

# Tables des matières

## CHAPITRE I : INTRODUCTION GENERALE

Introduction..... - 1 -

I. Eléments de description structurale des peptides ..... - 2 -

    I.A. Acides Aminés (AA)..... - 2 -

    I.B. Peptides ..... - 2 -

        I.B.1. Rappel : structure des peptides et des protéines ..... - 2 -

        I.B.2. Le problème conformationnel ..... - 4 -

        I.B.3. Structures des peptides ..... - 7 -

II. Le cation Na<sup>+</sup> en biochimie ..... - 8 -

    II.A. Na<sup>+</sup>, élément essentiel pour les êtres vivants ..... - 8 -

        II.A.1. Rôle fonctionnel du sodium : canaux à sodium..... - 8 -

        II.A.2. Rôle structurant du sodium ..... - 9 -

    II.B. Na<sup>+</sup>, élément souvent rencontré en biochimie en phase gazeuse ..... - 10 -

III. Chimie en phase gazeuse ..... - 11 -

    III.A. Méthodes de spectrométrie de masse..... - 11 -

    III.B. Intérêt de la chimie en phase gazeuse..... - 12 -

        III.B.1. Propriétés intrinsèques ..... - 12 -

        III.B.2. Conformation d'un peptide au sein d'une protéine..... - 13 -

        III.B.3. La micro-solvatation ..... - 13 -

    III.C. Méthodes de chimie théorique ..... - 15 -

    III.D. Conclusions ..... - 15 -

IV. Notre étude ..... - 15 -

Bibliographie..... - 17 -

## **CHAPITRE II : PRINCIPES EXPERIMENTAUX DE CARACTERISATION STRUCTURALE DE PETITS IONS BIOLOGIQUES EN PHASE GAZEUSE**

**Introduction..... - 20 -**

**I. Résumé des développements expérimentaux actuels..... - 21 -**

**I.A. Echange H/D..... - 21 -**

- I.A.1. Principe ..... - 21 -
- I.A.2. Un exemple d'échange H/D en solution ..... - 22 -
- I.A.3. Quelques résultats d'échange H/D en phase gazeuse ..... - 23 -
- I.A.4. Conclusions ..... - 24 -

**I.B. Technique de mobilité ionique ..... - 24 -**

- I.B.1. Principe ..... - 24 -
- I.B.2. Application aux cas d'oligomères d'acides aminés ..... - 25 -

**I.C. Techniques de détermination d'énergie de liaison ..... - 26 -**

- I.C.1. La méthode d'équilibre ..... - 26 -
  - I.C.2. La méthode cinétique de Cooks ..... - 27 -
  - I.C.3. Autre exemple : dissociation par radiation infrarouge du corps noir (Blackbody Infrared Radiative Dissociation : BIRD) ..... - 29 -
  - I.C.4. Mesure d'énergie au seuil de dissociation : méthode TCID ..... - 31 -
- I.D. Conclusion ..... - 32 -**

**II. Etude structurale par spectroscopie infrarouge en phase gazeuse ..... - 32 -**

**II.A. Introduction : spectres IR de molécules biologiques ..... - 32 -**

**II.B. Spectres IR de neutres en matrice inerte ..... - 33 -**

**II.C. Spectres IR de neutres en phase gazeuse ..... - 34 -**

**II.D. Spectres IR d'ions en phase gazeuse ..... - 34 -**

- II.D.1. Introduction ..... - 34 -

- II.D.2. Principe de l'IRMPD ..... - 35 -

- II.D.3. Tracé d'un spectre IRMPD ..... - 37 -

- II.D.4. Interprétation d'un spectre IRMPD : comparaisons aux spectres IR théoriques - 38 -

II.D.5. Quelques résultats .....	- 39 -
II.D.6. Conclusion .....	- 40 -
<b><u>Bibliographie</u></b> .....	<b>- 41 -</b>

## **CHAPITRE III : METHODES EXPERIMENTALES ET THEORIQUES**

<b><u>Introduction</u></b> .....	<b>- 45 -</b>
----------------------------------	---------------

<b><u>I. Le Laser à électrons libres CLIO</u></b> .....	<b>- 46 -</b>
---	---------------

I.A. Introduction .....	- 46 -
I.B. Principe du LEL.....	- 46 -
I.C. Structure temporelle du LEL.....	- 47 -
I.D. Principales caractéristiques du LEL CLIO .....	- 48 -

<b><u>II. Le spectromètre de masse MICRA</u></b> .....	<b>- 48 -</b>
--	---------------

II.A. Introduction .....	- 48 -
<b>II.B. Bref rappel sur la spectrométrie de masse FT-ICR</b> .....	- 48 -
II.B.1. Mouvement cyclotron .....	- 48 -
II.B.2. Principe de la FTICR.....	- 49 -
II.B.3. Optimisation de la détection et résolution .....	- 51 -
II.B.4. Cellules .....	- 52 -
<b>II.C. Originalités du montage MICRA en 2002:</b> .....	- 52 -
II.C.1. Montage expérimental .....	- 52 -
II.C.2. Formation des ions et introduction des espèces neutres.....	- 53 -
<b>II.D. Développement apporté lors de ces travaux</b> .....	- 54 -
II.D.1. Introduction .....	- 54 -
II.D.2. Aménagement de MICRA : le MALDI .....	- 54 -
II.D.3. Les échantillons MALDI .....	- 55 -

### **III. Méthodes théoriques..... - 59 -**

<b>III.A. Introduction.....</b>	<b>- 59 -</b>
III.A.1. Introduction .....	- 59 -
III.A.2. Stratégies selon la taille et la nature des systèmes .....	- 60 -
<b>III.B. Le champ de force AMBER.....</b>	<b>- 60 -</b>
III.B.1. Les champs de forces .....	- 60 -
III.B.2. AMBER .....	- 61 -
<b>III.C. Méthodes de chimie quantique .....</b>	<b>- 62 -</b>
III.C.1. Introduction .....	- 62 -
III.C.2. Méthodes Hartree-Fock (HF) et post-Hartree-Fock .....	- 63 -
III.C.3. Méthodes de la DFT .....	- 64 -
<b>III.D. La méthode ri-BLYP : principe et calibration.....</b>	<b>- 65 -</b>
III.D.1. Principe de l'approximation ri.....	- 66 -
III.D.2. Calibration de ri-BLYP pour le calcul d'énergies.....	- 67 -
III.D.3. Calibration de ri-BLYP pour le calcul des géométries.....	- 69 -
III.D.4. Calibration de ri-BLYP pour le calcul de fréquences de vibration .....	- 70 -
III.D.5. Conclusion .....	- 71 -

### **Bibliographie..... - 73 -**

## **CHAPITRE IV : ROLE STRUCTURANT DE Na<sup>+</sup> SUR LES ACIDES AMINES GLYCINE ET PROLINE**

### **I. Introduction..... - 75 -**

<b>I.A. Structures des acides aminés isolés .....</b>	<b>- 75 -</b>
I.A.1. Acides aminés isolés en solution aqueuse.....	- 75 -
I.A.2. Acides aminés isolés en phase gazeuse .....	- 75 -
<b>I.B. Stabilisation du zwitterion en phase gazeuse.....</b>	<b>- 78 -</b>
I.B.1. Microsolvatation .....	- 78 -
I.B.2. Complexation par un alcalin.....	- 78 -

I.B.3. Cas de Gly-Na <sup>+</sup> .....	- 79 -
I.B.4. Cas de Pro-Na <sup>+</sup> .....	- 80 -
<b>I.C. Conclusion .....</b>	<b>- 81 -</b>

**II. Spectroscopie infrarouge des complexes Gly-Na<sup>+</sup> et Pro-Na<sup>+</sup> ..... - 81 -**

<b>II.A. Résumé .....</b>	<b>- 81 -</b>
<b>II.B. Article.....</b>	<b>- 82 -</b>

**Bibliographie..... - 90 -**

**CHAPITRE V : ROLE STRUCTURANT DE Na<sup>+</sup> SUR DES PETITS PEPTIDES : DU DI- AU TETRA-PEPTIDE**

**Introduction..... - 94 -**

**I. Structure des complexes Na<sup>+</sup>-Gly<sub>n</sub> et Na<sup>+</sup>-Ala<sub>n</sub> (n≤4)..... - 95 -**

<b>I.A. Conformères des complexes Gly<sub>n</sub>-Na<sup>+</sup> et Ala<sub>n</sub>-Na<sup>+</sup> (n&lt;5) .....</b>	<b>- 95 -</b>
I.A.1. Gly <sub>2</sub> -Na <sup>+</sup> et Ala <sub>2</sub> -Na <sup>+</sup> .....	- 95 -
I.A.2. Gly <sub>3</sub> -Na <sup>+</sup> et Ala <sub>3</sub> -Na <sup>+</sup> .....	- 98 -
I.A.3. Gly <sub>4</sub> -Na <sup>+</sup> et Ala <sub>4</sub> -Na <sup>+</sup> .....	- 102 -
I.A.4. Conclusion .....	- 106 -

**I.B. Spectres IRMPD ..... - 106 -**

I.B.1. Echantillons MALDI .....	- 107 -
I.B.2. Spectroscopie IRMPD de Ala <sub>2</sub> -Na <sup>+</sup> et Gly <sub>2</sub> -Na <sup>+</sup> .....	- 107 -
I.B.3. Spectroscopie IRMPD de Gly <sub>3</sub> -Na <sup>+</sup> et Ala <sub>3</sub> -Na <sup>+</sup> .....	- 112 -
I.B.4. Conclusion .....	- 119 -

**II. Energies de liaison peptide-Na<sup>+</sup> ..... - 120 -**

**II.A. Echelle expérimentale d'affinités de peptides pour le sodium ..... - 120 -**

<b>II.B. Détermination théorique d'affinités de peptides pour le sodium.....</b>	<b>- 122 -</b>
II.B.1. Rappel sur la détermination d'énergie de liaison .....	122 -
II.B.2. Conformères de plus basse énergie de Gly <sub>2</sub> et Ala <sub>2</sub> .....	122 -
II.B.3. Conformères de plus basse énergie de Gly <sub>3</sub> et Ala <sub>3</sub> .....	123 -
II.B.4. Conformères de plus basse énergie de Gly <sub>4</sub> et Ala <sub>4</sub> .....	125 -
<b>II.C. Comparaison expériences-théorie .....</b>	<b>- 129 -</b>
<b>III. Cas des tautomères GlyHis et HisGly sodiés et isolés .....</b>	<b>- 131 -</b>
<b>III.A. Résumé .....</b>	<b>- 131 -</b>
<b>III.B. Article.....</b>	<b>- 131 -</b>
<b>Bibliographie.....</b>	<b>- 159 -</b>

## **CHAPITRE VI : ROLE STRUCTURANT DE Na<sup>+</sup> SUR LES HOMO-PEPTIDES Gly<sub>n</sub>-Na<sup>+</sup>, 5≤n≤10**

<b>Introduction.....</b>	<b>- 161 -</b>
<b>I. Propension à former des hélices <math>\alpha</math> .....</b>	<b>- 162 -</b>
<b>I.A. Propension à former des hélices <math>\alpha</math> .....</b>	<b>- 162 -</b>
I.A.1. Selon la nature des acides aminés .....	162 -
I.A.2. Augmenter la propension à former des hélices $\alpha$ .....	163 -
<b>I.B. Etudes de peptides dérivés de Ala<sub>n</sub> par l'équipe de M. Jarrold.....</b>	<b>- 164 -</b>
<b>I.C. Notre étude .....</b>	<b>- 165 -</b>
<b>II. Les différentes structures en hélices envisagées .....</b>	<b>- 165 -</b>
<b>II.A. Géométries .....</b>	<b>- 165 -</b>
<b>II.B. Energies .....</b>	<b>- 166 -</b>

<b><u>III. Les différentes structures globulaires.....</u></b>	<b>- 169 -</b>
<b>III.A. Tendances énergétiques générales .....</b>	<b>169</b>
<b>III.B. Géométries .....</b>	<b>170</b>
III.B.1. Descripteurs .....	170
III.B.2. Géométrie des structures les plus basses en énergie .....	171
III.B.3. Coordination au métal .....	183
III.B.4. Liaisons hydrogène .....	184
III.B.5. Conclusion .....	184
<b>III.C. Energies relatives des conformères globulaires.....</b>	<b>185</b>
<b><u>IV. Conclusion et perspectives.....</u></b>	<b>- 187 -</b>
<b><u>V. Annexe 1 : Calibration des méthodes de chimie quantique sur Gly<sub>6</sub>-Na<sup>+</sup> -</u></b>	<b>188 -</b>
<b>V.A. Méthodes et Bases de calculs employés.....</b>	<b>188 -</b>
<b>V.B. Energies .....</b>	<b>189</b>
V.B.1. Influence de la base.....	189
V.B.2. Influence de la méthode .....	190
<b><u>Bibliographie.....</u></b>	<b>- 192 -</b>
<b><u>CONCLUSIONS ET PERSPECTIVES.....</u></b>	<b>- 194 -</b>

## **Chapitre I :**

### **Introduction générale**

## Introduction

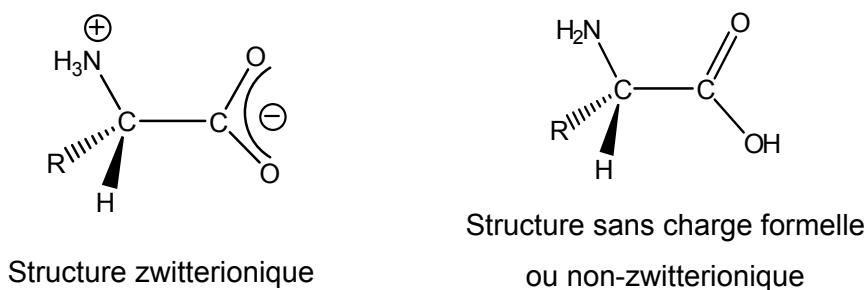
Cette thèse a pour but l'étude du rôle structurant du cation  $\text{Na}^+$  sur des molécules d'intérêt biologique, en phase gazeuse. Elle se trouve donc à l'interface entre la physico-chimie de par les méthodes d'analyse utilisées, et la biochimie en ce qui concerne les systèmes étudiés.

