl-amide (Ac-Gly-NH<sub>2</sub>) and Acetyl-Alanyl-amide (Ac-Ala-NH<sub>2</sub>), to uncover the basis of stability in NAPA. Specifically to increase understanding of the energetic contributions from the side chain and the intramolecular interactions involving the aromatic ring of Phe.

All molecular visualizations were constructed using the Molekel program package,<sup>65,66</sup> including structures and orbital population surfaces; the latter were constructed using a 0.098 cutoff value,<sup>67</sup> a fraction of the maximum physical extent of the electronic population distribution. Interpolated grid points with a value equal to this cutoff are considered to be the surface (points with a greater value are within the surface, lesser are without). The choice of cutoff is generally arbitrary; if it were too small the orbitals would engulf the molecule and little information could be discerned; if it were too high, then important features of the electron distribution (such as continuity over multiple atoms) might not be seen.<sup>68–71</sup>

All experimental IR results were obtained from the accompanying work.<sup>72</sup>

### 3. Results and Discussion

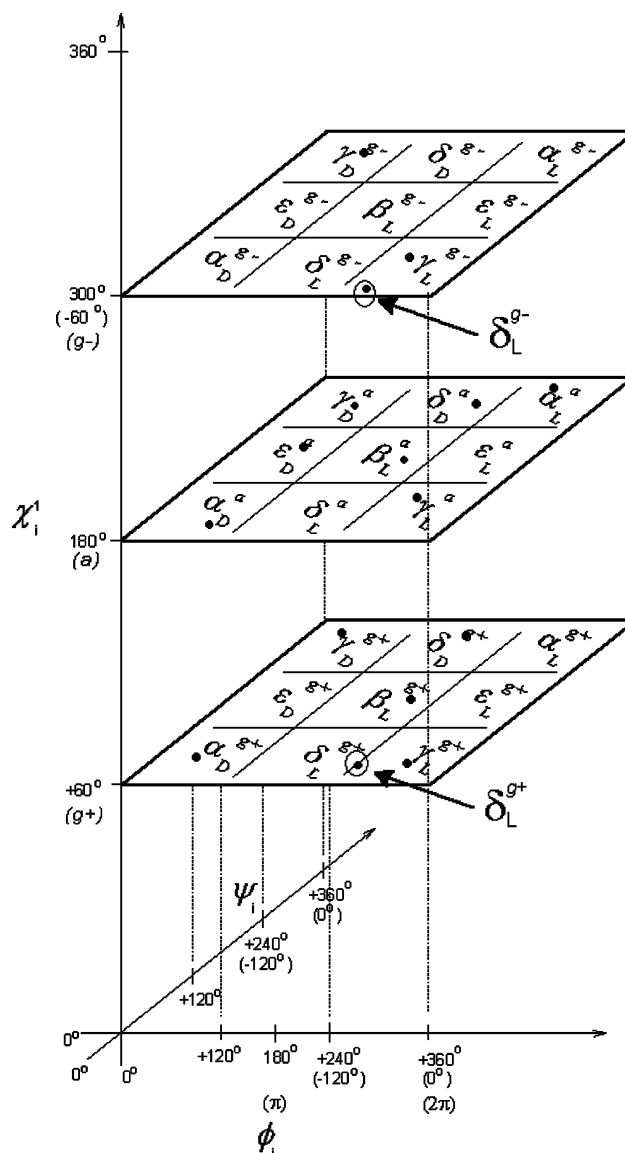
A qualitative proposal is made as to the foundations of the driving forces responsible for the conformational preferences of NAPA and other model peptide systems. The results are reported as backbone conformations following established conformational nomenclature.<sup>3,4,6,13</sup>

In total, 34 and 32 conformational minima, of the 81 MDCA-predicted and attempted, were found and confirmed by frequency analysis, for the L/61fp level and the B/6+ levels, respectively. Within the set of stable minima characterized, 17 and 16 were unique for the L/61fp and B/6+, respectively, with the  $\chi_i^2$  torsion providing the degeneracy, as mentioned in the methods section. No stable  $\epsilon_L$  backbone conformer was found. Figure 2 shows the approximate location of the B/6+ geometry optimized conformers on a traditional ‘cut’ (0° → 360°) of the Ramachandran map. Figure 3a,b shows the structure of the  $\delta_L^{g^+}$  and  $\delta_L^{g^-}$  BB conformers.

The three structures closest in agreement to experimental spectra<sup>72</sup> are displayed using visualization [Figure 4a–c].

**3.1. Geometric Parameters and Structure.** The optimized geometries for the structures at minima are listed in Table 1 and Supporting Information Table 1, displaying the results for the 6 most structural influential dihedral angles; specifically  $\omega_{i-1}$ ,  $\varphi_i$ ,  $\chi_i^1$ ,  $\chi_i^2$ ,  $\psi_i$ ,  $\omega_i$  (Figure 1).

Two exceptions were found for the cross-level agreement of the topologically probable (stable) set of conformers, with  $\beta_L^{g^-}$  being unstable at the B/6+ level and converging to the  $\gamma_L^{g^-}$  conformer, despite repeated computational attempts to locate a stable minimum in the former catchment region. The same results emerged for the  $\alpha_D^{g^-}$  conformer, which continually

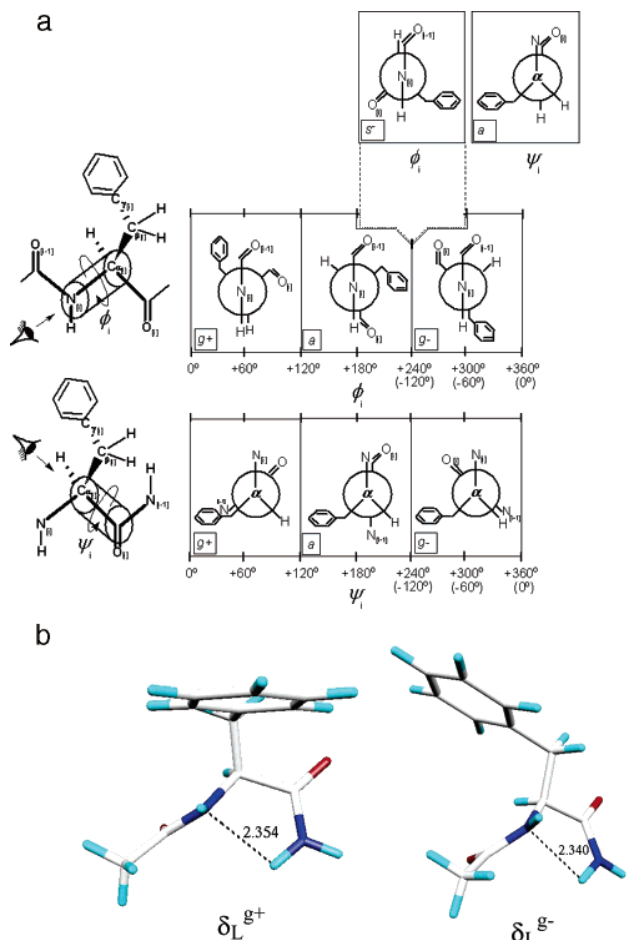


**Figure 2.** Traditional ‘cut’ (0° → 360°) of the Ramachandran map, showing the approximate location (dots) of stable backbone conformers, geometry optimized at the B3LYP/6-31+G(d) level of theory, for  $\chi_i^1 \equiv g^+$  (top),  $\chi_i^1 \equiv a$  (middle), for  $\chi_i^1 \equiv g^-$  (bottom). Backbone conformational nomenclature is ‘labeled’ in each catchment region. The two  $\delta_L$  backbone conformers are circled to highlight the ‘borderline’ eclipsing geometries for their  $\varphi_i$  dihedral angle ( $\varphi_i \approx +240^\circ$ ).

converged to the  $\gamma_L^{g^-}$  conformer. These attempts follow the methods described in the literature.<sup>17,40</sup>

The B/6+ geometry optimized results for the  $\beta_L^{g^-}$  and  $\alpha_D^{g^-}$  conformers are reported as not found (N/F) in its respective place in Table 1. The ‘missing conformers’ at the 6-31+G(d) level of theory are attributed to a ‘smoothing’ of these regions of the potential energy hypersurface as a result of the inclusion of diffuse functions.

Despite being traditionally reported as being planar and rigid, the peptide bond displays some perturbation of this ‘structural ideal’. The magnitude of this perturbation is on the order of 11.47°, 15.54°, 14.94°, and 13.01° for the  $\omega_{i-1}$  dihedral of the  $\delta_L^{g^-}$  L/6fp, B/6, B/6p, and B/6+ conformers, respectively. The  $\omega_i$  dihedral shows a maximum deviation from planarity of 8.48° for the  $\delta_D^{g^+}$  L/61fp conformer and 10.55°, 10.02°, and 9.22° for the  $\delta_L^{g^+}$  L/6fp, B/6, B/6p, and B/6+ conformers, respectively.



**Figure 3.** a. The resultant eclipsing  $\varphi_i$  rotor, where the C $_{\alpha}$ -H hydrogen and the C=O $_{[i-1]}$  oxygen are the eclipsing nuclei (top right). Despite being in an eclipsing conformation, the  $\delta_L^{g+}$  conformer is relatively close in energy to the lowest energy conformer (see right-most column in Table B). Detailed Newman projections and viewing angles are also included for all MDCA-predicted rotamers of  $\varphi_i$  (middle) and  $\psi_i$  (bottom), for the  $g^+$ ,  $a$ ,  $g$  rotamers (left, middle, right), respectively. The eclipsing  $\varphi_i \cong +240^\circ$  rotamer is fit relative to the three ‘ideal poses’ of the  $\varphi_i$  dihedral angle (dashed lines). b. Visual representations of the B3YLP/6-31+G(d) geometry optimized and frequency confirmed  $\delta_L^{g+}$  (left) and  $\delta_L^{g-}$  (right) conformers. The N $_{[i]} - \dots - \text{HN}_{[i+1]}$  interatomic distances (Å) are shown (dashed lines). These potential interactions form 5-membered intramolecular rings. Slight pyramidalization of the N $_{[i]}$  bond angles is apparent in both structures.

The origin of the nonplanarity in the  $\omega$  dihedral angles is unknown at this point; however, there seems to be a relationship with the proximal bond angles and the degree of pyramidalization or planarity shown at the amidic N $_{[i]}$  and C $_{\alpha[i]}$ , respectively. This phenomenon requires a separate study to determine its origin and influences upon structure.

Table 1 also shows that overall  $\varphi_i$  is structurally ‘well behaved’, whereas  $\psi$  is not. Furthermore, the  $\varphi_i$  dihedral adheres closely to the  $g^+$ ,  $a$ , and  $g$  ‘sectioning’ dictated by MDCA, with  $\delta_L$  being the only exception with  $\varphi_i = 120^\circ$ ; Figure 3 shows this special eclipsing case for  $\varphi_i$ . STDEVs are also shown in Table 1 for the  $\omega_{i-1}$ ,  $\chi_i^2$ , and  $\omega_i$  dihedrals;  $\chi_i^1$  was not subjected to this statistical treatment.