Ce chapitre a pour but d'une part de donner quelques éléments de biochimie indispensables à la compréhension des résultats qui vont suivre et d'autre part de situer la problématique de ce travail.

## I. Eléments de description structurale des peptides

### I.A. Acides Aminés (AA)

Les acides aminés sont les constituants principaux des protéines. De manière générale, ils ont pour formule  $\text{NH}_2\text{-CHR-COOH}$ , R désignant la chaîne latérale caractérisant les acides aminés. Cette chaîne peut comporter diverses fonctions comme un groupe alkyle (alanine, leucine, valine...), une fonction acide (acide aspartique ou glutamique), alcool (sérine ou thréonine), thiol (cystéine, méthionine), etc, le plus petit acide aminé étant la glycine qui ne comporte pas de chaîne latérale (R=H). Cependant en solution aqueuse, les acides aminés ne se présentent jamais sous cette forme, leur structure dépend du pH de la solution, et des fonctions de la chaîne latérale. Lorsqu'ils sont sous forme globalement neutre, ils possèdent en fait une fonction carboxylate et une fonction ammonium, on parle alors de forme zwitterionique. Au contraire lorsqu'ils sont sous forme gazeuse, les acides aminés ne présentent que des structures sans charge formelle. Ce comportement a été établi à la fois expérimentalement et théoriquement.



**Figure I-1 :** Isomérie des acides aminés : structure zwitterionique en solution et non-zwitterionique en phase gazeuse.

### I.B. Peptides

#### I.B.1. Rappel : structure des peptides et des protéines

Les protéines peuvent être formellement considérées comme des polymères obtenus par la condensation d'acides aminés, créant une fonction amide, dite fonction peptidique entre deux résidus. La liaison peptidique étant plane et généralement trans, seuls les angles dièdres  $\text{N-C}_\alpha$  (habituellement noté  $\Phi$ ) et  $\text{C}_\alpha\text{-CO}$  ( $\Psi$ ) seront les variables conformationnelles pour la chaîne principale (cf. Figure I-2).

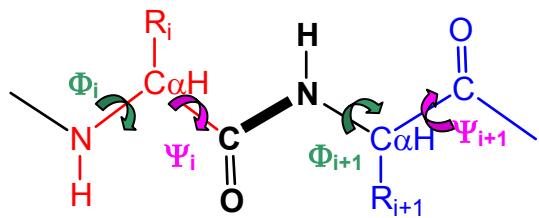


Figure I-2 : Les degrés de liberté de la chaîne principale d'un polypeptide.

On comprend aisément que plus le nombre de résidus augmente, plus le nombre de conformations accessibles devient important. Cependant toutes les valeurs des couples ( $\Phi$ ,  $\Psi$ ) ne sont pas permises stériquement. Le diagramme de Ramachandran représente l'espace conformationnel accessible. Les diagrammes de Ramachandran (Figure I-3) sont semblables pour tous les acides aminés (partie verte), sauf pour la glycine qui est beaucoup plus flexible (parties verte et saumon).

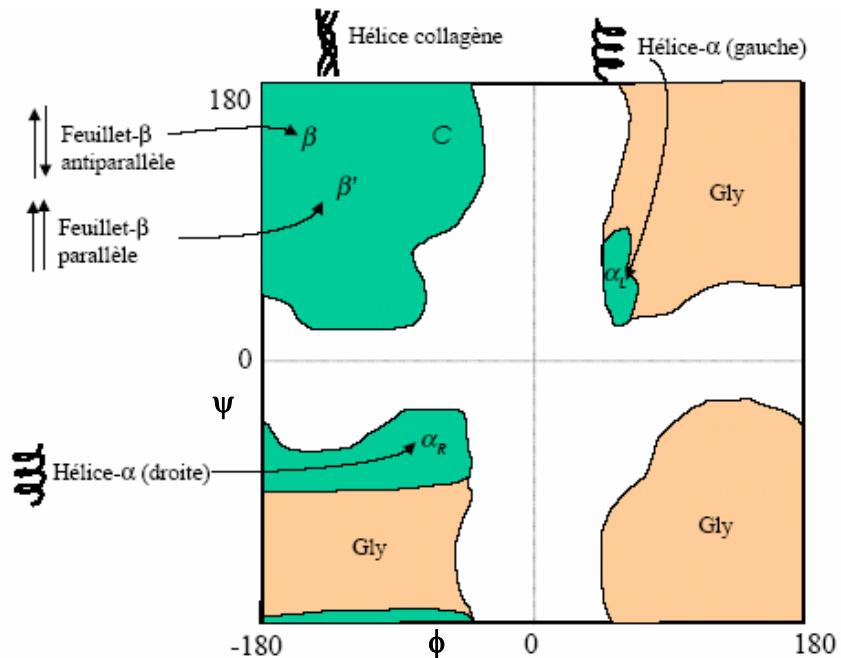


Figure I-3 : Carte de Ramachandran pour un polypeptide.

On distingue quatre niveaux de structure pour une protéine : structure primaire, secondaire, tertiaire et quaternaire. La structure primaire correspond à l'identification de la séquence des acides aminés composant la protéine. La structure secondaire correspond aux conformations locales du squelette protéique, dont les plus connues sont l'hélice  $\alpha$  et le feuillet  $\beta$ . L'agencement tridimensionnel des structures secondaires, impliquant des

interactions entre résidus distants dans la séquence, constitue la structure tertiaire. Enfin lorsque la protéine comporte plusieurs sous-unités, le positionnement relatif des différentes chaînes polypeptidiques définit la structure quaternaire.

### I.B.2. Le problème conformationnel <sup>1, 2</sup>

Dans le cas de petits peptides, les problèmes conformationnels restent les mêmes, les couples ( $\Phi, \Psi$ ) sont à déterminer pour chaque résidu. Cependant, la plupart des peptides courts sont **non structurés**. En effet, les différentes conformations possibles ont des énergies relativement proches (quelques  $\text{kJ.mol}^{-1}$  de différence), si bien que ces peptides adoptent différentes combinaisons de conformations. Le passage d'une conformation à une autre est généralement rapide (temps de vie de l'ordre de 1 ns) pour les angles  $\Phi$ , tandis qu'elle est beaucoup plus lente pour les liaisons amides.

Cependant certains petits peptides sont structurés. Ils peuvent alors adopter deux grands types de structures secondaires :

- d'une part, des structures à motif périodique, tels que les hélices. Des différences importantes existent avec les structures des protéines, puisque par exemple les peptides ne forment pas de feuillet  $\beta$ . Il existe cependant un équivalent qui est l'épingle à cheveux  $\beta$  ou  $\beta$ -hairpin.
- d'autre part, des conformations qui ne font pas apparaître de périodicité dans la structure tridimensionnelle, on parlera alors de **structure globulaire**.

Un résumé rapide des structures secondaires que l'on peut rencontrer dans le cas des petits peptides est présenté ci-dessous.

#### I.B.2.a. Hélices

Les peptides présentent les mêmes types d'hélices que les protéines. On distingue 3 grandes classes d'hélices, selon le nombre d'unités peptidiques par tour d'hélice (noté  $n$ ) et le pas de l'hélice (noté  $p$ ). Toutes ces hélices présentent une chiralité, mais seule la configuration droite est observée (en effet la configuration gauche impose un contact trop étroit entre les chaînes latérales et la chaîne peptidique). Leurs principales caractéristiques sont présentées dans le tableau I-1 et sur la figure I-4.

##### L'hélice $\alpha$

C'est de loin la plus fréquente. Comme toutes les liaisons peptidiques et les liaisons hydrogène sont parallèles, l'hélice  $\alpha$  est très polaire. Son moment dipolaire correspond approximativement à 3.5 D par résidu (la charge positive est du côté de l'extrémité N-

terminale, la charge négative du côté C-terminal). Les fragments de protéines en hélice  $\alpha$  comportent, en moyenne, 12 résidus.

### L'hélice $3_{10}$

L'hélice  $3_{10}$ , dans laquelle les liaisons hydrogène ne sont pas parallèles et la structure moléculaire particulièrement dense, n'existe que dans de rares cas (par exemple dans des séquences contenant l'acide  $\alpha$ -amino butyrique, ou à certaines extrémités d'hélices  $\alpha$ ), généralement sur un petit nombre de résidus. Son nom provient de ce qu'elle est constituée de trois résidus par tour, avec 10 atomes du donneur à l'accepteur de liaison hydrogène.

### L'hélice $\pi$

L'hélice  $\pi$  n'est que très rarement observée, et uniquement comme segment d'hélices plus longues. Sa conformation large et plate entraîne la présence d'un trou axial trop étroit pour admettre des molécules d'eau, mais trop large pour que s'établissent des interactions de van der Waals à travers l'axe de l'hélice.

Hélice	n (nombre de résidu par tour)	p (pas)	Liaison H	Translation par tour ( $\text{\AA}$ )
$3_{10}$	3	10	$i \rightarrow i+3$	6
$\alpha$	3.6	13	$i \rightarrow i+4$	5.4
$\pi$	4.4	16	$i \rightarrow i+5$	5.1

Tableau I-1 : Principales caractéristiques des différentes hélices.

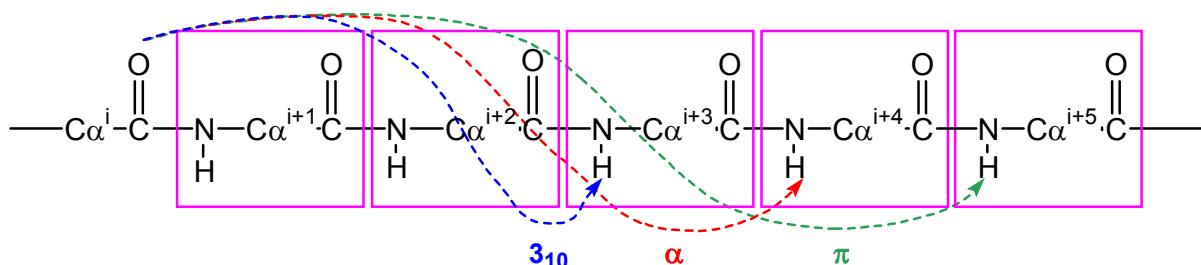


Figure I-4 : Schéma des liaisons hydrogène des principales hélices polypeptidiques.

### I.B.2.b. $\beta$ -hairpin

L'épinglé à cheveux de type  $\beta$  est l'équivalent du feuillet  $\beta$  dans les peptides. La structure du peptide peut être décomposée en trois parties (cf. figure I-5). Deux correspondent à des simples brins  $\beta$ , séparés par un coude  $\beta$  autrement appelé C<sub>10</sub> (ce type de motif sera décrit

dans la section suivante). Le coude permet un repliement de la chaîne peptidique, rendant possible l'interaction des deux brins  $\beta$  par de multiples liaisons hydrogène.

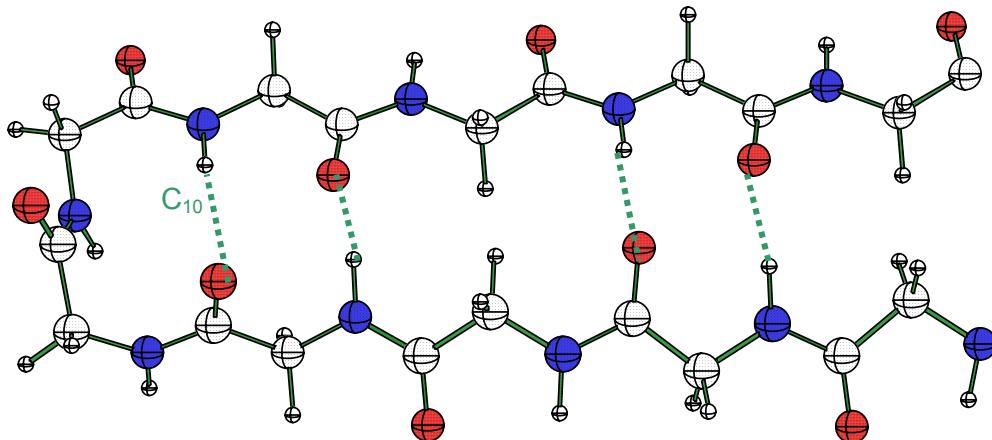


Figure I-5 : Forme d'une épingle à cheveux de type II ( $\beta$ -hairpin).

### I.B.2.c. Coudes ou turn, globule

Comme nous l'avons précisé auparavant, la structure tertiaire des protéines ou des peptides n'est pas complètement répétitive, mais correspond souvent à l'enchaînement de plusieurs structures secondaires ou à une structure globulaire. La jonction entre deux structures secondaires se fait par le biais de coudes ou turns, ils permettent un repliement de la chaîne peptidique. On distingue deux grandes classes de turns : les  $\gamma$ -turns et les  $\beta$ -turns qui présentent tous deux une liaison hydrogène intramoléculaire.

#### $\gamma$ -turns<sup>3</sup>

Les  $\gamma$ -turns présentent une liaison hydrogène entre la liaison C=O d'un résidu  $i$  et la liaison N-H du résidu  $i+2$ , formant un cycle à 7 atomes (cf figure I-6), que l'on note indifféremment  $\gamma$ -turn ou C<sub>7</sub>. On distingue deux types de  $\gamma$ -turns, selon l'orientation de la chaîne latérale. Le  $\gamma$ -turn inverse, qui présente un C<sub>7</sub> avec la chaîne latérale centrale en position équatoriale, est le plus couramment rencontré dans les protéines. Le  $\gamma$ -turn classique, qui présente un C<sub>7</sub> de type axial est, malgré son nom, extrêmement rarement observé.<sup>4</sup>

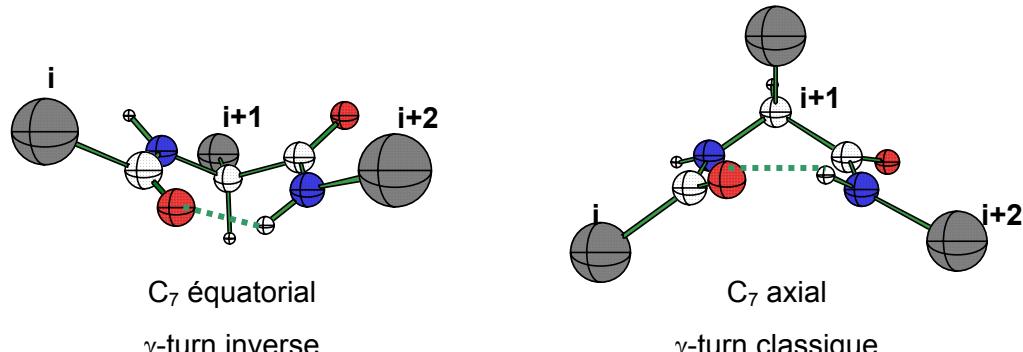


Figure I-6 : Structure des  $\gamma$ -turns inverse et classique.

### $\beta$ -turns

Les  $\beta$ -turns présentent une liaison hydrogène entre les résidus  $i$  et  $i+3$ , formant un cycle à 10 atomes (cf figure I-7). On les note indifféremment  $\beta$ -turn ou  $C_{10}$ . On distingue plusieurs types de  $\beta$ -turns, les deux catégories les plus importantes étant le  $\beta$ -turn de type I et celui de type II. Ils diffèrent par une inversion de l'orientation de la liaison peptidique centrale.

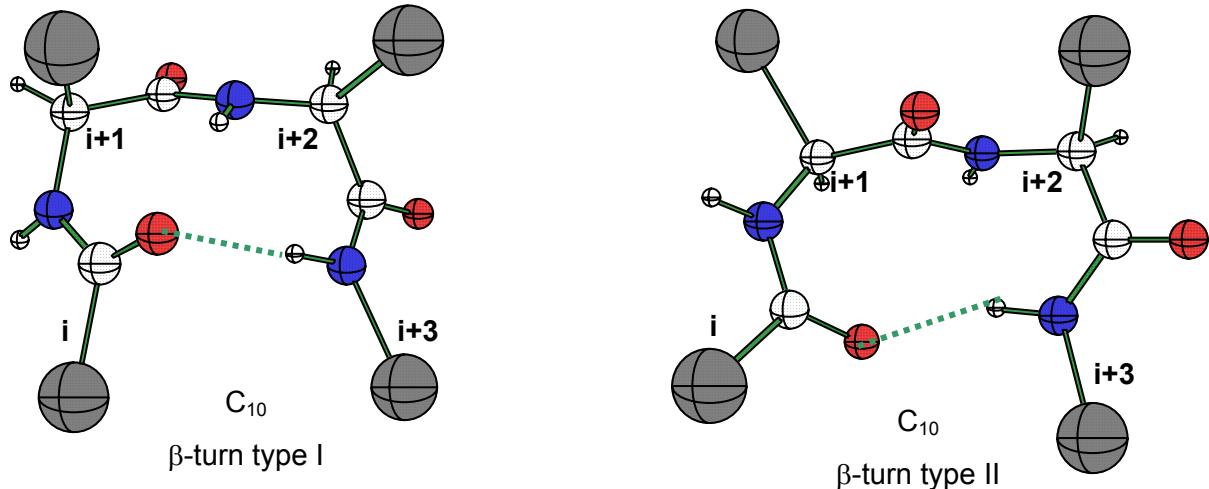


Figure I-7 : Structure des  $\beta$ -turns de types I et II.

Il y a une compétition entre la formation de  $\gamma$ -turns et la formation de  $\beta$ -turns ; les  $\gamma$ -turns sont entropiquement favorisés alors que les  $\beta$ -turns le sont enthalpiquement.

### I.B.3. Structures des peptides

La structure des protéines comme celle des peptides est gouvernée, entre autres, par les liaisons hydrogène et les interactions hydrophobes. Dans le cas des liaisons hydrogène, la structure résulte d'un compromis entre les liaisons hydrogène intramoléculaires (dont on a vu les différents motifs précédemment) appelées aussi l'auto-solvatation, et les liaisons hydrogène intermoléculaires entre le peptide et les molécules de solvant (la solvatation).

La plupart des petits peptides ne présentent pas de structure secondaire bien définie en solution (on parle alors de « random coils »). Ils peuvent aussi adopter une structure définie mais sans périodicité, on a donc alors affaire à des peptides globulaires. Cependant il est possible d'imposer artificiellement une structure hélicoïdale à un peptide, soit par modification chimique, soit par interaction avec un cation métallique ; ce point sera discuté au chapitre VI.

Par ailleurs, le solvant a comme dans le cas des acides aminés un rôle prépondérant dans la conformation des peptides. Ainsi, on peut induire une transition  $\beta$ -hairpin  $\rightarrow$  hélice selon la nature du solvant. C'est le cas de l'octa-peptide Boc-Leu-Val-Aib-Gly-Leu-Val-Val-OMe (Boc désigne le groupement tertiobutyloxycarbonyle : fonction de protection de l'amine terminale) qui adopte une conformation  $\beta$ -hairpin dans les solvants protiques  $((\text{CH}_3)_2\text{SO}$  ou  $\text{CH}_3\text{OH}$ ) et une structure en hélice dans les solvants qui interagissent peu avec la chaîne peptidique ( $\text{CHCl}_3$  et  $\text{CH}_3\text{CN}$ ).<sup>5</sup>

En phase gazeuse, les peptides doivent leur stabilité principalement aux liaisons hydrogène intramoléculaires, leurs conformations sont alors beaucoup plus structurées qu'en phase condensée. Ainsi on peut observer plus couramment des structures hélicoïdales, voire même des feuillets  $\beta$ .<sup>6</sup>

## II. Le cation $\text{Na}^+$ en biochimie

### II.A. $\text{Na}^+$ , élément essentiel pour les êtres vivants

#### II.A.1. Rôle fonctionnel du sodium : canaux à sodium

Le sodium est un élément essentiel pour les êtres vivants. C'est le principal cation inorganique extracellulaire chez les espèces animales. Il joue notamment un rôle prédominant dans la transmission des influx dans les tissus musculaires et nerveux. La transmission de l'influx nerveux se fait grâce à la présence, dans la membrane entourant le neurone, de canaux ioniques. La transmission correspond à la propagation d'un gradient de concentration d'ions ; les canaux ioniques s'activent et laissent entrer, en quelques millièmes de seconde, des ions sodium (des ions potassium sortent peu de temps après par d'autres canaux) : la face interne de la membrane devient alors chargée positivement. L'ouverture de canaux à sodium en une région donnée stimule la membrane de la région adjacente qui devient à son tour excitabile...et ainsi de suite.

Les canaux ioniques sont des complexes protéiques. Leur fonction est de faciliter ou catalyser la diffusion d'ions à travers les membranes biologiques. En effet, les membranes sont faiblement diélectriques et hydrophobes, le transport des ions à travers elles comporte donc une très haute barrière énergétique. Ces canaux présentent deux conformations, l'une ouverte qui autorise le passage des ions, et l'autre fermée qui l'interdit. Ils jouent donc le rôle de porte et régulent le transfert des ions à travers la membrane. Par ailleurs ils sont sélectifs vis-à-vis des ions. Il existe donc des canaux à sodium, à potassium, à calcium, à chlorures...

Ce modèle de l'influx nerveux dû au transport d'ions à travers une membrane est connu depuis les années 1950 grâce aux travaux de Hodgkin et Keynes, qui reçurent le prix Nobel de médecine en 1963.<sup>7</sup> Cependant ce n'est que depuis une dizaine d'années que la structure tridimensionnelle de certains de ces canaux a pu être déterminée. On peut notamment souligner le travail de R. MacKinnon qui a reçu le prix Nobel de chimie en 2003 pour son étude des canaux de potassium.<sup>8-11</sup> Dans le cas de la Kcsa, canal de potassium, le canal est constitué d'un tétramère, dont chaque unité est constituée de segments d'hélices  $\alpha$ . Les canaux à sodium sont des monomères quatre fois plus gros que les canaux à potassium.<sup>12, 13</sup> Cependant la structure des canaux à sodium et à potassium est relativement proche. Chaque segment d'hélice comporte une partie qui sert de porte dont la rotation autour d'une glycine permet d'adopter la conformation ouverte ou fermée, et une partie qui correspond au site sélectif constitué de 6 acides aminés. Le transport des ions, à travers les canaux, se fait par interactions successives avec les oxygènes des carbonyles peptidiques.

### II.A.2. Rôle structurant du sodium

L'activité biologique d'une protéine ou d'une enzyme est très sensible à sa structure tridimensionnelle, et aux détails de sa conformation autour de sites spécifiques. Cette structuration est souvent gouvernée par des liaisons hydrogène ou des interactions hydrophobes, comme nous l'avons vu précédemment, mais peut aussi l'être par des interactions électrostatiques, principalement de type charge-charge ou charge-dipôle. Les charges sont souvent liées aux chaînes latérales d'acides aminés (chargées positivement comme dans l'arginine ou la lysine, ou négativement comme dans l'aspartate ou le glutamate), mais peuvent aussi être dues à des cations métalliques. Les cations métalliques ont donc fréquemment un rôle structurant dans les macromolécules biologiques. Si les cations alcalins ont le plus souvent un rôle fonctionnel du type décrit dans le paragraphe précédent, il existe aussi des cas importants dans lesquels leur rôle est structural, en particulier dans des protéines allostériques (leurs activités sont modulées par la fixation non-covalente et réversible d'un ion sodium). L'exemple le mieux caractérisé jusqu'ici est celui de la thrombine,<sup>14, 15</sup> enzyme de la famille des protéases à sérine, qui joue un rôle essentiel dans la coagulation du sang. De nombreuses études cinétiques et cristallographiques, sur un grand nombre de mutants, ont permis d'établir que l'activation allostérique par le sodium se fait à distance du site catalytique ; les facteurs contrôlant la reconnaissance du cation et l'activation allostérique ne sont pas nécessairement les mêmes. La comparaison détaillée d'un grand nombre de structures et de leur activité catalytique a montré que les changements structuraux induits par la présence du sodium sont faibles, mais peuvent se propager sur de longues distances (au moins 15 Å dans ce cas, beaucoup plus dans

d'autres), assurant la transduction de l'attachement du cation en augmentation de l'activité catalytique. Un autre cas intéressant est celui de la dialkylglycine decarboxylase, qui catalyse la décarboxylation et la transamination de dialkylglycines telles que la  $\alpha$ -méthyl alanine.<sup>16, 17</sup> Cette enzyme possède deux sites à cations monovalents, dans lesquels se lient des couples ( $K^+$ ,  $Na^+$ ) ou ( $Na^+$ ,  $Na^+$ ). Dans l'un de ces sites, le sodium est positionné à l'extrémité C-terminale d'un fragment en hélice  $\alpha$ , créant une forte stabilisation de type charge-dipôle avec le macro-dipôle de l'hélice. Ce type de structure sera discuté en détail dans le chapitre VI.