**3.2. Total Energy, Zero-Point-Energy, and Relative Energy.** Energetic results for total energy (Hartree), ZPE, and scaled-ZPE adjusted relative energy (kcal $\cdot$ mol $^{-1}$ ) are listed in Table 2.

Considering the scaled-ZPE adjusted relative energy, one finds that the  $\gamma_L^{g+}$  conformer is the most stable (global

minimum) for the L/61fp and B/6 and B/6p levels of theory. With the inclusion of diffuse functions, using the B/6+ level of theory, the  $\beta_L^a$  structure is at the global minimum.

All BB conformers, with the exception of  $\beta_L$ , show an increase in relative energies with inclusion of corrected-ZPE. This is attributed to all NON-  $\beta_L$  BB conformers being in ‘deep’ minima and  $\beta_L$  being in a relatively flat part of the surface. The ‘flat topology’ of the  $\beta_L$  BB conformers are confirmed by these structures having low ZPE values.

The inclusion of diffuse functions has some influence on the relative energetic order of the conformers; a simple sketch of the molecular orbital (MO) overlaps expected between interacting atoms for each conformer is shown in Figure 5. The  $\beta_L$  conformer shows that the C=O $_{[i]}$  orbital must be evaluated further from the nucleus to have sufficient MO-overlap for effective sharing of the electron density. The MOs of the  $\gamma_L$  and  $\gamma_D$  conformers already have effective overlap between the orbitals, inherent to these geometries; this is displayed by their relatively close spatial proximities (Table 3).

Clearly Figure 5 is an extremely simplified and ‘human-centric’ interpretation of the MO overlap but does show how the inclusion of diffuse and polarization functions perhaps aids to more significantly stabilize the beta-L conformer, in relation to the others. The  $\beta_L^a$  B/6+ conformer is more stable than  $\gamma_L^{g+}$ , although not by a significant amount ( $\sim 0.6$  kcal $\cdot$ mol $^{-1}$  each). This ‘ $\beta$ -lowering’ is observed for the L/61fp, B/6p, and B/6+ levels. The qualitative analysis in Figure 5 also brings up the notion that a second diffuse function should be applied to all peptide model computations, as diffusivity, when applied to H atom, would be quite an important contributor to stability.

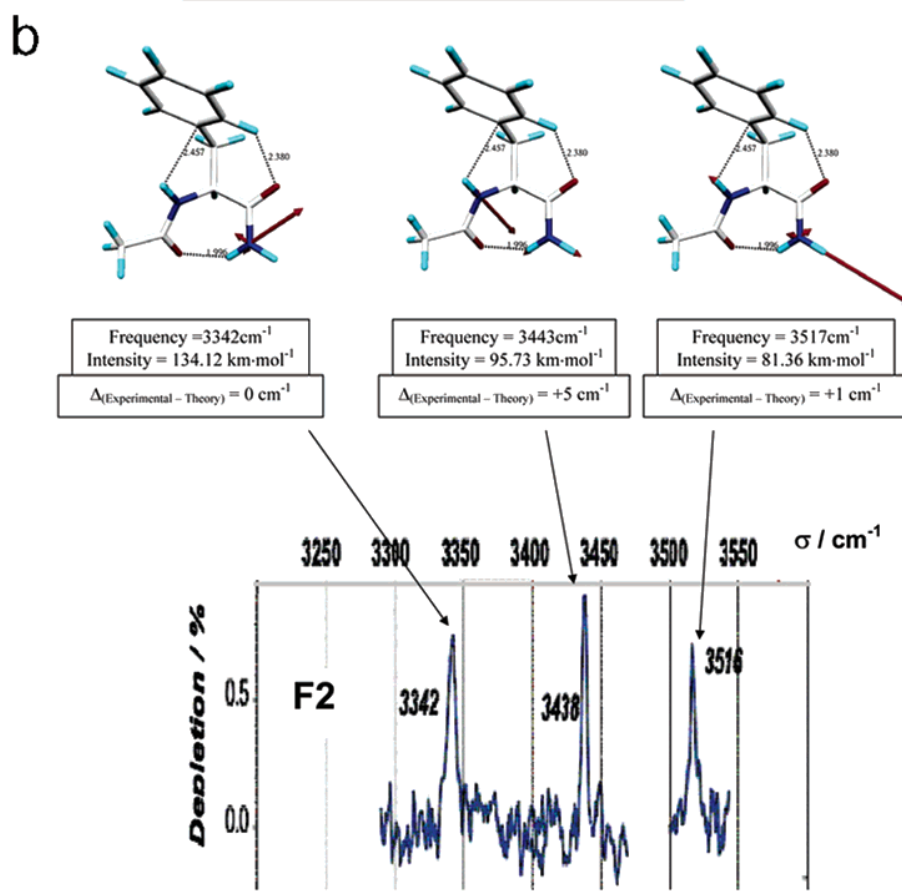
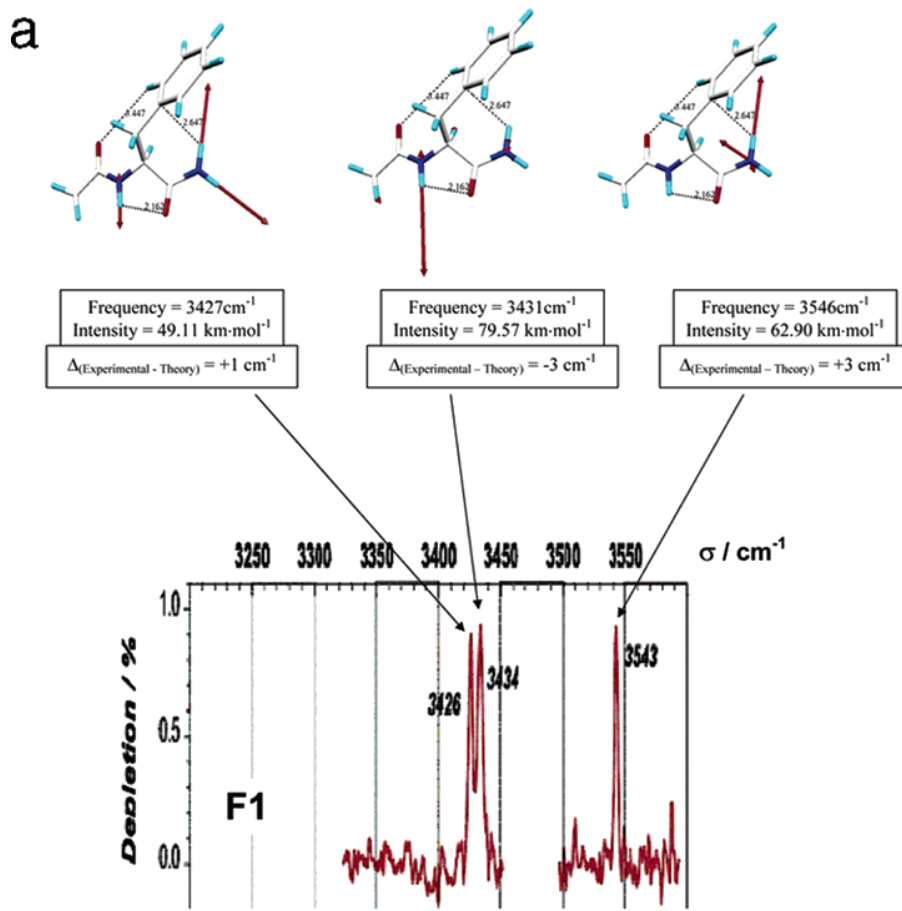
Classically, smaller rings are predicted as being more stable and thus one would expect the  $\beta_L$  structure to have a higher stability than the  $\gamma_L$  conformer.

The  $\delta_L$  conformer ‘defies’ classical theory by having a low relative energy, despite having  $\varphi_i \cong -120^\circ$  (better described as an eclipsing conformation, where the acetyl C=O $_{[i-1]}$  and the C $_{\alpha}$  hydrogen (C $_{\alpha}$ -H) are the atoms eclipsed). Due to the trigonal planar structure of the acetyl C (C=O $_{[i-1]}$ ), the C $_{\beta}$  side chain and C=O $_{[i]}$  and substituents do not eclipse the methyl carbon of the Ac group. Despite the ‘energetic debit’ due to the structural arrangement, the allowed intramolecular interactions stabilize the structure to the extent that it is at a genuine minimum, with a low relative energy.

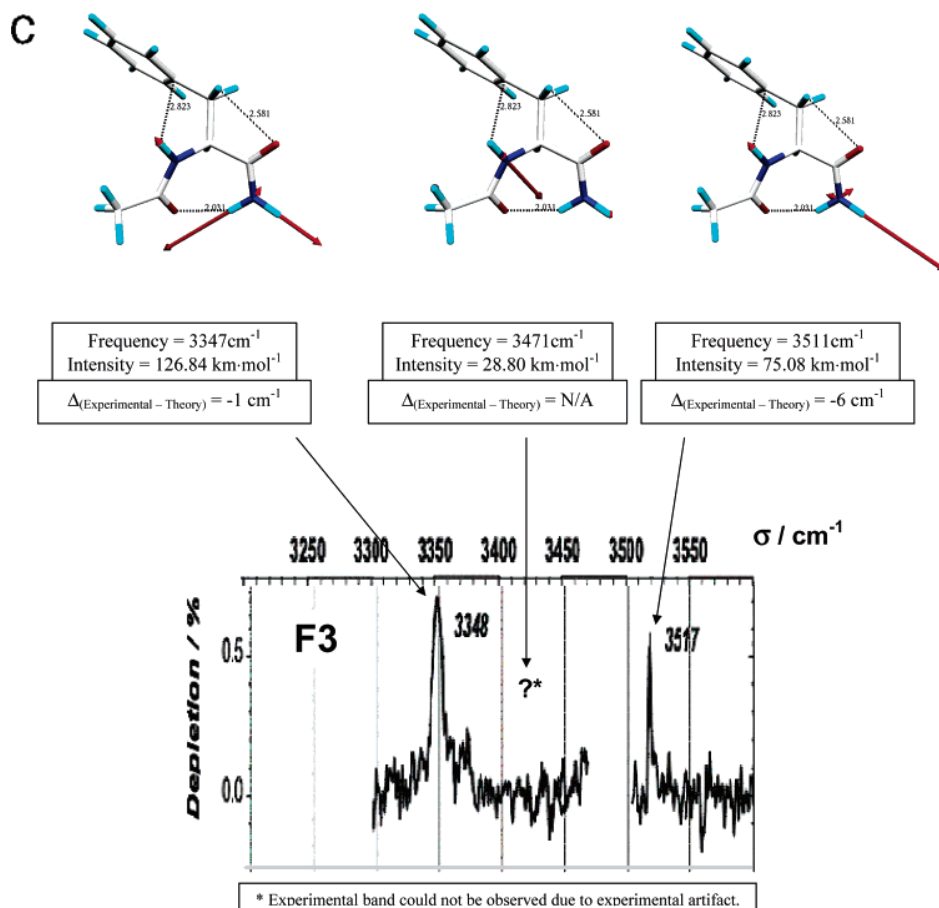
The  $\delta_D$  conformer does not show this  $\varphi_i$  structural probability due to the need for the acetyl C=O $_{[i-1]}$  and C $_{\beta}$ -(Ar) atoms to eclipse. This conformation creates a very ‘energetically costly’ steric arrangement that could not be possibly recovered by any interaction facilitated by the ‘classically forbidden’  $\varphi = +120.0^\circ$  backbone structure.