Des compilations de structure de sites de  $Na^+$  dans des protéines<sup>18, 19</sup> montrent que la chélation du sodium est presque toujours assurée par des atomes d'oxygène, le plus souvent sans charge formelle (oxygènes des fonctions peptidiques de chaîne principale ou carbonyle de chaînes latérales, ou de molécules d'eau, par opposition aux groupes carboxylates fréquemment rencontrés dans la coordination des ions calcium). La distance Na-O moyenne est de 2.33 Å (en bon accord avec les résultats de calculs présentés plus loin), les nombres de coordination les plus fréquents sont 5 et 6, même si des valeurs de 4 et 7 sont aussi rencontrées occasionnellement.

## **II.B. $Na^+$ , élément souvent rencontré en biochimie en phase gazeuse**

Dans toutes les techniques d'ionisation de molécules d'intérêt biologique en phase gazeuse (que ce soient les techniques de MALDI ou d'ESI), on observe des ions protonés mais aussi fréquemment des adduits avec le potassium et surtout avec le sodium. En effet, le sodium étant présent dans tous les détergents, il est parfois difficile de s'en affranchir. Dans la plupart des études de systèmes biologiques, le sodium est considéré comme une impureté dont la présence est indésirable. En effet une concentration trop importante de sel peut modifier un spectre de masse provoquant une diminution de la sensibilité voire une suppression de certains signaux.<sup>20</sup> Cependant dans certains cas, les adduits de sodium peuvent être intéressants du point de vue analytique, puisque ces complexes peuvent conduire à une fragmentation spécifique de la molécule en interaction. Ainsi dans l'étude des résidus antibiotiques  $\beta$ -lactame du lait, l'étude quantitative des pénicillines G et V et de la cloxacilline se fait par la fragmentation spécifique de l'adduit sodié, par couplage chromatographie liquide – spectrométrie de masse en tandem.<sup>21</sup>

### III. Chimie en phase gazeuse.

#### III.A. Méthodes de spectrométrie de masse

Les molécules biologiques pouvant être désormais produites en phase gazeuse grâce au développement de sources d'ions adaptées (ElectroSpray Ionisation : ESI ou Matrix Assisted Laser Desorption Ionisation : MALDI), l'étude de leur comportement en phase gazeuse est un domaine en plein développement. Les méthodes de spectrométrie de masse sont aujourd'hui au cœur de l'analyse de biomolécules, notamment l'analyse protéomique. Les protocoles actuels se basent sur la fragmentation initiale d'une protéine en peptides, en utilisant des enzymes protéolytiques. Puis la séquence de ces peptides est déterminée par leur mode de fragmentation, en utilisant diverses méthodes d'activation (par collision avec un gaz chimiquement inerte : CID pour Collision Induced Dissociation ; par photodissociation ou encore par recombinaison cation-électron : ECD pour Electron Capture Dissociation). Ces techniques de détermination de structure primaire sont aujourd'hui largement automatisées.

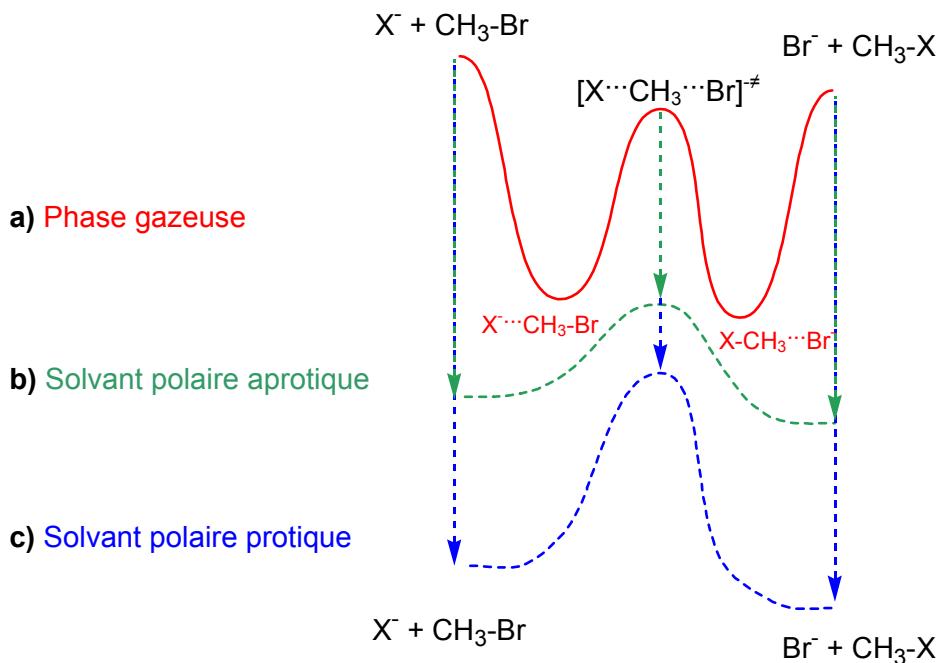
Les techniques de spectrométrie de masse pour la détermination des structures supérieures (secondaire, tertiaire et quaternaire) commencent à être développées, cependant elles sont encore loin d'être automatisées. Nos travaux se situent dans cette thématique. Nous avons développé l'identification d'ions biologiques gazeux par spectroscopie infrarouge. Ces expériences ont été menées en collaboration avec le groupe « Ions en Phase Gazeuse » du laboratoire de Chimie Physique (Orsay), à l'aide du spectromètre de masse de Résonance Cyclotronique Ionique à Transformée de Fourier (FT-ICR) MICRA, développé par cette équipe. Les spectromètres de masse FT-ICR présentent plusieurs avantages, notamment de très bonnes résolution et précision en masse. D'autre part une cellule ICR permet de piéger les ions pendant des temps variables, et ainsi d'être utilisée comme chambre de réaction entre l'ion choisi et une molécule neutre. MICRA (pour Mobile ICR Analyzer), appareil non commercial, a la particularité d'être transportable et peut donc être couplé à d'autres montages. La description de cette expérience originale sera présentée dans le chapitre III. Le développement de techniques de détermination structurale basées sur la spectrométrie de masse et adaptées au cas de biomolécules fera l'objet du chapitre II.

## III.B. Intérêt de la chimie en phase gazeuse

### III.B.1. Propriétés intrinsèques

Nous venons de voir (Partie I.A) que le solvant, le pH etc. jouent des rôles très importants dans la structure des acides aminés ou celle des peptides. Afin de comprendre et de quantifier les différentes contributions qui stabilisent les structures des peptides, il est nécessaire de faire la part des effets de solvant et de contre-ions, et des propriétés intrinsèques des peptides. La chimie en phase gazeuse permet d'étudier la réactivité et la structure intrinsèques d'espèces ioniques.

L'étude en phase gazeuse a, par le passé, montré qu'en donnant accès à la réactivité intrinsèque de certaines espèces, elle pouvait être indispensable pour comprendre leur réactivité en solution. L'exemple le plus important en chimie organique a été l'étude de la substitution nucléophile d'ordre 2 ( $S_N2$ ). En effet, la comparaison de la cinétique des réactions de  $S_N2$  d'un halogénure  $X^-$  sur le bromure de méthyle :  $X^- + CH_3Br \rightarrow CH_3X + Br^-$  a permis d'établir que la réaction est non seulement extrêmement rapide en phase gazeuse (lorsque l'énergie de l'état de transition est inférieure à celle des réactifs (figure I-8)), rapide en solvant polaire aprotique et lente en solvant polaire protique, mais aussi que l'ordre de nucléophilie des halogénures est inversé entre les réactions en solvant polaire protique ( $I^- > Br^- > Cl^- > F^-$ ) et en solvant polaire aprotique ou en phase gazeuse ( $F^- > Cl^- > Br^- > I^-$ ).<sup>22, 23</sup> Ainsi l'étude en phase gazeuse montre que la nucléophilie intrinsèque des halogénures est d'autant plus importante que le nucléophile est petit et chargé. Par contre en solution, la solvatation joue un rôle primordial dans le comportement nucléophile. En effet, en solvant protique, les nucléophiles anioniques sont d'autant plus stabilisés par liaisons hydrogène qu'ils sont petits. Ceci explique l'inversion de nucléophilie entre la phase gazeuse ou un solvant aprotique et un solvant protique, puisque les nucléophiles petits, excellents nucléophiles en phase gazeuse, se voient très stabilisés par solvatation, contrairement aux états de transition dans lesquels la charge est délocalisée. Ceci est illustré avec la figure I-8, qui présentent les profils réactionnels en phase gazeuse, en solvant polaire aprotique et en solvant polaire protique.



**Figure I-8 :** Profil énergétique de la réaction de  $S_N2$  :  $X^- + CH_3Br \rightarrow CH_3X + Br^-$  **a)** en phase gazeuse, **b)** en phase liquide avec un solvant polaire aprotique **c)** en phase liquide avec un solvant polaire protique.

### III.B.2. Conformation d'un peptide au sein d'une protéine

Les expériences en phase condensée montrent que les hélices  $\alpha$ , formées d'un nombre de résidus généralement assez petit (typiquement inférieur à 15), constituent environ 30% de la structure secondaire des protéines. Par contre les peptides correspondants aux mêmes éléments de séquence, ne forment généralement pas d'hélices dans l'eau. Ceci est en partie du à l'environnement protéique dans lequel se trouve la séquence peptidique. En effet, on évalue que la constante diélectrique à l'intérieur d'une protéine est d'environ 4. Rappelons que dans le cas de la phase aqueuse  $\epsilon_r=78$  et dans le vide  $\epsilon_r=1$ . Ainsi si on veut comprendre la conformation de certaines séquences peptidiques au sein d'une protéine, l'étude de peptides en phase gazeuse peut être une modélisation raisonnable.

### III.B.3. La micro-solvatation

Nous avons vu qu'il existe une transition entre les structures (exemple des acides aminés zwitterioniques ou sans charges formelles) ou entre la réactivité (exemple de la nucléophilie) en solution ou en phase gazeuse. Il est aujourd'hui possible de produire de

manière contrôlée (directement par MALDI, ESI ou par réaction d'association entre une molécule ionisée et une vapeur de solvant), des adduits solvatés par une, deux ou plusieurs molécules de solvant ; on parle alors de micro-solvatation. On peut donc simuler la transition entre phase liquide et phase gazeuse. Si cette approche s'est avérée peu efficace dans le cas de la réactivité, à cause de l'impossibilité de contrôler la micro-solvatation des réactifs neutres par rapport à celle des ions, ouvrant d'autres voies de réaction sans analogie avec la solvatation macroscopique (exemple de la  $S_N2$ ),<sup>24-27</sup> elle a fait ses preuves dans le cas des études structurales. Deux stratégies peuvent être appliquées. La première consiste à former directement des adduits par électronébullisation puis à les fragmenter afin de déterminer l'énergie de liaison molécule-solvant. L'alternative est de produire la molécule seule en phase gazeuse puis de former l'adduit par association avec des molécules de solvant, et d'étudier la thermodynamique de cette association en mesurant  $\Delta_rH^\circ$  et  $\Delta_rS^\circ$ . Les premières expériences de micro-solvatation de petits peptides protonés  $Gly_nH^+$  ( $n=1-4$ ) ont été menées par Kebarle *et al.* en 1995.<sup>28</sup> Depuis lors, les expériences de micro-solvatation sur des systèmes biologiques ont fait l'objet de nombreux articles. Ainsi, Bowers *et al.* ont montré, sur des di-, tri-, tétra- et penta-peptides protonés contenant l'alanine et l'arginine, que les énergies successives d'attachement d'une molécule d'eau sont de l'ordre de 28 à 60 kJ/mol, les plus grandes valeurs étant pour l'association de la première molécule d'eau. D'autre part, les petits peptides sont mieux solvatés par la première molécule d'eau que les plus grands, parce qu'il y a moins d'auto-solvatation pour les petits peptides.<sup>29</sup> Jarrold *et al.* ont étudié l'hydratation de  $Ac-KA_nH^+$  et  $Ac-A_nKH^+$   $n=15,20$  présentant une structure globulaire et une structure hélicoïdale, respectivement. Ils ont montré que l'hydratation des structures globulaires était, thermodynamiquement, plus favorable que celles des hélices.<sup>30</sup> Ce résultat a été ensuite exploité pour déterminer le nombre de résidus à partir duquel une structure en hélice devient plus stable qu'une structure globulaire.<sup>31</sup> Par ailleurs, les expériences de micro-solvatation permettent d'étudier la transition structurale entre la phase gazeuse et la phase liquide. L'équipe d'Evan Williams a étudié entre autres l'effet de l'hydratation sur les complexes valine - cation alcalin. En phase liquide la valine se présente sous forme zwitterionique dans les complexes  $Val-Li^+$ ,  $Val-Na^+$ ,  $Val-K^+$ , alors qu'en phase gazeuse elle présente une structure sans charge formelle.<sup>32, 33</sup> Ces auteurs ont récemment montré que dans le complexe  $Val-Li^+$  solvaté par trois molécules d'eau, la valine a une structure zwitterionique. Ceci montre que l'on peut simuler la structure de molécules en phase liquide par micro-solvatation en phase gazeuse.<sup>34</sup>

### III.C. Méthodes de chimie théorique

Les outils de chimie quantique permettent depuis une dizaine d'années de calculer les propriétés électroniques et vibrationnelles des molécules organiques ou organométalliques en phase gazeuse, et ce avec une excellente précision. Il existe aussi des modèles permettant de simuler la solvatation, par micro-solvatation explicite et/ou en modélisant le solvant par un continuum diélectrique. De plus, aujourd'hui, le développement des algorithmes, des codes et des ordinateurs permet d'envisager la modélisation d'assez grands systèmes, notamment biologiques, avec un temps de calcul raisonnable, que ce soit par mécanique moléculaire, par calculs *ab-initio* ou par des méthodes hybrides. Nous verrons par la suite (chapitre II) que toutes les méthodes expérimentales de détermination structurale d'ions gazeux font appel à des calculs (simulation de spectres infrarouges, de sections efficaces de collision, d'énergie de liaison...) pour valider ou interpréter les résultats expérimentaux. Nous avons donc combiné, lors de ce travail, approches expérimentale et théorique. Pour la partie théorique nous avons utilisé plusieurs logiciels : Gaussian 03<sup>35</sup>, HyperChem<sup>36</sup> et Turbomole<sup>37</sup> dont les caractéristiques seront développées au chapitre III.

### III.D. Conclusions

L'étude en phase gazeuse de la structure d'ions biologiques est d'une part, relativement aisée en principe d'un point de vue théorique, d'autre part accessible expérimentalement grâce aux développements de nouvelles techniques couplées à la spectrométrie de masse. Les résultats obtenus en phase gazeuse permettent de définir les propriétés intrinsèques des molécules. Dans le cas de peptides, l'étude en phase gazeuse peut être une bonne alternative pour comprendre leurs conformations au sein d'une protéine où règne une constante diélectrique plus proche de celle du vide que celle d'une solution aqueuse.

## IV. Notre étude

Nous nous sommes intéressés à l'interaction du cation sodium avec de petites molécules biologiques (acides aminés et oligopeptides) en phase gazeuse, afin de comprendre son rôle structurant. En effet, la complexation par le sodium peut induire une isomérisation comme le passage de la structure sans charge formelle d'un acide aminé à une structure zwitterionique. Il peut aussi induire des changements de conformation comme le passage de la structure globulaire d'un peptide nu à une structure hélicoïdale. Nous nous sommes donc intéressés à ces deux facettes : isomérisation d'acides aminés et changement de

conformation de peptides en phase gazeuse, par la complexation avec  $\text{Na}^+$ . Pour ce faire, nous avons combiné approches expérimentales par spectrométrie de masse FT-ICR et théoriques. Dans le cadre d'une collaboration avec le groupe « Ions en Phase Gazeuse » du Laboratoire de Chimie Physique d'Orsay, nous avons adapté une technique de couplage entre un spectromètre de masse FT-ICR et un laser infrarouge, pour obtenir les premiers spectres infrarouges d'ions biologiques. Une autre collaboration, avec l'équipe du Professeur Chrys Wesdemiotys (Akron, Etats Unis), nous a permis de déterminer l'affinité de petits peptides pour  $\text{Na}^+$ . Nous avons réalisé la partie théorique de ce travail.

Ce mémoire suit le schéma suivant :

Le chapitre II décrira les principes des techniques, par spectrométrie de masse, de détermination structurale de biomolécules. Nous insisterons particulièrement sur les deux expériences utilisées dans ce travail.

Le troisième chapitre présentera le montage et les développements expérimentaux ainsi que les méthodes théoriques utilisés. L'accent sera mis sur les aspects originaux de ces techniques : les montages non commerciaux et les méthodes de calculs non usuelles (le logiciel TURBOMOLE).

Les trois derniers chapitres présentent les résultats sur le rôle structurant de  $\text{Na}^+$ . Ils sont classés par taille croissante de peptides. Le chapitre IV est consacré aux complexes entre acides aminés et  $\text{Na}^+$ , le cinquième aux complexes sodiés de di-peptides aux tétrapeptides, principalement d'oligo-glycines et d'oligo-alanines, et enfin le dernier, du penta-peptide au déca-peptide :  $\text{Gly}_n\text{-Na}^+$ ,  $5 \leq n \leq 10$ .

## Bibliographie

- [1] Voet, D.; Voet, J. G., *Biochimie*, DeBoeck Université, 2ième édition. **1998**.
- [2] Venkatraman, J.; Shankaramma, S. C.; Balaram, P., Design of folded peptides. *Chem. Rev.* **2001**, 101, (10), 3131-3152.
- [3] Stachel, S. J.; Hu, H. T.; Van, Q. N.; Shaka, A. J.; Van Vranken, D. L., Stabilization of a C<sub>7</sub> equatorial gamma turn in DMSO-d<sub>6</sub> by a ditryptophan crosslink. *Bioorg. Med. Chem.* **1998**, 6, (9), 1439-1446.
- [4] Milnerwhite, E. J.; Ross, B. M.; Ismail, R.; Belhadjmostefa, K.; Poet, R., One Type of Gamma-Turn, Rather Than the Other Gives Rise to Chain-Reversal in Proteins. *J. Mol. Biol.* **1988**, 204, (3), 777-782.
- [5] Awasthi, S. K.; Shankaramma, S. C.; Raghothama, S.; Balaram, P., Solvent-induced beta-hairpin to helix conformational transition in a designed peptide. *Biopolymers* **2001**, 58, (5), 465-476.
- [6] Dugourd, P.; Antoine, R.; Breaux, G.; Broyer, M.; Jarrold, M. F., Entropic stabilization of isolated beta-sheets. *J. Am. Chem. Soc.* **2005**, 127, (13), 4675-4679.
- [7] Hodgkin, A. L.; Huxley, A. F., A Quantitative Description of Membrane Current and Its Application to Conduction and Excitation in Nerve. *J. Physiol.-London* **1952**, 117, (4), 500-544.
- [8] MacKinnon, R., Potassium channels and the atomic basis of selective ion conduction (Nobel lecture). *Angew. Chem.-Int. Edit.* **2004**, 43, (33), 4265-4277.
- [9] Zhou, Y.; MacKinnon, R., Ion binding affinity in the cavity of the KcsA potassium channel. *Biochemistry* **2004**, 43, (17), 4978-4982.
- [10] Jiang, Y. X.; Lee, A.; Chen, J. Y.; Cadene, M.; Chait, B. T.; MacKinnon, R., The open pore conformation of potassium channels. *Nature* **2002**, 417, (6888), 523-526.
- [11] Jiang, Y. X.; Lee, A.; Chen, J. Y.; Ruta, V.; Cadene, M.; Chait, B. T.; MacKinnon, R., X-ray structure of a voltage-dependent K<sup>+</sup> channel. *Nature* **2003**, 423, (6935), 33-41.
- [12] Catterall, W. A., From ionic currents to molecular mechanisms: The structure and function of voltage-gated sodium channels. *Neuron* **2000**, 26, (1), 13-25.
- [13] Balser, J. R., Structure and function of the cardiac sodium channels. *Cardiovasc. Res.* **1999**, 42, (2), 327-338.
- [14] Pineda, A. O.; Carrell, C. J.; Bush, L. A.; Prasad, S.; Caccia, S.; Chen, Z. W.; Mathews, F. S.; Di Cera, E., Molecular dissection of Na<sup>+</sup> binding to thrombin. *J. Biol. Chem.* **2004**, 279, (30), 31842-31853.
- [15] Di Cera, E., Thrombin: a paradigm for enzymes allosterically activated by monovalent cations. *C. R. Biol.* **2004**, 327, (12), 1065-1076.

- [16] Toney, M. D.; Hohenester, E.; Cowan, S. W.; Jansonius, J. N., Dialkylglycine Decarboxylase Structure - Bifunctional Active-Site and Alkali-Metal Sites. *Science* **1993**, 261, (5122), 756-759.
- [17] Toney, M. D.; Hohenester, E.; Keller, J. W.; Jansonius, J. N., Structural and Mechanistic Analysis of 2 Refined Crystal-Structures of the Pyridoxal Phosphate-Dependent Enzyme Dialkylglycine Decarboxylase. *J. Mol. Biol.* **1995**, 245, (2), 151-179.
- [18] Harding, M. M., The architecture of metal coordination groups in proteins. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **2004**, 60, 849-859.
- [19] Glusker, J. P., Structural Aspects of Metal Liganding to Functional-Groups in Proteins. *Adv. Protein Chem.* **1991**, 42, 1-76.
- [20] Lavarone, A. T.; Udekwu, O. A.; Williams, E. R., Buffer loading for counteracting metal salt-induced signal suppression in electrospray ionization. *Anal. Chem.* **2004**, 76, (14), 3944-3950.
- [21] Holstege, D. M.; Puschner, B.; Whitehead, G.; Galey, F. D., Screening and mass spectral confirmation of beta-lactam antibiotic residues in milk using LC-MS/MS. *J. Agric. Food Chem.* **2002**, 50, (2), 406-411.
- [22] Olmstead, W. N.; Brauman, J. I., Gas-Phase Nucleophilic Displacement-Reactions. *J. Am. Chem. Soc.* **1977**, 99, (13), 4219-4228.
- [23] Laerdahl, J. K.; Uggerud, E., Gas phase nucleophilic substitution. *Int. J. Mass Spectrom.* **2002**, 214, (3), 277-314.
- [24] Bohme, D. K.; Rakshit, A. B., Gas-Phase Measurements of the Influence of Stepwise Solvation on the Kinetics of Nucleophilic Displacement-Reactions with  $\text{CH}_3\text{Cl}$  and  $\text{CH}_3\text{Br}$  at Room-Temperature. *J. Am. Chem. Soc.* **1984**, 106, (12), 3447-3452.
- [25] Seeley, J. V.; Morris, R. A.; Viggiano, A. A., Temperature dependences of the rate constants and branching ratios for the reactions of  $\text{F}-(\text{H}_2\text{O})_{0-5}$  with  $\text{CH}_3\text{Br}$ . *J. Phys. Chem. A* **1997**, 101, (25), 4598-4601.
- [26] Morokuma, K., Potential-Energy Surface of the  $\text{S}_{\text{N}}2$  Reaction in Hydrated Clusters. *J. Am. Chem. Soc.* **1982**, 104, (13), 3732-3733.
- [27] Raugei, S.; Cardini, G.; Schettino, V., Microsolvation effect on chemical reactivity: The case of the  $\text{Cl}^- + \text{CH}_3\text{Br}$   $\text{S}_{\text{N}}2$  reaction. *J. Chem. Phys.* **2001**, 114, (9), 4089-4098.
- [28] Klassen, J. S.; Blades, A. T.; Kebarle, P., Determinations of Ion-Molecule Equilibria Involving Ions Produced by Electrospray - Hydration of Protonated Amines, Diamines, and Some Small Peptides. *J. Phys. Chem.* **1995**, 99, (42), 15509-15517.
- [29] Liu, D. F.; Wyttenbach, T.; Barran, P. E.; Bowers, M. T., Sequential hydration of small protonated peptides. *J. Am. Chem. Soc.* **2003**, 125, (28), 8458-8464.
- [30] Kaleta, D. T.; Jarrold, M. F., Noncovalent interactions between unsolvated peptides. *J. Phys. Chem. A* **2002**, 106, (42), 9655-9664.

- [31] Kohtani, M.; Jarrold, M. F., Water molecule adsorption on short alanine peptides: How short is the shortest gas-phase alanine-based helix? *J. Am. Chem. Soc.* **2004**, 126, (27), 8454-8458.
- [32] Lemoff, A. S.; Bush, M. F.; Williams, E. R., Binding energies of water to sodiated valine and structural isomers in the gas phase: The effect of proton affinity on zwitterion stability. *J. Am. Chem. Soc.* **2003**, 125, (44), 13576-13584.
- [33] Jockusch, R. A.; Lemoff, A. S.; Williams, E. R., Effect of metal ion and water coordination on the structure of a gas-phase amino acid. *J. Am. Chem. Soc.* **2001**, 123, (49), 12255-12265.
- [34] Lemoff, A. S.; Williams, E. R., Binding energies of water to lithiated valine: Formation of solution-phase structure in vacuo. *J. Am. Soc. Mass Spectrom.* **2004**, 15, (7), 1014-1024.
- [35] Gaussian03, revision B.05, Frisch, M. J. et al., Gaussian, Inc., Pittsburgh, PA, 2003.
- [36] HyperChem 6.0 : Hypercube, Inc., 1115 NW 4th Street, Gainesville, FL 32601 (USA).
- [37] TURBOMOLE 5.6 , Ahlrichs, R. et al, Quantum Chemistry Group, University of Karlsruhe (Germany).