Accordingly, its relative energy is also much higher than both of the  $\delta_L$  conformers, clearly indicating that there are very powerful stabilizing forces attributed to the  $\delta_L$  structure, absent in the  $\delta_D$  BB conformers.

**3.3. Hydrogen Bonding and Other Intramolecular Interactions.** The presence or absence of intramolecular H-bonding and weakly polar interaction may be qualitatively observed, for all optimized conformers, in Table 3. However, proximity of oppositely polarizable centers is not the sole determining factor of the presence, absence, or strength of H-bonding or weakly polar interaction. Geometric threshold separation may only be used as an indicator of a potential for interaction between polarizable centers. Sufficient MO overlap is required to support the claim of a stabilizing interaction; a reiteration of the observations made for the energetic results (Figures 5 and 6).







**Figure 4.** Comparison between experimental and computed IR spectra for the B3LYP/6-31+G(d) geometry optimized  $\beta_L^\alpha$ ,  $\gamma_L^{g+}$ , and  $\gamma_L^{g-}$  conformers of NAPA (parts a–c, respectively). Systematic conformational analysis resulted in accurate identification of these conformers as being candidate to most effectively reproduce experimentally determined normal modes of vibration. Dashed lines show selected intramolecular interaction distances (Å) within each conformer. Arrows show direction of force gradients associated with each normal mode of vibration. Computed frequencies have been corrected by a scaling factor of 0.960. Computed peak intensities and the difference between experimentally and theoretically determined vibrations are also shown.

Perusing Table 3, one finds that a trend is displayed, where interactions ‘turn on and off’ like switches, with classical structural ideals satisfied.

A number of the following 3.3 subsection results are presented and discussed in the context of AIM analyses (Section 3.5). The reader is encouraged to peruse Section 3.5 and to oscillate between all figures, tables, and discussions to have a complete set of results and observations.

**3.3.1. Traditional Backbone–Backbone Interaction: C=O...H–N.** The distances between the carbonyl oxygen (C=O<sub>[i-1]</sub>, C=O<sub>[i]</sub>) and amide H (N<sub>[i]</sub>-H, N<sub>[i+1]</sub>-H, N<sub>[i+1]</sub>-H') atoms are listed in Table 3. Both the  $\beta_L$  and  $\gamma_L$  BB conformations display distances common to this well-established BB–BB H-bond interaction (~1.8–2.2 Å). More specifically C=O<sub>[i-1]</sub>...H–N<sub>[i]</sub> and C=O<sub>[i-1]</sub>...H–N<sub>[i+1]</sub> atoms are within these distances, for  $\beta_L$  and  $\gamma_L/\gamma_D$ , respectively; no other BB conformer is a candidate for these interaction types. These interactions can be observed in Figure 4a–c (dashed lines), Figure 5 (overlapping orbital ‘lobes’), Figure 6 (density overlap in orbitals, indicated by ovals), and Figure 7a–c (indicated by arrows).

**3.3.2. Aromatic-Amide Side Chain–Backbone Interaction: Ar...H–N.** The interaction has previously been characterized as involving the C<sub>γ</sub> atoms of the Phe side chain aromatic ring.<sup>18</sup> The distances between the aromatic C<sub>γ</sub> and the three amide H (N<sub>[i]</sub>-H, N<sub>[i+1]</sub>-H, N<sub>[i+1]</sub>-H') atoms are listed for each level of theory, in columns 4, 5, and 6 in Table 3, defined as (Ar<sub>[i]</sub>)–C<sub>γ</sub>...H–N<sub>[i+1]</sub>. These interactions can be observed in Figure

6 (density overlap in orbitals, indicated by ovals) and Figures 8 and 9 (indicated in the legends).

The  $\beta_L^\alpha$ ,  $\gamma_L^{g+}$ ,  $\gamma_L^{g-}$ ,  $\delta_L^{g+}$ ,  $\delta_D^{g+}$ , and  $\epsilon_D^a$  conformers all show a potential for this interaction; the first three also having the three lowest corrected-ZPE-scaled relative energies. A distance <3.0 Å was used as a ‘scoring’ criterion. Only the  $\beta_L$  conformer has AIM confirmed data for the existence of the interaction.

**3.3.3. Alternate Backbone–Backbone Interaction: N...H–N.** The distances between the amidic nitrogen (N<sub>[i]</sub>, N<sub>[i+1]</sub>) and amide hydrogen (N<sub>[i]</sub>-H, N<sub>[i+1]</sub>-H, N<sub>[i+1]</sub>-H') atoms are listed for each level of theory, in columns 7, 8, and 9 in Table 3. The amide...amidic-hydrogen interaction appears to be strongly stabilizing and directly responsible for the observed energetic trends of the  $\delta_L$  conformers and their resultant low relative energy. The visualization of the  $\delta_L$  structures, Figure 3b, shows the slight pyramidalization of the amidic N<sub>[i]</sub>. Once again, satisfaction of geometric threshold separation may not be used as a sole determining factor for the presence of a true N<sub>[i]</sub>...H–N<sub>[i+1]</sub> electrostatic interaction. An affirmation of an exchange of electronic probability density must be accurately identified and characterized.

The  $\alpha_L$ ,  $\alpha_D$ , and  $\epsilon_L$  conformers also show similar small geometric threshold separation (Table 3, columns 7, 8, and 9), making them a candidate for possible N<sub>[i]</sub>...H–N<sub>[i+1]</sub> interaction. The AIM analysis was not extended to the  $\delta_L$ ,  $\delta_D$ ,  $\alpha_L$ ,  $\alpha_D$ , and  $\epsilon_L$  BB conformers; the ideal is a complete analysis of all stable conformers, including AIM analysis.

**TABLE 1: Selected Backbone and Side Chain Dihedral Angles (Degrees) for N-Ac-Phe-NH<sub>2</sub> Geometry Optimized at the B3LYP/6-31G+(d) and BLYP/6-311G(df,p) Levels of Theory<sup>a</sup>**

BB	$\chi^1$	$\chi^2$	$\omega_{i-1}$		$\phi_i$		$\chi_i^1$		$\chi_i^2$		$\psi_i$		$\omega_i$	
			B3LYP/ 6-31+G (d)	BLYP/ 6-311G (df,p)	B3LYP/ 6-31+G (d)	BLYP/ 6-311G (df,p)	B3LYP/ 6-31+G (d)	BLYP/ 6-311G (df,p)	B3LYP/ 6-31+G (d)	BLYP/ 6-311G (df,p)	B3LYP/ 6-31+G (d)	BLYP/ 6-311G (df,p)	B3LYP/ 6-31+G (d)	BLYP/ 6-311G (df,p)
$\beta$	+	+	176.73	175.94	-156.17	-155.95	58.81	60.28	89.59	89.871	167.47	167.82	-175.31	-178.80
$\beta$	a	+	177.57	175.40	-158.32	-157.13	194.26	-159.15	67.55	71.25	162.34	166.44	172.44	174.37
$\beta$	-	+	N/F <sup>b</sup>	174.28	N/F <sup>b</sup>	-118.77	N/F <sup>b</sup>	-61.03	N/F <sup>b</sup>	94.80	N/F <sup>b</sup>	145.72	N/F <sup>b</sup>	179.60
	mean		-179.71	177.15	175.21	-157.25	-143.95	126.54	-173.30	78.57	85.31	164.91	159.99	178.57
	STDEV		5.46	0.59	0.85	1.52	21.81	95.78	119.97	15.58	12.42	3.63	12.38	8.66
$\gamma_L$	+	+	-175.49	-175.20	-81.88	-81.71	-316.29	-317.91	78.54	79.12	54.86	59.21	175.62	179.05
$\gamma_L$	a	+	179.74	-179.12	-82.62	-82.14	-164.96	-162.04	87.03	91.49	82.09	78.42	-170.76	-172.08
$\gamma_L$	-	+	-173.13	-173.85	-84.17	-83.17	-55.80	-53.38	114.72	112.04	73.44	73.21	-172.97	-173.66
	mean		-176.29	-176.29	-176.06	-82.89	-82.34	-179.02	-177.78	93.43	94.21	70.13	70.28	-56.04
	STDEV		3.63	3.63	2.74	1.17	0.75	130.81	132.96	18.92	16.63	13.91	9.94	200.62
$\gamma_D$	+	+	172.95	171.32	55.99	53.09	67.88	68.67	80.97	83.37	-25.36	-27.78	-174.57	-175.27
$\gamma_D$	a	+	176.73	176.98	73.48	72.34	-168.77	-170.18	82.50	84.59	-67.22	-66.50	173.68	175.16
$\gamma_D$	-	+	173.15	174.07	73.82	72.43	-58.95	-59.19	103.65	101.77	-52.42	-54.85	-177.98	-178.61
	mean		174.28	174.28	174.12	67.76	65.95	-173.28	-173.57	89.04	89.91	-48.33	-49.71	-179.62
	STDEV		2.13	2.13	2.83	10.20	11.14	116.65	116.11	12.68	10.29	21.23	19.86	6.04
$\alpha_L$	+	+	-169.64	-171.66	-119.01	-121.85	53.87	54.02	81.47	83.66	15.83	18.81	171.84	174.52
$\alpha_L$	-	+	-166.99	-168.53	-107.53	-114.15	-60.86	297.05	112.54	104.80	5.01	12.99	172.38	172.82
	mean		-168.32	-168.32	-170.10	-113.27	-118.00	-3.50	175.53	97.01	94.23	10.42	15.90	172.11
	STDEV		1.87	1.87	2.21	8.12	5.44	81.13	171.84	21.97	14.95	7.65	4.12	0.38
$\alpha_D$	+	+	171.61	170.93	-163.73	-168.39	59.91	57.01	96.48	92.56	-38.57	-36.97	-170.78	-171.52
$\alpha_D$	a	+	175.23	175.57	-154.28	-153.58	-174.47	-173.44	75.13	77.54	-59.87	-63.12	-172.65	-174.49
	mean		173.42	173.42	173.25	-159.01	-160.99	122.72	121.79	85.81	85.05	-49.22	-50.04	-171.72
	STDEV		2.56	2.56	3.28	6.68	10.47	88.83	91.60	15.10	10.62	15.06	18.49	1.32
$\alpha_L$	a	+	-169.71	N/F <sup>b</sup>	-86.13	N/F <sup>b</sup>	-170.33	N/F <sup>b</sup>	75.43	N/F <sup>b</sup>	-24.12	N/F <sup>b</sup>	171.67	N/F <sup>b</sup>
$\alpha_D$	+	+	165.23	167.71	49.50	47.39	51.49	48.70	81.70	80.74	41.01	45.16	-172.58	-172.47
$\alpha_D$	a	+	170.90	170.87	68.18	63.77	-131.20	-137.08	102.38	100.44	31.85	38.62	-173.38	-174.09
$\alpha_D$	-	+	N/F <sup>b</sup>	167.87	N/F <sup>b</sup>	69.04	N/F <sup>b</sup>	-62.32	N/F <sup>b</sup>	101.32	N/F <sup>b</sup>	27.45	N/F <sup>b</sup>	-174.89
	mean		170.37	168.07	168.82	58.84	60.07	140.15	69.77	92.04	94.17	36.43	37.08	-172.98
	STDEV		4.89	4.01	1.78	13.21	11.29	125.38	143.78	14.62	11.63	6.48	8.95	0.57
$\epsilon_D$	a	+	-163.67	-162.52	64.85	65.95	-157.24	-158.10	60.64	61.13	-168.84	-167.54	-179.66	-176.93

<sup>a</sup> The means and standard deviations for each backbone conformer are also shown where applicable. Note: means and standard deviations are not computed or shown for  $\chi_i^1$  due to the organization of the data by backbone conformation. <sup>b</sup> N/F indicates conformer not found at the level of theory.

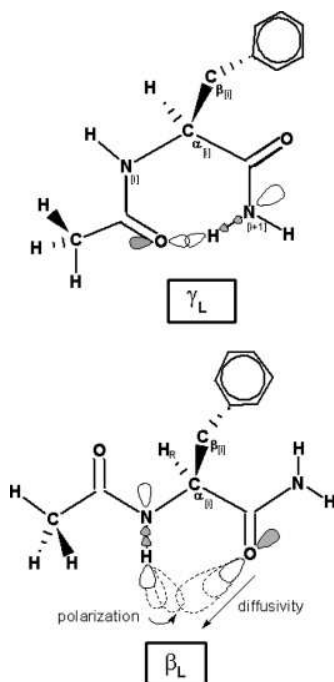
**TABLE 2: Total Energy (Hartrees), Relative Energy (kcal\*mol<sup>-1</sup>), and Corrected ZPE (Hartree)-Adjusted Rel. E. (kcal\*mol<sup>-1</sup>) for Stable Backbone and Side Chain Conformations of N-Ac-Phe-NH<sub>2</sub> at the B3LYP/6-31G+(d) and BLYP/6-311G(df,p) Levels of Theory**

BB	$\chi^1$	$\chi^2$	total energy (Hartrees)		relative energy (kcal*mol <sup>-1</sup> )		zero-point energy ZPE (Hartree/particle)		scaled by 0.967 ZPE corrected SCF (Hartrees)		scaled ZPE-corrected relative energy (kcal*mol <sup>-1</sup> )	
			B3LYP/ 6-31+G (d)	BLYP/ 6-311G (df,p)	B3LYP/ 6-31+G (d)	BLYP/ 6-311G (df,p)	B3LYP/ 6-31+G (d)	BLYP/ 6-311G (df,p)	B3LYP/ 6-31+G (d)	BLYP/ 6-311G (df,p)	B3LYP/ 6-31+G (d)	BLYP/ 6-311G (df,p)
$\beta$	+	+	-687.62181	-687.54838	2.32	2.60	0.239391	0.231188	-687.39032	-687.324821	2.16	1.95
$\beta$	a	+	-687.62550	-687.55163	0.00	0.56	0.239646	0.231567	-687.39376	-687.327709	0.00	0.14
$\beta$	-	+	N/F <sup>a</sup>	-687.54756	N/F <sup>a</sup>	3.12	N/F <sup>a</sup>	0.231040	N/F <sup>a</sup>	-687.324144	N/F <sup>a</sup>	2.37
$\gamma_L$	+	+	-687.62523	-687.55253	0.17	0.00	0.240465	0.232271	-687.3927	-687.327925	0.67	0.00
$\gamma_L$	a	+	-687.62470	-687.55174	0.50	0.50	0.240053	0.231899	-687.39257	-687.327494	0.75	0.27
$\gamma_L$	-	+	-687.62494	-687.55166	0.35	0.55	0.240002	0.231795	-687.39286	-687.327512	0.57	0.26
$\gamma_D$	+	+	-687.61407	-687.54125	7.18	7.08	0.240479	0.231843	-687.38152	-687.317053	7.68	6.82
$\gamma_D$	a	+	-687.61959	-687.54648	3.71	3.80	0.240161	0.231927	-687.38735	-687.322204	4.02	3.59
$\gamma_D$	-	+	-687.62281	-687.54921	1.69	2.09	0.240446	0.232184	-687.3903	-687.324683	2.17	2.03
$\delta_L$	+	+	-687.62278	-687.54939	1.71	1.97	0.240102	0.231643	-687.3906	-687.325389	1.99	1.59
$\delta_L$	-	+	-687.62085	-687.54702	2.92	3.46	0.239457	0.231032	-687.38929	-687.323614	2.81	2.70
$\delta_D$	+	+	-687.61598	-687.54196	5.97	6.63	0.239767	0.231727	-687.38413	-687.317878	6.04	6.30
$\delta_D$	a	+	-687.61382	-687.54020	7.33	7.74	0.239241	0.230933	-687.38247	-687.316885	7.09	6.93
$\alpha_L$	a	+	-687.61489	N/F <sup>a</sup>	6.66	N/F <sup>a</sup>	0.239242	N/F <sup>a</sup>	-687.38354	N/F <sup>a</sup>	6.41	N/F <sup>a</sup>
$\alpha_D$	+	+	-687.61104	-687.53854	9.08	8.78	0.239876	0.231436	-687.37908	-687.314741	9.21	8.27
$\alpha_D$	a	+	-687.61391	-687.54124	7.28	7.08	0.239768	0.231248	-687.38205	-687.317626	7.35	6.46
$\alpha_D$	-	+	N/F <sup>a</sup>	-687.54441	N/F <sup>a</sup>	5.09	N/F <sup>a</sup>	0.231271	N/F <sup>a</sup>	-687.320774	N/F <sup>a</sup>	4.49
$\epsilon_D$	a	+	-687.61468	-687.54074	6.79	7.40	0.239557	0.231304	-687.38303	-687.317068	6.74	6.81

<sup>a</sup> N/F indicates conformer not found at the level of theory.

3.3.4. Alternate Side Chain-Backbone Interactions: C=O...H-R. The distances between the carbonyl oxygen (C=O<sub>[i-1]</sub>), C=O<sub>[i]</sub>) and phenyl H (H-C<sub>3</sub>(Ar)) atoms are not tabulated nor

listed in this work. The data were not collected nor presented, as the interaction was not predicted as being significant in the NAPA system. However, the potential for the C=O<sub>[i-1]</sub>-H-



**Figure 5.** Schematic representation of the expected energetic stabilization attributed to the basis set employed in the geometry optimization of the  $\gamma_L$  (top) and  $\beta_L$  backbone conformers (bottom) of NAPA. Although qualitative, the diagram shows how employing basis sets including polarization and diffuse functions can more significantly stabilize one conformer relative to another. This is due to atomic and molecular orbital alignments and symmetries of the  $p$ -orbitals of O and the  $\sigma^*$  orbitals of the amidic H, influenced by their geometries. In the  $\gamma_L$  conformer, there may be a higher degree of orbital (MO) overlap inherent to the structure, relative to the  $\beta_L$  conformer. Therefore, one may predict from simple diagrams and basic theoretical principles that the  $\gamma_L$  conformer would be more significantly stabilized than the  $\beta_L$  conformer, using less mathematically complete basis sets. More complete basis sets may therefore lower the energy of the  $\beta_L$  conformer more significantly than the  $\gamma_L$  conformer.

### CHART 3

	Experimental		B3LYP/6-31+G(d)		BLYP/6-311G(df,p)		
A	1694	1725	1695	1725	1661	1694	$\beta_L^a$
B	1682	1718	1685	1725	1648	1694	$\gamma_L^{g+}$
C	1680	1719	1687	1736	1650	1706	$\gamma_L^{g-}$
			→	1722/1726	→	1693/1694	$\delta_L^{g+}$

$C_\delta$ -(Ar<sub>[i]</sub>) and (C=O<sub>[i]</sub>) - -H-C<sub>δ</sub>-(Ar<sub>[i]</sub>) interactions to be present emerged from AIM analyses of the  $\beta_L^a$  and  $\gamma_L^{g+}$  conformers. These interactions can be observed in Figure 4a–c (dashed lines), Figure 6 (density overlap in orbitals, indicated by ovals), and Figure 7a,b (indicated by arrows). Further investigation is required to more accurately characterize the existence and nature of these interactions.

**3.4. Experimental and Theoretical IR Frequencies.** The computed frequencies of selected normal modes of vibration are listed in Table 4; N-H<sub>[i+1]</sub>-symmetric, N-H<sub>[i]</sub>, N-H<sub>[i+1]</sub>-anti-symmetric stretches are found in columns 1, 3, and 5, respectively. The C=O<sub>[i-1]</sub>, C=O<sub>[i]</sub> stretches and N-H<sub>[i]</sub>, N-H<sub>[i+1]</sub>, N-H'<sub>[i+1]</sub> are found in Table 5 and Supporting Information tables, respectively. Corresponding intensities (km<sup>2</sup>mol<sup>-1</sup>) are found to the right of each frequencies' columnated results.

The experimentally determined N–H stretch spectral region (3200–3600 cm<sup>-1</sup>) are displayed in Figure 4–c. Carbonyl stretches are reported in Chart 3. Qualitatively, one may observe that the theoretically determined IR intensities match well with

the visual size of the absorption peaks. More in-depth description, presentation, and analyses of these experimentally determined results can be found in the accompanying work.<sup>72</sup>

Established literature recommends that frequency scaling-factors of 0.96 should be applied to the B/6+ geometry optimized frequencies.<sup>44</sup> The 0.96-fraction scales each of the 9 N–H stretch frequencies, of the  $\beta_L^a$ ,  $\gamma_L^{g+}$ , and  $\gamma_L^{g-}$  conformers to within 0–6 cm<sup>-1</sup>.

The 0.96-fraction does not provide numerical agreement between the experimental and for the theoretically determined B/6+ frequencies. Rather, if one applies a 0.98 scaling-factor, the C=O<sub>[i-1]</sub>, C=O<sub>[i]</sub> stretches show acceptable numerical agreement with the experimental ones. Using the 0.98 scaling-factor in conjunction with the frequencies in Table 5, one finds that the experimentally C=O-stretch-determined conformers A, B, and C agree relatively well with the  $\beta_L^a$ ,  $\gamma_L^{g+}$ , and  $\gamma_L^{g-}$  structural conformers, respectively. Chart 3 shows the experimental and 0.98-scaled theoretical C=O<sub>[i-1]</sub> and C=O<sub>[i]</sub> frequencies (cm<sup>-1</sup>) for the B/6+ level, the lower and higher values, respectively. Unscaled L/61fp results are also shown.

The C=O<sub>[i-1]</sub> and C=O<sub>[i]</sub> frequencies for the  $\delta_L^{g+}$  conformer are also shown scaled by a factor of 0.98 for the B/6+ level and unscaled for L/61fp results. The C=O<sub>[i-1]</sub> and C=O<sub>[i]</sub> frequencies are shifted to approximately 1722 cm<sup>-1</sup> and 1726 cm<sup>-1</sup>, respectively (Chart 3), matching well with the experimentally determined C=O<sub>[i]</sub> stretch. No other conformer shows this C=O near 'stretch degeneracy' (Table 5). Additionally, the conformer shows degeneracy in its H–N<sub>[i]</sub> and H–N<sub>[i+1]</sub> symmetric stretches.