# **Chapitre II :**

## **Principes expérimentaux de**

## **caractérisation structurale de**

## **petits ions biologiques**

### **en phase gazeuse**

## Introduction

Il existe, aujourd’hui, de nombreuses et efficaces méthodes de caractérisation structurale en phase condensée de composés biologiques, que ce soit la RMN, la diffraction des rayons X ou le dichroïsme circulaire. Depuis une dizaine d’années, ce champ d’études s’est étendu à la phase gazeuse, notamment grâce aux nouvelles performances de techniques d’ionisation de composés biologiques en phase gazeuse (ESI et MALDI) et de l’amélioration des performances des spectromètres de masse. La spectrométrie de masse est devenue un élément indispensable dans l’élucidation de la structure primaire des protéines. Nous allons voir, dans ce deuxième chapitre, que la spectrométrie de masse est aussi un outil important dans l’étude des structures secondaires et tertiaires des peptides et des protéines, en phase gazeuse. De nombreuses stratégies ont été mises en œuvre ces dix dernières années. Nous ne ferons donc pas une description exhaustive de tous les nouveaux principes expérimentaux développés pour la caractérisation tridimensionnelle d’ions biologiques, mais nous mentionnerons les plus utilisés actuellement, dans une première partie. La seconde partie de ce chapitre est consacrée au principe de la spectroscopie IRMPD (spectroscopie de Dissociation induite par l’absorption de MultiPhotons InfraRouge) d’ions en phase gazeuse, technique utilisée lors de ce travail.

## I. Résumé des développements expérimentaux actuels

Dans ce paragraphe, nous allons voir trois grandes classes d'expériences de détermination structurale d'ions biologiques en phase gazeuse : la méthode de la cinétique d'échange hydrogène-deutérium (échange H/D), la méthode de mesure de mobilité ionique et les méthodes de détermination d'énergie de liaison.

### I.A. Echange H/D

#### I.A.1. Principe

##### I.A.1.a. Principe de cette méthode

Ce type d'expérience est basé sur l'hypothèse que tous les hydrogènes labiles d'une protéine ne s'échangent pas à la même vitesse ; ainsi la connaissance de la vitesse d'échange et le nombre d'hydrogènes échangés sont des indicateurs des propriétés structurales de la protéine ou du peptide. La vitesse d'échange est dépendante du type d'hydrogène : ammonium, amine, amide, acide, etc, et de l'environnement de cet hydrogène : est-il impliqué dans une liaison hydrogène, est-il stériquement accessible ? Deux stratégies sont élaborées selon que l'on s'intéresse à la structure de la molécule en solution ou en phase gazeuse. On peut donc, soit effectuer l'échange H/D en solution à l'aide d'un solvant deutéré et suivre la cinétique grâce à la spectrométrie de masse, soit effectuer l'échange H/D en phase gazeuse en faisant réagir l'ion d'intérêt (par exemple  $\text{MH}^+$ ) avec une molécule réactive deutérée, typiquement  $\text{D}_2\text{O}$ ,  $\text{ND}_3$  ou  $\text{CH}_3\text{OD}$ , ce qui permet de sonder la conformation en phase liquide ou en phase gazeuse, respectivement. Dans le cas de la phase gazeuse, le processus d'échange H/D reste encore assez mal caractérisé. Il a été étudié notamment par l'équipe de Beauchamp<sup>1</sup> sur des oligomères de glycine protonés, cinq mécanismes différents ont été identifiés. Ces mécanismes dépendent notamment de l'affinité protonique du réactif deutéré. Certains travaux ont mis en évidence que les hydrogènes d'ammonium s'échangent le plus rapidement, suivis par les hydrogènes amidiques ne participant pas à une liaison hydrogène, les hydrogènes d'acides carboxyliques s'échangent lentement, enfin les hydrogènes amidiques impliqués dans une liaison H ont une cinétique d'échange extrêmement lente.<sup>2, 3</sup> Cette tendance n'est à priori pas générale, parce qu'il existe plusieurs mécanismes d'échange, selon l'environnement du site échangeable.

La cinétique d'échange H/D permet de déterminer le type d'hydrogène accessible au solvant, en étudiant la cinétique. D'autre part, à un temps assez long, la distribution de pics isotopiques permet de déterminer le nombre de H échangés pour une conformation donnée. Un hydrogène accessible au solvant deutéré s'échange, contrairement à un H inaccessible ou impliqué dans une liaison hydrogène. Ainsi les conformations les plus ouvertes, exposant plus d'hydrogènes au solvant qu'une conformation compacte, verront leur nombre d'hydrogènes échangés plus important. Le nombre de H échangé donne donc une indication sur la forme de la molécule. Dans le cas de protéines cela permet de différencier la conformation native, plus compacte, d'une structure dénaturée. Par ailleurs, la superposition de plusieurs distributions est la signature de la présence de plusieurs conformations.

### I.A.1.b. Limites de cette méthode

Si la technique d'échange H/D en phase gazeuse s'est révélée efficace dans certains cas, les équipes de Beauchamp et de Bowers ont montré qu'elle a certaines limites. L'étude d'oligoglycines sodiées,<sup>4</sup> et de penta-peptides protonés<sup>5</sup> contenant quatre alanines et une arginine, par cinétique d'échange H/D en phase gazeuse, montre qu'il peut exister des étapes d'isomérisations intermédiaires dans le profil réactionnel, ce qui peut conduire à des interprétations erronées sur l'identification des hydrogènes labiles dans la molécule initiale. De plus le mécanisme, donc les vitesses d'échange H/D, sont dépendants du réactif deutéré (D<sub>2</sub>O, ND<sub>3</sub>...). Il a été ainsi montré que cette technique d'échange H/D, ne permet pas de sonder la conformation de minima, mais celle de l'état de transition de l'étape limitante d'échange H/D.

### I.A.2. Un exemple d'échange H/D en solution<sup>6</sup>

Katta et Chait se sont intéressés à la conformation de l'ubiquitine et de la lysozyme d'œuf de poule, dans différentes conditions de solvant. Ils ont montré que, avec 1% de CH<sub>3</sub>COOD dans D<sub>2</sub>O, l'ubiquitine comme la lysozyme échangeaient 62 à 65% de leurs hydrogènes labiles en 30 minutes. Par contre, dans des conditions méthanoliques : mélange 1 :1 de CH<sub>3</sub>OD : D<sub>2</sub>O avec 1% CH<sub>3</sub>COOD, l'ubiquitine échange 94% de ses hydrogènes labiles alors que la lysozyme n'en échange que 62%. Par ailleurs, le rendement d'échange H/D de la lysozyme, en condition réductrice, est de 96% en 30 minutes. Ceci permet de déduire les conditions dénaturantes de chaque protéine, le rendement d'échange H/D étant bien inférieur pour une protéine native que pour une protéine dénaturée. En effet, une protéine native a une structure relativement compacte, certains de ces hydrogènes labiles se trouvent piégés dans la partie hydrophobe de la protéine, inaccessibles au solvant, ou impliqués dans une liaison hydrogène, contrairement à une structure dénaturée, beaucoup plus ouverte.

### I.A.3. Quelques résultats d'échange H/D en phase gazeuse

#### I.A.3.a. Le cytochrome c<sup>7,8</sup>

L'équipe de McLafferty a étudié l'effet de l'échange H/D sur les états de charge +6 à +19 du cytochrome c (104 acides aminés, M=12300 g/mol, 198 hydrogènes labiles), produits par ESI. Ils ont montré, par échange H/D en phase gazeuse, que la distribution de masse de l'état de charge +7 était très large, signature d'au moins sept conformations distinctes, alors que l'état de charge +9 ne présentait qu'une distribution isotopique, donc une seule conformation. Ils ont identifié sept conformations sur l'ensemble des états de charge. L'étude des fragmentations et de l'échange H/D des différents fragments a permis d'identifier la structure tertiaire de chaque conformère, notamment plusieurs segments d'hélices  $\alpha$ .

#### I.A.3.b. Gly<sub>2</sub>H<sup>+</sup>

L'équipe de Marshall a attribué chaque constante de vitesse d'échange H/D aux différents hydrogènes de Gly<sub>2</sub>H<sup>+</sup>. La cinétique d'échange H/D a montré que trois hydrogènes s'échangeaient très rapidement à la même vitesse ( $k_1=6.8\pm0.6\ 10^{-11}\ \text{cm}^3\text{s}^{-1}\text{molécule}^{-1}$ ), deux autres constantes de vitesse ont été déterminées, l'une très faible ( $k_2=0.072\pm0.015\ 10^{-11}\ \text{cm}^3\text{s}^{-1}\text{molécule}^{-1}$ ) et l'autre intermédiaire ( $k_3=1.2\pm0.2\ 10^{-11}\ \text{cm}^3\text{s}^{-1}\text{molécule}^{-1}$ ). La réaction la plus rapide correspond à l'échange des trois hydrogènes de l'amine terminale protonée. Afin d'attribuer les deux autres constantes de vitesse à l'hydrogène amidique ou l'hydrogène acide, ils les ont comparées à celles de deux dérivés méthylés de Gly<sub>2</sub>H<sup>+</sup> : GlySarH<sup>+</sup> et Gly<sub>2</sub>OMeH<sup>+</sup> (Schéma II-1). La constante intermédiaire correspond à la constante mesurée dans le cas de Gly<sub>2</sub>OMeH<sup>+</sup>, alors que la plus faible correspond à celle de GlySarH<sup>+</sup>. On peut ainsi identifier la constante de chaque hydrogène et en conclure que la vitesse d'échange est la plus grande pour un H d'ammonium puis un H amidique et enfin un H acide dans le cas des dipeptides protonés.

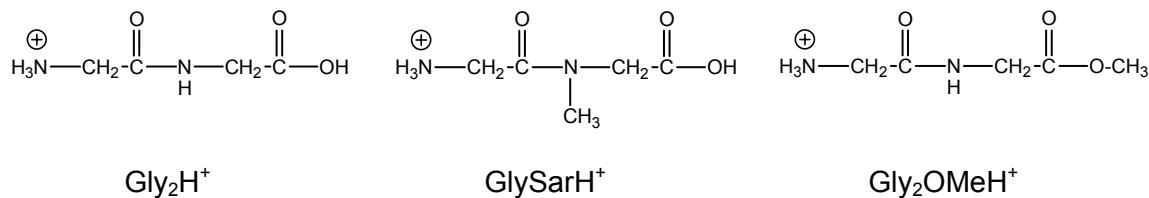


Schéma II-1

### I.A.4. Conclusions

La technique d'échange H/D donne, dans la plupart des cas, une indication sur le type et le nombre d'hydrogènes labiles exposés au solvant ou à la surface de la molécule et donc sa forme générale (compacte ou dépliée). Cependant cette méthode a des limites : il n'est pas encore possible de relier systématiquement le nombre et la vitesse d'échange H/D à des caractéristiques structurales. Il est donc souvent nécessaire de coupler cette technique à d'autres techniques d'identification structurale.

## I.B. Technique de mobilité ionique

Cette technique a été développée dans les années 90, principalement par trois groupes américains : ceux de Bowers, Jarrold et Clemmer. Elle a pour second nom « la chromatographie ionique ».

### I.B.1. Principe

#### I.B.1.a. Principe<sup>9</sup>

Son principe est basé sur le déplacement d'un ion soumis à un champ électrique faible dans un tube rempli d'un gaz tampon statique. L'ion se déplace à travers le tube, sa vitesse étant limitée par de multiples collisions avec le gaz tampon. Le temps de transit à l'intérieur du tube est alors caractéristique de sa section efficace de collision : plus celle-ci est importante, plus la vitesse de l'ion sera faible et donc le temps de rétention de l'ion à l'intérieur du tube sera plus important. La section efficace de collision dépend de la forme et donc de la conformation de l'ion. Le temps de transit dépend des conditions expérimentales, notamment du champ électrique, de la pression et de la masse du gaz statique, de la longueur du tube... ; il peut être relié à la section efficace de collision moyenne ou mobilité. Ainsi la détermination de la section efficace de collision est un indicateur de la conformation de l'ion, notamment une forme allongée comme une hélice a une section efficace de collision beaucoup plus importante que celle d'une structure compacte.

#### I.B.1.b. Technique couplée à la modélisation moléculaire

En parallèle des résultats expérimentaux, il est possible de déterminer la section efficace de collision moyenne de manière théorique pour une conformation donnée. Ceci implique de mener une recherche conformationnelle théorique de l'ion étudié. Dans le cas de grands peptides et a fortiori de protéines, la méthode de choix pour une étude conformationnelle est la dynamique moléculaire ou la recherche de Monte Carlo, couplées à un champ de forces pour

les calculs d'énergie. Pour chaque conformation trouvée, la section efficace de collision est calculée à l'aide d'un modèle développé par Bowers *et al.*<sup>10</sup>

La comparaison entre expérience et théorie permet de déterminer les types de conformations expérimentalement présentes. En effet, lorsque la mobilité calculée pour un conformère est en bon accord avec l'expérience, ce conformère devient un candidat pour la structure de l'ion observé.

### I.B.1.c. Limite de cette méthode

La faiblesse de cette approche est que la structure d'une molécule, quelle que soit sa complexité, n'est caractérisée que par un paramètre unique. Or il existe, en général, plusieurs conformations ayant des sections efficaces de collision proches. Dans le cas de petites molécules comme les complexes acide aminés- $\text{Na}^+$ , leurs conformations de plus basse énergie ont des géométries relativement proches et donc des mobilités très similaires, il est alors impossible de les différencier expérimentalement.<sup>11</sup> Dans les cas de grands peptides ou protéines, on ne peut distinguer, avec certitude, que des grandes familles de structure comme, par exemple, les formes globulaires des formes hélicoïdales.