The  $\delta_L^{g+}$  conformer requires a more thorough theoretical characterization, as mentioned in the structural analysis (section 3.1), if only to uncover the basis of the near-degeneracy of the two C=O and H–N stretches.

**3.5. Electron Populations and Orbital Overlap.** Surface plots of the electron density distribution are shown in Figure 6. Although the diagrams only provide a qualitative illustration of the density distribution of each static geometric 'pose', they provide a feeling for the specificity and limited nature of certain intramolecular interactions. The results also show that the level of theory, and more particularly the influence of the basis set applied, plays a dominant role in determining the degree of density overlap observed between polarized centers. Once again a lengthy text description is much inferior to a quick observation of the difference between the structures on the left and right portions of the upper part of Figure 6.

The structures in Figure 6 have been aligned in viewing space to show the interactions, where they exist. The AIM predicted C=O<sub>[i-1]</sub> - -H-C<sub>δ</sub>-(Ar<sub>[i]</sub>) interaction in the  $\beta_L^a$  conformer was not confirmed by electron population analysis as no overlap in density was observed. Inversely, the  $\gamma_L^{g+}$  and  $\gamma_L^{g-}$  conformers show density overlap consistent with the (Ar)-C<sub>γ</sub> - -H-N<sub>[i]</sub> interaction, where the AIM analysis does not show the existence of this interaction.

**3.6. Atoms-In-Molecules (AIM) Analysis.** The bond paths emerging from Atoms-In-Molecules analysis of the input wave functions are structurally depicted in Figure 7a–c. These diagrams show the Bond Critical Points (BCPs) as (red) dots, found on the line joining two atoms. Ring Critical Points (RCPs) are also shown as (yellow) dots, found at the center of intramolecular rings formed either via covalent or ionic bonding.

Figure 7a shows the  $\beta_L^a$  conformer as having 5-, 6-, and 8-membered rings formed by the following intramolecular

**TABLE 3: Interatomic Distances for Selected Intramolecular Interactions (Å) for Geometry Optimized Conformers of N-Ac-Phe-NH<sub>2</sub> at the B3LYP/6-31G+(d) and BLYP/6-311G(df,p) Levels of Theory**

BB	$\chi^1$	$\chi^2$	C=O <sub>[i-1]</sub> -H-N <sub>[i+1]</sub> (Å)		C=O <sub>[i-1]</sub> -H'-N <sub>[i+1]</sub> (Å)		C=O <sub>[i]</sub> -H-N <sub>[i]</sub> (Å)		Ar-C <sub><math>\gamma</math>[i]</sub> -H-N <sub>[i]</sub> (Å)		Ar-C <sub><math>\gamma</math>[i]</sub> -H-N <sub>[i+1]</sub> (Å)			
			B3LYP/6-31+G	BLYP/6-311G	B3LYP/6-31+G	BLYP/6-311G	B3LYP/6-31+G	BLYP/6-311G	B3LYP/6-31+G	BLYP/6-311G	B3LYP/6-31+G	BLYP/6-311G	B3LYP/6-31+G	BLYP/6-311G
			(d)	(df,p)	(d)	(df,p)	(d)	(df,p)	(d)	(df,p)	(d)	(df,p)	(d)	(df,p)
$\beta$	+	+	6.05	6.10	5.04	5.08	2.19	2.18	3.36	3.40	4.07	4.01		
$\beta$	a	+	6.08	6.11	5.08	5.11	2.16	2.15	4.51	4.55	2.65	2.66		
$\beta$	-	+	N/F <sup>a</sup>	5.48	N/F <sup>a</sup>	4.40	N/F <sup>a</sup>	2.42	3.66	N/F <sup>a</sup>	3.66	N/F <sup>a</sup>	4.56	
$\gamma_L$	+	+	3.75	3.74	2.00	1.99	3.88	3.86	2.46	2.46	4.96	5.00		
$\gamma_L$	a	+	3.76	3.74	2.16	2.10	3.44	3.53	4.11	4.12	4.59	4.65		
$\gamma_L$	-	+	3.71	3.72	2.03	2.02	3.72	3.74	2.82	2.82	5.05	5.07		
$\gamma_D$	+	+	3.58	3.58	1.83	1.80	4.25	4.29	3.96	3.95	3.52	3.49		
$\gamma_D$	a	+	3.66	3.67	1.97	1.96	3.77	3.82	4.49	4.51	3.94	3.94		
$\gamma_D$	-	+	3.69	3.69	1.94	1.93	3.99	4.00	3.44	3.46	4.43	4.42		
$\delta_L$	+	+	4.94	5.04	3.53	3.60	4.04	4.03	2.65	2.70	4.39	4.44		
$\delta_L$	-	+	4.70	4.86	3.39	3.46	4.18	4.11	3.03	3.24	5.01	5.09		
$\delta_D$	+	+	5.96	5.99	4.81	4.77	3.59	3.60	3.66	3.72	2.76	2.87		
$\delta_D$	a	+	6.00	6.05	5.01	5.05	3.45	3.46	4.53	4.54	3.65	3.65		
$\alpha_L$	a	+	4.40	N/F <sup>a</sup>	3.41	N/F <sup>a</sup>	4.35	N/F <sup>a</sup>	4.00	N/F <sup>a</sup>	4.70	N/F <sup>a</sup>		
$\alpha_D$	+	+	3.51	3.54	2.67	2.84	4.46	4.48	3.55	3.57	4.76	4.85		
$\alpha_D$	a	+	3.94	3.92	3.10	3.23	4.42	4.46	4.05	4.11	5.15	5.12		
$\alpha_D$	-	+	N/F <sup>a</sup>	3.97	N/F <sup>a</sup>	3.08	N/F <sup>a</sup>	4.46	N/F <sup>a</sup>	3.25	N/F <sup>a</sup>	5.11		
$\epsilon_D$	a	+	4.76	4.82	4.46	4.53	2.76	2.86	4.55	4.57	2.40	2.40		

BB	$\chi^1$	$\chi^2$	Ar-C <sub><math>\gamma</math>[i]</sub> -H'-N <sub>[i+1]</sub> (Å)		N <sub>[i]</sub> -H-N <sub>[i+1]</sub> (Å)		N <sub>[i]</sub> -H'-N <sub>[i+1]</sub> (Å)		N <sub>[i+1]</sub> -H-N <sub>[i]</sub> (Å)			
			B3LYP/6-31+G(d)	BLYP/6-311G(df,p)	B3LYP/6-31+G(d)	BLYP/6-311G(df,p)	B3LYP/6-31+G(d)	BLYP/6-311G(df,p)	B3LYP/6-31+G(d)	BLYP/6-311G(df,p)	B3LYP/6-31+G(d)	BLYP/6-311G(df,p)
			(d)	(df,p)	(d)	(df,p)	(d)	(df,p)	(d)	(df,p)	(d)	(df,p)
$\beta$	+	+	4.59	4.56	3.98	4.01	4.43	4.46	3.79	3.79		
$\beta$	a	+	4.02	4.12	3.99	4.01	4.43	4.46	3.76	3.77		
$\beta$	-	+	N/F <sup>a</sup>	5.57	N/F <sup>a</sup>	3.88	N/F <sup>a</sup>	4.43	N/F <sup>a</sup>	3.89	3.89	
$\gamma_L$	+	+	5.33	5.39	2.74	2.77	3.90	3.93	3.68	3.74		
$\gamma_L$	a	+	4.99	5.08	2.96	2.92	3.94	3.95	3.83	3.83		
$\gamma_L$	-	+	5.89	5.92	2.86	2.86	3.93	3.95	3.89	3.90		
$\gamma_D$	+	+	4.44	4.41	2.71	2.72	3.98	4.03	3.85	3.92		
$\gamma_D$	a	+	4.52	4.52	2.84	2.85	3.95	3.99	3.88	3.91		
$\gamma_D$	-	+	5.54	5.56	2.77	2.79	3.96	3.99	3.83	3.87		
$\delta_L$	+	+	4.91	4.95	2.35	2.36	3.76	3.78	3.11	3.11		
$\delta_L$	-	+	5.83	5.90	2.34	2.36	3.76	3.78	3.22	3.16		
$\delta_D$	+	+	4.03	4.12	2.69	2.65	3.79	3.80	2.42	2.44		
$\delta_D$	a	+	4.43	4.46	2.94	2.98	3.88	3.91	2.59	2.64		
$\alpha_L$	a	+	4.92	N/F <sup>a</sup>	2.36	N/F <sup>a</sup>	3.75	N/F <sup>a</sup>	3.29	N/F <sup>a</sup>		
$\alpha_D$	+	+	5.21	5.29	2.50	2.52	3.83	3.84	3.70	3.71		
$\alpha_D$	a	+	5.51	5.48	2.39	2.43	3.76	3.78	3.46	3.51		
$\alpha_D$	-	+	N/F <sup>a</sup>	5.91	N/F <sup>a</sup>	2.39	N/F <sup>a</sup>	3.79	N/F <sup>a</sup>	3.52	3.52	
$\epsilon_D$	a	+	4.08	4.10	4.05	4.08	4.48	4.52	4.12	4.19		

<sup>a</sup> N/F indicates conformer not found at the level of theory.

interactions: C=O<sub>[i]</sub>-H-N<sub>[i]</sub>, (Ar<sub>[i]</sub>)-C <sub>$\gamma$</sub> -H-N<sub>[i+1]</sub>, C=O<sub>[i-1]</sub>-H-C <sub>$\delta$</sub> (Ar<sub>[i]</sub>), respectively.

Figure 7b shows the  $\gamma_L^{g+}$  conformer as having two 7-membered rings formed by the following intramolecular interactions: C=O<sub>[i-1]</sub>-H-N<sub>[i+1]</sub> and C=O<sub>[i]</sub>-H-C <sub>$\delta$</sub> (Ar<sub>[i]</sub>), respectively.

Figure 7c shows the  $\beta_L^a$  conformer as having a 7-membered ring formed by the C=O<sub>[i]</sub>-H-N<sub>[i]</sub> intramolecular interaction.

**3.7. Comparison of Structure and Energy of Ac-Gly-NH<sub>2</sub>, Ac-Ala-NH<sub>2</sub>, and Ac-Phe-NH<sub>2</sub>.** The nature and magnitude of the structural and energetic trends for Ac-Gly-NH<sub>2</sub>, Ac-Ala-NH<sub>2</sub>, and Ac-Phe-NH<sub>2</sub>, geometry optimized at the B3LYP/6-31+G(d) and MP2/6+ levels of theory, may be observed in Table 6. Energetic trends are in qualitative agreement with one another in regards to the ordering by relative energy. Quantitatively, the largest difference is found between the B/6+ and M/6+ geometry optimized Gly  $\beta_L$  conformer ( $\Delta$ Energy  $\cong$  1.31 kcal\*mol<sup>-1</sup>). MP2/6+ frequency calculations were not completed for the Phe systems; therefore, frequency and ZPE values were not obtained for the Phe systems, thus no ZPE values are presented.