## I.B.2. Application aux cas d'oligomères d'acides aminés

L'équipe de Jarrold s'est particulièrement intéressée à la structure d'oligomères de glycine et d'alanine en fonction du nombre de résidu, par étude de mobilité ionique. Ils ont montré que  $\text{Ala}_n\text{H}^+$  pour  $n=5-20$  avait une structure globulaire, alors que  $\text{Ala}_n\text{LysH}^+$  avait une structure hélicoïdale à partir de  $n=7$ .<sup>12</sup> La figure II-1 présente les sections efficaces calculées et mesurées relatives à une structure en hélice  $\alpha$  idéale, de telle manière que la section efficace de collision relative d'une hélice  $\alpha$  soit indépendante du nombre de résidus. La comparaison entre les sections efficaces de collision relatives mesurées et calculées indique des structures en hélice pour  $\text{Ac-Ala}_n\text{-LysH}^+$ . A l'inverse,  $\text{Ala}_n\text{H}^+$  présente une structure globulaire. Quant aux structures de  $\text{Gly}_n\text{LysH}^+$ , elles présentent toutes des structures globulaires.<sup>13</sup> Ceci a permis de montrer que la glycine a une plus faible propension à former des hélices que l'alanine en phase gazeuse, tendance identique aux structures en solution. Une discussion sur les tendances des différents peptides à former des hélices sera présentée au chapitre VI.

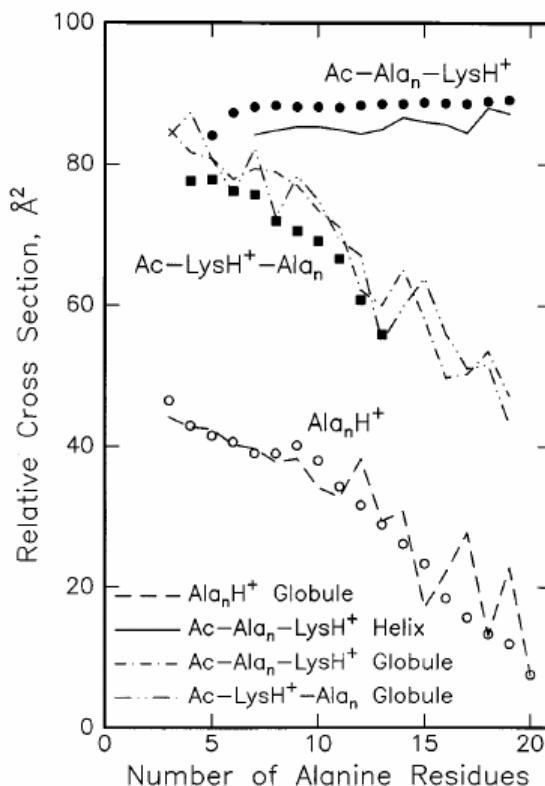


Figure 5. Plot of the relative cross sections for  $\text{Ala}_n\text{H}^+$  (○),  $\text{Ac-Ala}_n\text{-LysH}^+$  (●), and  $\text{Ac-LysH}^+-\text{Ala}_n$  (■) peptides against the number of alanine residues. The relative cross section scale is given by  $\Omega_{\text{av}}^{(1,1)} - 14.50n$  where the cross section,  $\Omega_{\text{av}}^{(1,1)}$ , is in  $\text{\AA}^2$  and  $14.50 \text{\AA}^2$  is the calculated average cross section per residue for an ideal polyalanine  $\alpha$ -helix. Cross sections calculated for  $\text{Ac-Ala}_n\text{-LysH}^+$  helices and  $\text{Ala}_n\text{H}^+$ ,  $\text{Ac-Ala}_n\text{-LysH}^+$ , and  $\text{Ac-LysH}^+-\text{Ala}_n$  globules are shown for comparison with the measured cross sections.

Figure II-1 : Sections efficaces calculées et mesurées relatives à une structure en hélice  $\alpha$  idéale, pour  $\text{Ala}_n\text{H}^+$ ,  $\text{Ac-Ala}_n\text{-LysH}^+$  et  $\text{Ac-LysH}^+-\text{Ala}_n$ .<sup>12</sup>

### I.C. Techniques de détermination d'énergie de liaison

Il existe de nombreuses méthodes basées sur la spectrométrie de masse, pour déterminer des grandeurs thermodynamiques. Nous décrirons brièvement, dans cette partie, quatre de ces méthodes. Si elles ne donnent, en général, pas accès directement à des informations structurales, leur couplage avec des méthodes théoriques permet de mettre en relation grandeurs thermodynamiques et paramètres structuraux.

#### I.C.1. La méthode d'équilibre<sup>14, 15</sup>

La mesure idéale d'affinité cationique ( $\Delta_r\text{H}^\circ_{\text{cat}}$ ) est la détermination directe de grandeurs thermodynamiques absolues, en étudiant par exemple l'équilibre d'association entre une molécule M et un cation  $\text{Cat}^+$ :



On peut alors déterminer les variations de la constante d'équilibre  $K_{\text{eq}}$  en fonction de la température, ce qui donne accès à  $\Delta_r H^\circ_{\text{cat}}$  et  $\Delta_r S^\circ_{\text{cat}}$ . Par exemple Jarrold *et al.* et Bowers *et al.* ont déterminé par cette technique les énergies de solvatation de plusieurs peptides protonés en phase gazeuse.<sup>16, 17</sup> Cependant dans le cas d'interaction de biomolécules avec des cations métalliques, l'enthalpie standard de réaction de complexation ( $\Delta_r H^\circ_{\text{cat}} \geq 150 \text{ kJ/mol}$ ) est trop importante pour obtenir l'équilibre à des températures usuelles.

Par ailleurs  $K_{\text{eq}}$  va dépendre de l'intensité des pics des ions  $\text{Cat}^+$  et  $M\text{-Cat}^+$  et de la pression de  $M$ . Il est alors nécessaire d'avoir une mesure de la pression partielle de  $M$  précise, ce qui limite son utilisation aux molécules suffisamment volatiles pour lesquelles on peut obtenir une pression constante. Les biomolécules ont une pression de vapeur saturante très faible, de plus elles se dégradent facilement quand on les vaporise. Bien que les techniques d'équilibre ont fait leur preuve dans la détermination d'énergie de liaison<sup>18</sup>, elles se révèlent inefficaces dans le cas de détermination d'énergie de liaison d'une biomolécule avec un cation. On a alors recours à des méthodes de mesure d'affinités relatives. On détermine la différence d'affinité pour un cation de la molécule d'intérêt et d'un composé de référence :  $M + \text{Ref-Cat}^+ \rightleftharpoons M\text{-Cat}^+ + \text{Ref}$ . G. Ohanessian et T. McMahon ont ainsi déterminé une échelle d'affinités pour le sodium de différentes molécules organiques.<sup>19</sup>

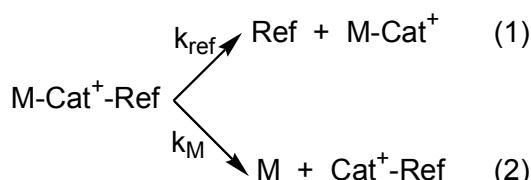
Dans le cas de biomolécules on utilise principalement des méthodes dites cinétiques, dont voici deux exemples.

### I.C.2. La méthode cinétique de Cooks

#### I.C.2.a. Principe<sup>14, 20</sup>

##### *Principe*

Le principe de cette méthode repose sur l'étude de la décomposition (métastable ou induite par collisions) d'un dimère cationisé  $M\text{-Cat}^+\text{-Ref}$  entre la molécule choisie  $M$  et un composé de référence Ref. Il existe alors les deux voies de dissociation suivantes :



La voie (1) correspond à la dissociation de la liaison  $\text{Cat}^+\text{-Ref}$  avec la constante de vitesse  $k_{\text{ref}}$ . Selon la voie (2) on dissocie la liaison  $M\text{-Cat}^+$  avec la constante de vitesse  $k_M$ . Le profil réactionnel thermodynamique est donné sur la figure II-2.

D'après la théorie de l'état de transition, on peut relier le rapport des constantes de vitesse à la différence d'énergie libre d'activation des deux voies de dissociation compétitives selon l'équation suivante :

$$\begin{aligned}\ln(k_M/k_{\text{Ref}}) &= -(\Delta_r G^{\#}_{\text{cat}}(M) - \Delta_r G^{\#}_{\text{cat}}(\text{Ref}))/RT_{\text{eff}} \\ &= (\Delta_r S^{\#}_{\text{cat}}(M) - \Delta_r S^{\#}_{\text{cat}}(\text{Ref}))/R - (\Delta_r H^{\#}_{\text{cat}}(M) - \Delta_r H^{\#}_{\text{cat}}(\text{Ref}))/RT_{\text{eff}}\end{aligned}$$

où R est la constante des gaz parfaits et  $T_{\text{eff}}$  est la température effective de dissociation du dimère.

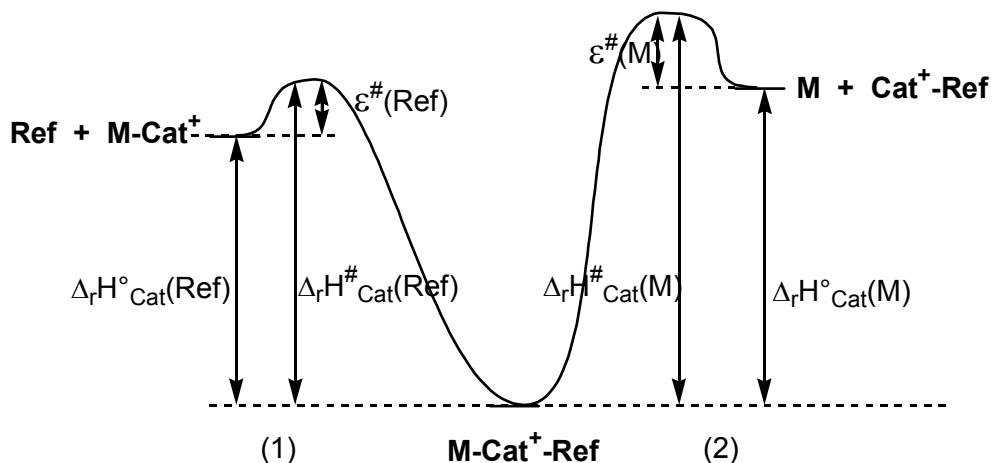


Figure II-2 : Profil réactionnel de la dissociation du dimère  $\text{M-Cat}^+-\text{Ref}$  selon les voies (1) et (2).

Dans le cas de la rupture simple d'une liaison de nature électrostatique, on peut considérer que les enthalpies d'activation des réactions d'association notées  $\varepsilon^{\#}(\text{Ref})$  et  $\varepsilon^{\#}(M)$  sur la figure II-2 sont négligeables. On peut alors considérer que  $\ln(k_M/k_{\text{Ref}}) = (\Delta_r S^{\#}_{\text{cat}}(M) - \Delta_r S^{\#}_{\text{cat}}(\text{Ref}))/R - (\Delta_r H^{\#}_{\text{cat}}(M) - \Delta_r H^{\#}_{\text{cat}}(\text{Ref}))/RT_{\text{eff}}$

Dans une seconde approximation, si la molécule M et le composé de référence Ref ont des structures similaires, c'est à dire qu'ils impliquent le même type de liaison avec le cation, alors on peut considérer que la différence d'entropie d'activation des réactions (1) et (2) est négligeable.

On obtient alors l'équation suivante qui est l'équation classique de la méthode cinétique :

$$\boxed{\ln(k_M/k_{\text{Ref}}) = -(\Delta_r H^{\circ}_{\text{cat}}(M) - \Delta_r H^{\circ}_{\text{cat}}(\text{Ref}))/RT_{\text{eff}} = \ln(I_{\text{M-Cat}^+}/I_{\text{Cat}^+-\text{Ref}})}$$

qui relie le rapport des constantes de vitesse aux affinités des molécules M et Ref pour le cation. Le rapport des constantes de vitesse est déterminé expérimentalement par le rapport des intensités des pics correspondant aux ions  $\text{M-Cat}^+$  et  $\text{Cat}^+-\text{Ref}$ .

#### *Limite*

La technique de détermination d'énergie de liaison par la méthode cinétique de Cooks a fait l'objet de plusieurs critiques, la principale étant l'utilisation de la température effective. En effet, même lorsque les hypothèses ci-dessus sont satisfaites, il reste à déterminer la température

effective, qui n'est pas une grandeur thermodynamique, puisque les ions gazeux ne sont pas en conditions d'équilibre. En pratique, la détermination de multiples affinités relatives à l'intérieur d'une série de composés analogues permet de vérifier la cohérence interne des résultats et de déterminer une valeur raisonnable pour la température effective.