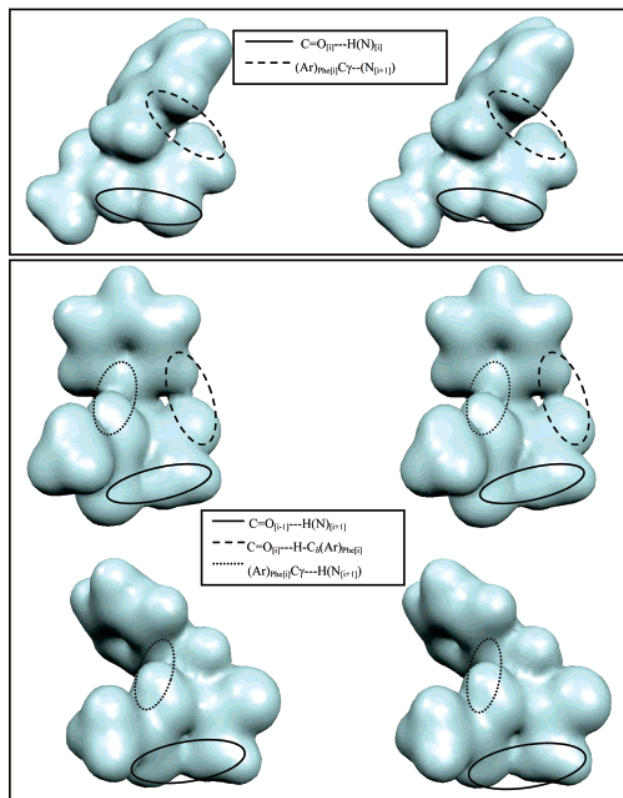
**3.7.1. Structural Trends of Ac-Gly-NH<sub>2</sub>, Ac-Ala-NH<sub>2</sub>, and Ac-Phe-NH<sub>2</sub>.** Backbone dihedral angles for Ala and Phe have more similarity than those of Gly. This may be attributed to the lack of a heavy atom in the  $\beta$ -position of the side chain of Gly. The  $\omega_{i-1}$  and  $\omega_i$  dihedral angles show enantiospecific trends in their deviation from planarity of the peptide bond.

**3.7.2. Energetic Trends of Ac-Gly-NH<sub>2</sub>, Ac-Ala-NH<sub>2</sub>, and Ac-Phe-NH<sub>2</sub>.** The Gly model shows the  $\gamma_L$  conformer to be 0.92 and 2.23 kcal\*mol<sup>-1</sup> lower in energy than the  $\beta_L$  conformer, at the B/6+ and M/6+ levels of theory, respectively. The Ala model shows the  $\gamma_L$  conformer to be 1.19 and 1.72 kcal\*mol<sup>-1</sup> lower in energy than the  $\beta_L$  conformer, at the B/6+ and M/6+ levels of theory, respectively. The lack of an aromatic side chain in Gly and Ala may be responsible for this difference in energetic ordering. Further investigation is necessary to provide an accurate energetic partitioning of all.

#### 4. Conclusion

An accurate synergy has emerged between theory and experiment, owing its accuracy to the exhaustive MDCA-directed



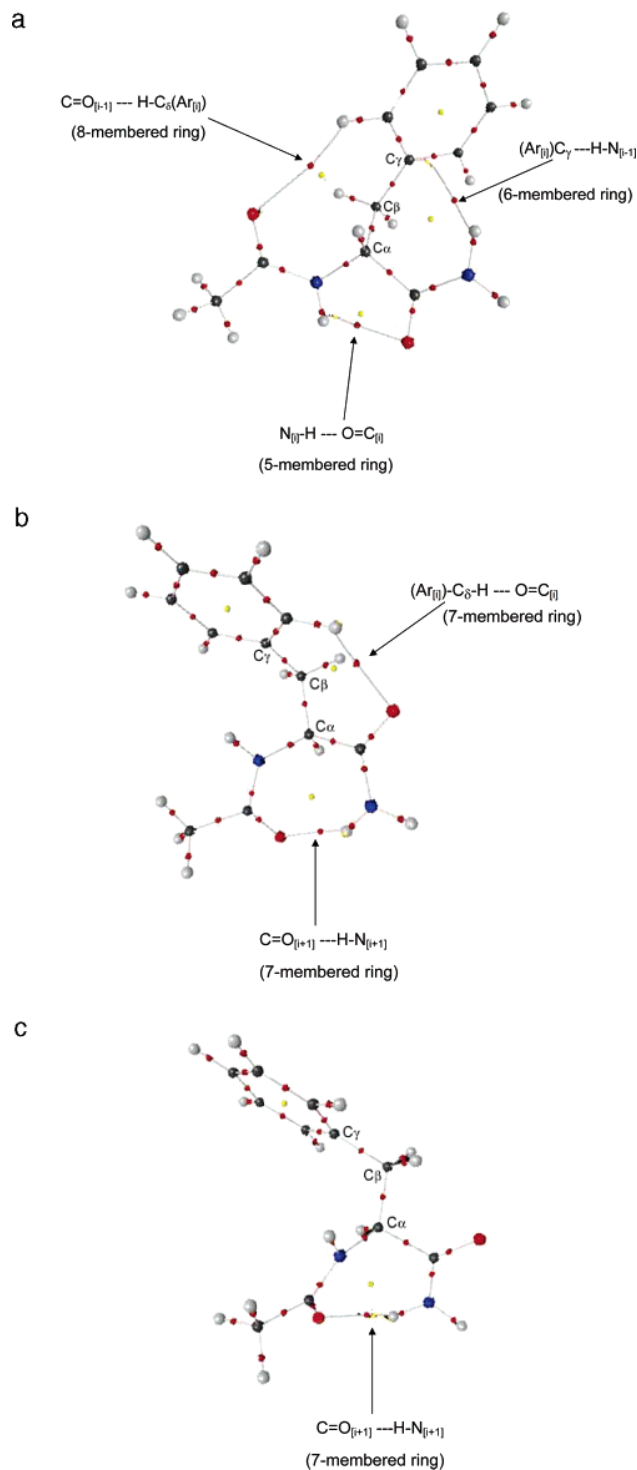


**Figure 6.** Molecular orbital (MO) population surfaces constructed from the electron density matrices emerging from the population analyses of the B3LYP/6-31G(d,p) (*left side*) and B3LYP/6-31+G(d) (*right side*) geometry optimized conformers of NAPA. Surfaces were constructed using a cutoff value of 0.0098. The  $\beta_L^{\alpha}$  conformer (*top*) shows orbital overlap and density for the  $C=O_{[i-1]}-H(N)_{[i]}$  interaction; the  $(Ar)_{[i]}C_{\gamma}-H(N)_{[i+1]}$  interaction shows overlap only in the population analysis of the 6-31G(d,p) optimized geometry (*top left*) but not in the 6-31+G(d) optimized geometry (*top right*). The  $\gamma_L^{s+}$  conformer (*middle*) shows orbital overlap and density between the following pairs of nuclei:  $C=O_{[i-1]}-H(N)_{[i+1]}$ ,  $C=O_{[i]}-H-C_{\delta}(Ar)_{[i+1]}$ . The  $\gamma_L^{s-}$  conformer (*bottom*) shows orbital overlap and density between the following pairs of atoms:  $C=O_{[i-1]}-H(N)_{[i+1]}$ ,  $C=O_{[i]}-H-C_{\beta}$ . All interactions are identified by an ellipse with their corresponding labels.

conformational searches carried out combined with exhaustive and continual refinements of experimental techniques and procedures. Theoretical precision was achieved through the use of an established, modular, explicit, and numeric methodology<sup>3,4</sup> and shows all expected trends for transferability to other model peptide systems.

Most promising for the numerical technique is its ability to predict, locate, and help to optimize transition state structures, between stable backbone and side chain conformers. Interconformational transition states are currently under investigation and are expected to help in the characterization of the kinetics that these systems undergo at higher temperatures.

The conformational searches for stable minima were successful at both characterizing the conformational probability of the NAPA model as well as providing insight into the electrostatic origins of stability, specifically, characterization of the aromatic-amide ( $Ar-H-N$ ) interaction as involving the  $C_{\gamma}$  of the phenylalanyl-side chain aromatic ring, in agreement with the literature<sup>18</sup> and experimental values for the H-N stretch. This is in addition to the novel characterization of the  $(Ar)_{[i]}C_{\delta}-H-N=O_{[i+1]}$  and  $N_{[i]}-H-N_{[i+1]}$  interactions. The combined use of conformational, orbital population, and AIM



**Figure 7.** Atoms-In-Molecules (AIM) analysis of the bonding in NAPA emerging from the wave function generated from the geometry optimized  $\beta_L^{\alpha}$ ,  $\gamma_L^{s+}$ , and  $\gamma_L^{s-}$  conformers (parts a–c, respectively), at the B3LYP/6-31G(d,p) level of theory. Bond-Critical-Points (BCP), Ring-Critical-Points (RCP), and Cage-Critical-Points (CCP) are shown, with  $\nabla\rho\bar{e} = 0$  (No flux in electron density) between two nuclei defining a bond. BCPs are observed between the following sets of atoms for each conformer, providing further evidence for the following intramolecular interactions:  $\beta_L^{\alpha} \rightarrow C=O_{[i-1]}-H(N)_{[i]}$ ,  $C=O_{[i-1]}-H-C_{\delta}(Ar)_{[i]}$ ,  $(Ar)_{[i]}C_{\gamma}-H(N)_{[i+1]}$ ,  $\gamma_L^{s+} \rightarrow C=O_{[i-1]}-H(N)_{[i]}$ ,  $C=O_{[i-1]}-H-C_{\delta}(Ar)_{[i]}$ ,  $\gamma_L^{s-} \rightarrow C=O_{[i]}-H(N)_{[i]}$ .

analyses allowed for efficient—yet incomplete—characterization of the interactions.

**TABLE 4: Frequency and Corresponding Peak Intensities for N<sub>[i]</sub>-H and N<sub>[i+1]</sub>-H Stretches for Stable Backbone and Side Chain Conformations of N-Ac-Phe-NH<sub>2</sub> Geometry Optimized at the B3LYP/6-31G+(d) and BLYP/6-311G(df,p) Levels of Theory**

BB	$\chi^1$	$\chi^2$	N <sub>[i+1]</sub> -H symmetric stretch frequency (cm <sup>-1</sup> )		corresponding IR intensity (km/mol)		N <sub>[i]</sub> -H symmetric stretch frequency (cm <sup>-1</sup> )		corresponding IR intensity (km/mol)		N <sub>[i+1]</sub> -H antisymmetric stretch frequency (cm <sup>-1</sup> )		corresponding IR intensity (km/mol)	
			B3LYP/6-31+G	BLYP/6-311G	B3LYP/6-31+G	BLYP/6-311G	B3LYP/6-31+G	BLYP/6-311G	B3LYP/6-31+G	BLYP/6-311G	B3LYP/6-31+G	BLYP/6-311G	B3LYP/6-31+G	BLYP/6-311G
			(d)	(df,p)	(d)	(df,p)	(d)	(df,p)	(d)	(df,p)	(d)	(df,p)	(d)	(df,p)
$\beta$	+	+	3577.9270	3469.1707	54.1110	70.1378	3582.7820	3486.4606	67.9710	28.6121	3702.1570	3614.1901	38.3670	26.5730
$\beta$	a	+	3570.5100	3458.1358	49.1070	82.1055	3574.8430	3471.6650	79.5710	29.5084	3694.3950	3598.0702	62.9030	51.0813
$\beta$	-	+	N/F <sup>a</sup>	3478.3780	N/F <sup>a</sup>	23.2372	N/F <sup>a</sup>	3489.1229	N/F <sup>a</sup>	34.2192	N/F <sup>a</sup>	3607.7910	N/F <sup>a</sup>	21.8266
$\gamma_L$	+	+	3481.0720	3353.6832	134.1160	135.9209	3586.6540	3486.8953	95.7330	83.9932	3663.3300	3560.5681	81.3590	53.6333
$\gamma_L$	a	+	3513.3110	3388.5335	84.1020	90.7427	3609.4880	3511.1472	24.6110	16.7312	3658.6850	3554.2049	65.5130	47.2269
$\gamma_L$	-	+	3486.3340	3364.3381	126.8380	122.6115	3615.2220	3517.0455	28.7990	22.1489	3657.1000	3552.9638	75.0850	47.9273
$\gamma_D$	+	+	3417.5970	3264.7907	266.5220	307.4132	3616.6430	3506.8267	23.1400	14.5113	3639.2550	3538.9353	83.4360	52.8538
$\gamma_D$	a	+	3477.9060	3352.3669	158.0650	158.2938	3623.6600	3525.6895	27.0860	18.6639	3653.4230	3550.9215	77.2090	51.3970
$\gamma_D$	-	+	3475.8900	3346.8369	165.8500	165.4869	3629.3260	3529.8588	26.1650	18.4935	3663.5760	3558.0135	92.9980	58.7556
$\delta_L$	+	+	3583.6080	3488.3733	41.6690	18.7995	3587.2150	3489.4493	38.1900	30.5442	3715.2250	3619.8833	58.2830	42.3169
$\delta_L$	-	+	3585.2910	3487.0192	31.1950	16.3998	3593.9990	3499.1132	16.7540	9.5762	3712.3050	3616.4718	61.3700	42.2104
$\delta_D$	+	+	3538.4830	3444.7415	30.1070	15.3795	3610.1300	3506.5499	33.0880	23.4875	3652.0990	3561.9769	44.0810	27.5724
$\delta_D$	a	+	3560.4220	3466.3187	26.6230	13.6282	3617.3670	3519.1395	31.0500	20.6570	3678.9560	3588.5221	31.1770	20.2206
$\alpha_L$	a	+	3571.3510	N/F <sup>a</sup>	28.9140	N/F <sup>a</sup>	3594.4220	N/F <sup>a</sup>	18.5030	N/F <sup>a</sup>	3691.6420	N/F <sup>a</sup>	46.4330	N/F <sup>a</sup>
$\alpha_D$	+	+	3569.6400	3474.7956	26.3530	14.7373	3603.7480	3503.4283	14.8290	9.0077	3684.6580	3591.0447	35.5590	24.2336
$\alpha_D$	a	+	3578.8920	3480.8619	29.9250	15.6998	3609.0520	3502.3932	17.8760	8.9525	3700.0040	3603.4496	45.1670	31.2483
$\alpha_D$	-	+	N/F <sup>a</sup>	3485.9118	N/F <sup>a</sup>	14.4586	N/F <sup>a</sup>	3505.2101	N/F <sup>a</sup>	10.3340	N/F <sup>a</sup>	3610.7527	N/F <sup>a</sup>	34.3217
$\epsilon_D$	a	+	3568.4340	3461.9913	75.8570	62.8132	3601.2010	3500.3882	19.9700	11.6278	3694.3010	3591.6445	128.1230	93.0635

<sup>a</sup> N/F indicates conformer not found at the level of theory.**TABLE 5: Frequency and Corresponding Peak Intensities for C=O<sub>[i-1]</sub> and C=O<sub>[i]</sub> Stretches, for Stable Backbone and Side Chain Conformations of N-Ac-Phe-NH<sub>2</sub> Geometry Optimized at the B3LYP/6-31G+(d) and BLYP/6-311G(df,p) Levels of Theory**

BB	$\chi^1$	$\chi^2$	C=O <sub>[i-1]</sub> stretch frequency (cm <sup>-1</sup> )		corresponding IR intensity (km/mol)		C=O <sub>[i]</sub> stretch frequency (cm <sup>-1</sup> )		corresponding IR intensity (km/mol)	
			B3LYP/6-31+G(d)	BLYP/6-311G(df,p)	B3LYP/6-31+G(d)	BLYP/6-311G(df,p)	B3LYP/6-31+G(d)	BLYP/6-311G(df,p)	B3LYP/6-31+G(d)	BLYP/6-311G(df,p)
			$\beta$	+	+	1731.66302	1665.0533	359.9030	270.6661	1769.527
$\beta$	a	+	1729.5300	1660.699	357.5110	256.6636	1760.0930	1693.6116	267.7880	208.0635
$\beta$	-	+	N/F <sup>a</sup>	1683.7046	N/F <sup>a</sup>	302.0460	N/F <sup>a</sup>	1700.4956	N/F <sup>a</sup>	118.6508
$\gamma_L$	+	+	1719.0870	1647.7092	210.0160	139.3216	1759.9200	1694.0799	408.2510	325.1595
$\gamma_L$	a	+	1724.1610	1651.9727	183.3480	125.9077	1770.0250	1700.8955	378.7280	302.9082
$\gamma_L$	-	+	1721.1080	1649.8019	188.0930	130.8634	1769.9580	1706.4150	414.9870	310.5188
$\gamma_D$	+	+	1742.9140	1673.6443	112.1490	78.8194	1753.5560	1692.7594	425.8650	321.0451
$\gamma_D$	a	+	1727.0960	1653.4694	135.6340	94.4640	1767.7660	1708.0507	410.0870	297.5544
$\gamma_D$	-	+	1724.3460	1651.0624	128.3170	91.5684	1758.8310	1702.5270	459.9810	329.5387
$\delta_L$	+	+	1757.6780	1692.6629	305.0980	218.2842	1761.9610	1694.2007	348.7500	279.2317
$\delta_L$	-	+	1763.4590	1697.4515	267.3380	174.3204	1765.5880	1701.0380	343.3440	280.6550
$\delta_D$	+	+	1750.3430	1680.6280	238.5280	131.8623	1783.3430	1715.5527	363.9710	249.3254
$\delta_D$	a	+	1748.7270	1679.6394	229.9410	170.2479	1790.5300	1721.2942	298.6180	233.3680
$\alpha_L$	a	+	1765.0800	N/F <sup>a</sup>	231.7410	N/F <sup>a</sup>	1783.3220	N/F <sup>a</sup>	336.9710	N/F <sup>a</sup>
$\alpha_D$	+	+	1766.5990	1696.5259	210.45503	131.862	1774.4290	1706.2830	275.8570	249.3254
$\alpha_D$	a	+	1761.331	1693.0211	193.8800	129.3567	1776.0030	1708.4913	336.1940	279.0296
$\alpha_D$	-	+	N/F <sup>a</sup>	1692.0334	N/F <sup>a</sup>	146.4486	N/F <sup>a</sup>	1708.9333	N/F <sup>a</sup>	269.8098
$\epsilon_D$	a	+	1756.8300	1693.2256	454.4770	282.1142	1768.8960	1706.8686	147.5480	173.484

<sup>a</sup> N/F indicates conformer not found at the level of theory.

The long-standing and ongoing debate about which method and basis set is best used (in order to achieve the minimal accuracy required to properly characterize the model peptide systems) has not been brought any closer to being satisfied. An excellent qualitative correlation with experimental results was achieved using the B3LYP method, employing the 6-31G(d), 6-31G(d,p), and 6-31+G(d) levels of theory, this is reflected at the BLYP/6-311G(df,p) level. Quantitative ‘certainty’ would require, by definition, an infinite basis set.

Future and continued work on the complete characterization of NAPA would benefit from an ‘automated iterative process’ of steps, until theoretical refinement suffers from a ‘law of

diminishing returns’; instead of where accuracy would not be sufficiently ameliorated to justify the computational resources expended. Clearly this is an ‘era-dependent’ factor, where computational resources are continually evolving, facilitating today that which was computationally unattainable yesterday.

A second ‘offshoot’ of this work may include the evaluation of conformer- and computational level-dependent force constants, associated with each degree of freedom. The level of theory generating results most accurately and precisely in agreement with experimental values may be used to reparameterize existing force-fields for peptide systems. Once the complete set of N- and C-protected amino acids have been so



- (30) Armstrong, K. M.; Fairman, R.; Baldwin, R. L. *J. Mol. Biol.* **1993**, *230*, 284–291.
- (31) Chakrabarti, P.; Uttamkumar, S. *J. Mol. Biol.* **1995**, *251*, 9–14.
- (32) Fan M-F.; Lin, Z.; McGrady, J. E.; Mingos, D. M. P. *J. Chem. Soc.* **1996**, Perkin 2, 563.
- (33) Chipot, C.; Maignet, B.; Pearlman, D. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1996**, *118*, 2998–3005.
- (34) McGaughey, G. B.; Gagnes, M.; Rappe A. K.  $\pi$ -Stacking Interactions: Alive and Well in Proteins **1998**, *273*, No. 25, issue of June 19, 15458–15463.
- (35) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. *J. Am. Chem. Soc.* **2000**, *122*, 3746–3753.
- (36) Toth, G.; Murphy, R. F.; Lovas, S. *J. Am. Chem. Soc.* **2001**, *123*, 11782–11790.
- (37) Toth, G.; Murphy, R. F.; Lovas, S. *Protein Eng.* **2001**, *14*, no. 8, 543–547.
- (38) Steiner, T.; Koellner, G. *J. Mol. Biol.* **2001**, *305*, 535–557.
- (39) Shi, Z.; Olson, C. A.; Kallenbach, N. R. *J. Am. Chem. Soc.* **2002**, *124*, 3284–3291.
- (40) Chass, G. A.; Marai, C. N. J.; Harrison, A. G.; Csizmadia, I. G. *J. Phys. Chem. A* **2002**, *106*, 9695–9704.
- (41) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.11; Gaussian, Inc.: Pittsburgh, PA, 1998.
- (42) Chass, G. A. *J. Mol. Struct. (THEOCHEM)* **2003**, *667*–667, 61.
- (43) Tehrani, F. S. *Comput. Phys. Commun.* **2001**, *140*, 58.
- (44) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Theory*; Wiley: New York, 1986.
- (45) Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69.
- (46) Ditchfield, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724.
- (47) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257.
- (48) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.
- (49) Perczel, A.; Angyan, J. G.; Kajtar, M.; Viviani, W.; Rivail, J. L.; Marcoccia, J. F.; Csizmadia, I. G. *J. Am. Chem. Soc.* **1991**, *113*, 6256.
- (50) Almeida, D. R. P.; Pisterzi, L. F.; Chass, G. A.; Torday, L. L.; Varro, Papp, J. Gy.; Csizmadia, I. G. *J. Phys. Chem. A* **2002**, *106* (43), 10423–10436.
- (51) Pisterzi, L. F.; Almeida, D. R. P.; Chass, G. A.; Torday, L. L.; Varro, Papp, J. Gy.; Csizmadia, I. G. *Chem. Phys. Lett.* **2002**, *365*, 542–551.
- (52) Bombasaro, J. A.; Suvire, F. D.; Chasse, G. A.; Zamarbide, G. N.; Estrada, M. R. *J. Mol. Struct. (THEOCHEM)* **2001**, *548*, 39–46.
- (53) Villagra, S. E.; Santillan, M. B.; Rodriguez, A. M.; Chasse, G. A.; Freile, M. L.; Zacchino, S.; Enriz, R. D.; Matyus, P. *J. Mol. Struct. (THEOCHEM)* **2001**, *549*, 217–228.
- (54) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *Chem. Phys. Lett.* **1988**, *153*, 503.
- (55) Scoot, A. P.; Radom, L. *J. Phys. Chem.* **1996**, *100*, 16502.
- (56) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
- (57) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- (58) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
- (59) Scott, A. P.; Radom, L. *J. Phys. Chem.* **1996**, *100*, 16502.
- (60) Bader, R. F. W. *Atoms in Molecules. A Quantum Theory*; Clarendon Press: Oxford, 1990; *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 620.
- (61) Bader, R. E. W.; Biegler-König, F. W.; Cheeseman, J. R.; Duke, J. A.; Keith, T. A.; Laidig, K. E.; Langdon, B. *AIMPAC* (a set of programs for the theory of atoms in molecules); McMaster University: Hamilton, ON, Canada, 1994. Biegler-König, F. W.; Bader, R. F. W.; Tang, T.-H. *J. Comput. Chem.* **1982**, *3*, 317–328.
- (62) Fang, D.-C.; Tang, T.-H. *AIM98PC* (a modified PC version of *AIMPAC*); Beijing Normal University: Beijing, China, 1998.
- (63) Biegler-König, F. W.; Schönbohn, J.; Bayles, D. *J. Comput. Chem.* **2002**, *22*, 545–559. Biegler-König, F. W.; Schönbohn, J. *J. Comput. Chem.* **2002**, *23*, 1489–1494.
- (64) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.
- (65) Molekel: Flükiger, P. F. Development of the molecular graphics package MOLEKEL and its application to selected problems in organic and organometallic chemistry, Thèse No 2561, Département de chimie physique, Université de Genève, Genève, 1992.
- (66) Portman, S.; Lüthi, H. P. MOLEKEL: An Interactive Molecular Graphics Tool. *Chimia* **2000**, *54*, 766–770.
- (67) MacDiarmid, A. G. *Rev. Mod. Phys.* **2001**, *73*, 701.
- (68) Ozaki, M.; Takahashi, A. *Chem. Phys. Lett.* **1986**, *127*, 242.
- (69) Newton, M. D.; Stanton, R. E. *J. Am. Chem. Soc.* **1986**, *108*, 2469.
- (70) Lothi, H. P.; Alml, F. *J. Chem. Phys. Lett.* **1987**, *135*, 357.
- (71) Nelson, D.; Perng, T. *Proc. Oklahoma Acad. Sci.* **2000**, *80*, 141.
- (72) Chin, W.; Mons, M.; Dognon, J.-P.; Mirasol, R. S.; Chass, G.; Dimicoli, I.; Piuze, F.; Butz, P.; Tardivel, B.; Compagnon, I.; von Helden, G.; Meijer, G. *J. Phys. Chem A* **2005**, *109*, 5281.





## Annexe C

# Tripeptides

C.1 Énergies d'une sélection de structures calculées pour le tripeptide Ac-Phe-Gly-Gly-NH<sub>2</sub>

C.2 Énergies et paramètres géométriques d'une sélection de structures calculées pour les tripeptides de type -Ala-Ala-Ala-

Ac-Phe-Gly-Gly-NH <sub>2</sub>	$\Delta E$ MP2 + ZPE
<b>coude <math>\beta</math> - coude <math>\beta</math></b>	
C <sub>10</sub> -C <sub>10</sub>	
II(a)-I'	0
II(g)-I'	1,50
hélice 3 <sub>10</sub> (g-)	2.19
hélice 3 <sub>10</sub> (g+)	2.83
hélice 3 <sub>10</sub> (a)	5.65
<b>coude <math>\gamma</math> - coude <math>\beta</math></b>	
C <sub>7</sub> -C <sub>10</sub>	
II(a) - $\gamma_D$	1.61
II(a) - $\gamma_L$	4.05
I(g+) - $\gamma_D$	3.88
I(g+) - $\gamma_L$	4.43
C <sub>10</sub> -C <sub>7</sub>	
$\gamma_L$ (g+) - I	5.19
<b>2 liaisons H</b>	
$\gamma_L$ (g+) - C <sub>13</sub>	3.89
C <sub>7</sub> // C <sub>13</sub>	
$\gamma_L$ (a) - C <sub>11</sub>	7.79
C <sub>7</sub> // C <sub>11</sub>	
$\beta_L$ (a)- $\gamma_L$ - $\gamma_L$	5.31
C <sub>5</sub> -C <sub>7</sub> -C <sub>7</sub>	
<b>3 liaisons H</b>	
$\gamma_L$ (g-)- $\gamma_D$ - $\gamma_D$	5.92
<b>0 liaison H</b>	
$\beta_L$ (a)- $\beta_L$ - $\beta_L$	8.43

TAB. C.1 – Énergies relatives calculées au niveau MP2/6-31+G(d) à partir de géométries B3LYP pour différentes structures du système Ac-Phe-Gly-Gly-NH<sub>2</sub>.

Tripeptide	$\Delta E$ (kcal/mol)	Angles dièdres ( $^{\circ}$ )					
		$\phi_{i+1}$	$\psi_{i+1}$	$\phi_{i+2}$	$\psi_{i+2}$	$\phi_{i+3}$	$\psi_{i+3}$
<b>Ala-Ala-Ala</b>							
$3_{10}$	0.00	-69	-20	-69	-10	-96	2
II'-I	0.12	55	-125	-69	-11	98	5
II-I'	1.66	-60	125	58	28	68	21
$\gamma_L$ - type I	1.85	-83	73	-73	-16	-97	4
$\gamma_L$ - $\gamma_L$ - $\gamma_L$	2.32	-83	73	-85	68	-84	71
<b>Phe-Ala-Ala</b>							
$3_{10}$ (g+)	0.00	-75	-9	-68	-13	-97	3
$3_{10}$ (g-)	0.85	-68	-22	-68	-11	-96	2
$3_{10}$ (a)	3.71	-64	-31	-68	-8	-98	3
II-I' (a)	1.80	-61	125	58	28	68	21
II-I' (g-)	3.94	-59	124	58	26	67	21
$\beta_L$ (a)- $\gamma_L$ - $\gamma_L$	1.92	-156	166	-82	71	-84	71
$\gamma_L$ (g+)- $\gamma_L$ - $\gamma_L$	3.74	-83	57	-84	66	-84	71
$\gamma_L$ (g-)- $\gamma_L$ - $\gamma_L$	4.26	-85	70	-85	67	-84	71
<b>Ala-Phe-Ala</b>							
$3_{10}$ (g+)	0.00	-68	-21	-74	-1	-100	2.5
$3_{10}$ (g-)	3.24	-68	-22	-68	-13	-94	1
II'-I (g+)	1.39	55	-127	-75	-1	-100	3
II'-I (g-)	3.66	55	-125	-69	-12	-97	3
$\gamma_L$ -type I (g+)	3.31	-83	78	-77	-9	-97	2
<b>Ala-Ala-Phe</b>							
$3_{10}$ (g-)	0.00	-69	-19	-68	-12	-95	1
$\gamma_L$ -type I (g+)	0.11	-83	74	-73	-14	-98	8
$\gamma_L$ -type I (g-)	2.94	-83	73	-71	-18	-93	0.4
$\gamma_L$ - $\gamma_L$ - $\gamma_L$ (g-)	2.50	-83	68	-85	63	-85	71

TAB. C.2 – Énergies relatives MP2/6-31+G(d) single point et angles dièdres  $\phi$  et  $\psi$  (B3LYP) de structures  $C_{10}$ - $C_{10}$ ,  $C_7$ - $C_{10}$ ,  $C_7$ - $C_7$ - $C_7$  et  $C_5$ - $C_7$ - $C_7$  pour les tripeptides -Ala-Ala-Ala-.

Tripeptide	Moment dipolaire $\mu$ $\mu$ (Debye)		Distances (Å)				
<b>Ala-Ala-Ala</b>							
$3_{10}$	12.7	2.28	C <sub>10</sub>	2.18	C <sub>10</sub>	2.2	
type II'-I	9.8	2.07	C <sub>10</sub>	2.16	C <sub>10</sub>		
type II-I'	9.9	2.1		2.11			
$\gamma_L$ - type I	9.6	2.02	C <sub>7</sub>	2.12	C <sub>10</sub>		
$\gamma_L$ - $\gamma_L$ - $\gamma_L$	7.3		C <sub>7</sub>		C <sub>7</sub>		C <sub>7</sub>
<b>Phe-Ala-Ala</b>							
$3_{10}$ (g+)	12.5	2.26		2.22			
$3_{10}$ (g-)	13	2.14		2.19			
$3_{10}$ (a)	12.5	2.12		2.15			
type II-I' (a)	9.8	2.115		2.10			
type II-I' (g-)	11.3	2.05		2.11			
$\beta_L$ (a)- $\gamma_L$ - $\gamma_L$	2.7	2.16		2.08		2.055	
$\gamma_L$ (g+)- $\gamma_L$ - $\gamma_L$	7.4	1.97	C <sub>7</sub>	2.03	C <sub>7</sub>	2.045	C <sub>7</sub>
$\gamma_L$ (g-)- $\gamma_L$ - $\gamma_L$	7.6	2.015	C <sub>7</sub>	2.04	C <sub>7</sub>	2.05	C <sub>7</sub>
<b>Ala-Phe-Ala</b>							
$3_{10}$ (g+)	12.1	2.135		2.24		2.635	NH- $\pi$
$3_{10}$ (g-)	12.8	2.20		2.15		2.84	NH- $\pi$
type II'-I (g+)	9.9	2.05	C <sub>10</sub>	2.18	C <sub>10</sub>		
type II'-I (g-)	9.6	2.09		2.12	C <sub>10</sub>		C <sub>10</sub>
$\gamma_L$ -type I (g+)	9.4	2.11	C <sub>7</sub>	2.14	C <sub>10</sub>	2.76	NH- $\pi$
<b>Ala-Ala-Phe</b>							
$3_{10}$ (g-)	12.4	2.19	C <sub>10</sub>	2.21			
$\gamma_L$ -type I (g+)	9.3	2.03		2.11			
$\gamma_L$ -type I (g-)	9.5	2.03		2.12			
$\gamma_L$ - $\gamma_L$ - $\gamma_L$ (g-)	7.0	2.03	C <sub>7</sub>	2.03	C <sub>7</sub>	2.055	C <sub>7</sub>

TAB. C.3 – Moments dipolaires et distances des interactions remarquables calculés pour différentes structures des tripeptides -Ala-Ala-Ala-, avec la méthode B3LYP/6-31+G(d).