### I.C.2.b. Technique couplée à des calculs de chimie quantique

L'affinité d'une molécule pour un cation peut aussi être obtenue par calcul de chimie quantique en déterminant les propriétés électroniques du complexe, de la molécule nue et du cation pour différentes conformations de la molécule isolée et complexée. Lorsque l'énergie de liaison calculée est en bon accord avec la valeur expérimentale, la conformation du complexe devient un candidat pour la structure de l'ion observé.

### I.C.2.c. Un échelle d'affinités des acides aminés et de Gly<sub>2</sub> pour Na<sup>+</sup><sup>20, 21</sup>

L'équipe de Wesdemiotis a étudié les affinités de plusieurs molécules d'intérêt biologique avec Na<sup>+</sup>, par la méthode cinétique de Cooks. Un premier article est paru en 2003 qui donnait l'affinité pour le sodium de 17 des 20 acides aminés. En 2004, des expériences et calculs complémentaires ont permis de déterminer l'affinité pour Na<sup>+</sup> du dipeptide Gly<sub>2</sub>. Ils ont obtenu expérimentalement une valeur de  $\Delta_r H^\circ_{\text{Na}^+}(\text{Gly}_2) = 203 \pm 8 \text{ kJ/mol}$  compatible avec les valeurs calculées pour deux structures. Elles présentent toutes deux une forme non-zwitterionique avec les deux carbonyles complexés au sodium. Ces deux structures sont très proches en énergie (1 kJ/mol), et diffèrent seulement par l'orientation de l'amine terminale (schéma II-2).

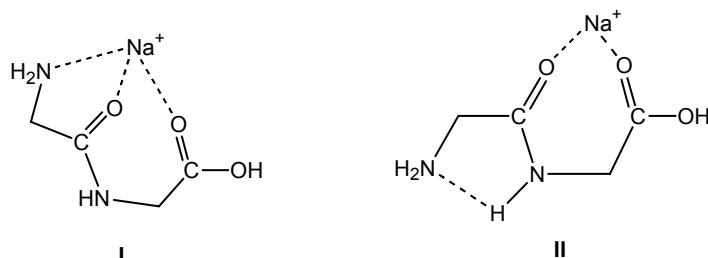


Schéma II-2

L'extension de ce travail à une série de peptides sera décrite au chapitre V.

## I.C.3. Autre exemple : dissociation par radiation infrarouge du corps noir (Blackbody Infrared Radiative Dissociation : BIRD)<sup>22</sup>

### I.C.3.a. Principe

Il est possible d'accéder expérimentalement aux énergies de seuil de dissociation, notamment grâce à la technique de dissociation par radiation du corps noir, initialement mise en place par

l'équipe de T. McMahon. La dissociation unimoléculaire d'un complexe est engendrée par l'absorption de radiations infrarouges dues au rayonnement du corps noir (la cellule, dans laquelle sont piégés les ions) à une température donnée. En étudiant la cinétique de dissociation, on peut en déduire la constante de vitesse à cette température. Cette constante de

vitesse suit la loi d'Arrhenius dans le cas de grands systèmes :  $k = A \exp\left(-\frac{Ea}{RT}\right)$  où A est le

facteur préexponentiel et Ea l'énergie d'activation pour la dissociation. Grâce à cette technique, on a alors accès à Ea qui peut être considérée comme égale à la valeur de  $\Delta_rH^\circ$ , si on fait l'hypothèse que la réaction d'association se produit sans barrière d'activation (cette approximation a déjà été invoquée dans la méthode cinétique de Cooks).

### I.C.3.b. Application à la détermination d'énergies de solvatation<sup>23-25</sup>

Cette technique a été largement utilisée par l'équipe de Williams, sous le nom de BIRD. Nous avons déjà mentionné dans le paragraphe sur la micro-solvatation (Chapitre I III.B.2), qu'ils s'étaient intéressés à la forme zwitterionique ou non-zwitterionique de la valine complexée avec Li<sup>+</sup> et solvatée par plusieurs molécules d'eau. Rappelons que les acides aminés en phase gazeuse se trouvent sous une forme non-zwitterionique. Cependant nous verrons dans le chapitre IV que la complexation avec un cation alcalin permet de stabiliser notamment la forme zwitterionique, comme la micro-solvatation. Ils ont étudié la dissociation unimoléculaire d'une molécule d'eau du complexe Val-Li<sup>+</sup> solvaté. Ils ont montré, en comparant les énergies d'activation de désolvatation de Val-Li<sup>+</sup>(H<sub>2</sub>O)<sub>n</sub>, Bet-Li<sup>+</sup>(H<sub>2</sub>O)<sub>n</sub> (Bet = (CH<sub>3</sub>)<sub>3</sub>N<sup>+</sup>-CH<sub>2</sub>-COO<sup>-</sup> molécule modèle du zwitterion) et AlaOEt-Li<sup>+</sup>(H<sub>2</sub>O)<sub>n</sub> (molécule modèle de la forme sans charge formelle) obtenues par des expériences de BIRD, que Val-Li<sup>+</sup>(H<sub>2</sub>O)<sub>3</sub> présente une forme zwitterionique de la valine.

### I.C.3.c. Technique apparentée

Il existe d'autres outils de dissociation donnant accès à l'énergie d'activation de dissociation que le rayonnement du corps noir. On peut notamment utiliser la technique de dissociation par absorption multiphotonique infrarouge (ou IRMPD), basée sur l'absorption de plusieurs photons infrarouges. On utilise souvent, pour ce faire, un laser à CO<sub>2</sub>, non accordable, à fluence élevée. Marshall *et al.* ont comparé les mécanismes de dissociation par BIRD et IRMPD, ils ont montré que ces mécanismes étaient identiques avec une distribution d'énergie interne de la molécule irradiée équivalente. Ils ont ainsi déterminé l'énergie d'activation pour la dissociation de la bradykinine protonée par IRMPD et ont trouvé un résultat similaire à celui déterminé par BIRD.<sup>26</sup>

### I.C.4. Mesure d'énergie au seuil de dissociation : méthode TCID

#### I.C.4.a. Principe <sup>27</sup>

Une autre technique qui permet d'accéder expérimentalement aux énergies de seuil de dissociation a été développée par P. Armentrout. Les ions M-Cat<sup>+</sup> sont produit par ESI puis transférés dans une cellule où règne une pression de 5 à 10 Torr où ils y sont thermalisés par collision. Ces ions sont ensuite accélérés, dans une zone de faible pression, avec une énergie cinétique connue  $E_{cin}$ , puis dissociés par collision avec un gaz rare introduit à une pression de l'ordre 0.1 mTorr. L'intensité des ions fragment, qui est proportionnelle à la section efficace de réaction de dissociation, est alors dépendante de l'énergie cinétique  $E_{cin}$  apportée aux ions. L'énergie au seuil de dissociation  $E_0$  correspond à l'énergie nécessaire pour casser la liaison M-Cat<sup>+</sup>, en utilisant toute l'énergie de collision et l'énergie thermique du complexe dans le seul mode de dissociation, formant les fragments à 0 K. Pour relier la variation de la section efficace de collision à  $E_0$ , P. Armentrout a proposé la relation suivante :

$$\sigma = \sigma_0 \sum_i g_i \frac{(E_{cin}^{CM} + E_i - E_0)^n}{E_{cin}^{CM}} \propto I_{Cat^+} . \sigma$$
 est la section efficace de la réaction de dissociation, qui

est proportionnelle à l'intensité de l'ion fragment (Cat<sup>+</sup>).  $E_{cin}^{CM}$  est l'énergie cinétique dans le référentiel du centre de masse,  $E_i$  est l'énergie ro-vibrationnelle de l'ion précurseur M-Cat<sup>+</sup> et  $g_i$  son abondance relative à une température donnée.  $n$  et  $\sigma_0$  sont des constantes déterminées de manière empirique. Les valeurs de  $E_0$  peuvent être ensuite converties en enthalpie standard de dissociation  $\Delta_f H^\circ(T)$ , en tenant compte de la correction d'énergie thermique entre 0 K et T, ce terme étant déterminé à l'aide de fréquences calculées ou estimées.

#### I.C.4.b. Limite

Ces auteurs estiment une erreur de 0.1 eV (9.6 kJ/mol) sur les valeurs de  $E_0$ .

#### I.C.4.c. Application à la détermination des énergies d'hydratation de Gly-Na<sup>+</sup> <sup>28</sup>

Armentrout *et al.* ont déterminé les quatre énergies de premières hydratations de Gly-Na<sup>+</sup>, par la méthode de détermination d'énergie de seuil de dissociation par collision. Ils ont observé une décroissance monotone des énergies de liaison Gly-Na<sup>+</sup>-(H<sub>2</sub>O)<sub>n</sub>, pour n allant de 1 à 4. Parallèlement, ils ont effectué une recherche conformationnelle théorique des complexes Gly-Na<sup>+</sup>-(H<sub>2</sub>O)<sub>n</sub>, 1≤n≤4. Ils ont montré que Gly-Na<sup>+</sup>-(H<sub>2</sub>O)<sub>1</sub> avait le même mode de complexation que Gly-Na<sup>+</sup> : par l'amine et le carbonyle de la fonction acide (N,O). A partir de Gly-Na<sup>+</sup>-(H<sub>2</sub>O)<sub>2</sub>, la structure de plus basse énergie présente une complexation du peptide uniquement par le carbonyle (O), les molécules d'eau étant en interaction soit avec Na<sup>+</sup>, soit avec Na<sup>+</sup> et l'acide aminé. La comparaison entre les valeurs expérimentales et les valeurs théoriques montre que

dans tous les complexes Gly-Na<sup>+</sup>-(H<sub>2</sub>O)<sub>n</sub>, la glycine présente une structure non-zwitterionique. Cependant les résultats expérimentaux ne permettent pas de conclure sur le mode de liaison (O) ou (N,O) du sodium.

### I.D. Conclusion

Il existe, aujourd’hui, des techniques de spectrométrie de masse qui permettent d’avoir accès avec plus ou moins de précision à la structure tridimensionnelle et donc à la conformation des molécules biologiques. Nous avons vu que les expériences de cinétique d’échange H/D et de mobilités ioniques permettaient de caractériser des classes de structures : compacte (comme les protéines dans leur conformation native, ou les formes globulaires) et dépliée (comme les protéines dénaturées ou les formes hélicoïdales). Cependant ces techniques ont montré leurs limites notamment dans l’élucidation des formes zwitterionique et non-zwitterionique des acides aminés et des petits peptides complexés au sodium. Nous avons vu aussi que dans le cas des expériences de mobilités ioniques et de mesure d’énergie de liaisons, les calculs, soit de modélisation moléculaire soit de chimie quantique, étaient nécessaires à la détermination structurale. De plus, il est souvent indispensable de coupler plusieurs de ces techniques, pour avoir une réponse univoque sur la conformation de molécules biologiques. Toutes ces méthodes sont des techniques indirectes qui permettent de sonder la structure tridimensionnelle d’ions.

Il existe aujourd’hui une technique directe qui permet d’étudier la structure d’ions en phase gazeuse : la spectroscopie infrarouge.

## II. Etude structurale par spectroscopie infrarouge en phase gazeuse

### II.A. Introduction : spectres IR de molécules biologiques <sup>29</sup>

La spectroscopie vibrationnelle (infrarouge (IR) et Raman) est appliquée depuis la fin des années 60 comme une approche complémentaire aux expériences de spectroscopie de dichroïsme circulaire (CD), dans la détermination des structures secondaires de polypeptides ou protéines en phase condensée. Les techniques de spectroscopie IR et de CD ont un temps caractéristique court (de l’ordre de 10<sup>-13</sup> s), ce qui permet d’étudier les interconversions conformationnelles, contrairement aux expériences de RMN (temps caractéristique 10<sup>2</sup> à 1 s) qui présentent des signaux moyennés. Il a été mis en évidence plusieurs vibrations caractéristiques, en particulier pour la fonction amide :

-amide A correspondant à l'elongation N-H (stretching) entre 3300 et 3450  $\text{cm}^{-1}$

-amide I correspondant à un couplage entre l'elongation C=O, l'elongation C-N et la vibration angulaire (bending) dans le plan C-N-H entre 1610 et 1695  $\text{cm}^{-1}$ .

-amide II correspondant à un couplage entre la vibration angulaire (bending) dans le plan C-N-H et l'elongation C-N entre 1480  $\text{cm}^{-1}$  et 1575  $\text{cm}^{-1}$ .

L'étude de la vibration amide I permet de déterminer l'environnement spatial de la fonction carbonyle (figure II-3). Ainsi l'hélice alpha présente des bandes amides I entre 1648 et 1662  $\text{cm}^{-1}$ , l'hélice  $3_{10}$  entre 1660 et 1666  $\text{cm}^{-1}$ , le feuillet bêta présente deux bandes caractéristiques, l'une entre 1620 et 1640  $\text{cm}^{-1}$ , l'autre entre 1670 et 1690  $\text{cm}^{-1}$  avec un rapport d'intensité de l'ordre de 10 : 1. De même les coudes  $\beta$  et  $\gamma$  ont des fréquences caractéristiques (figure II-3).

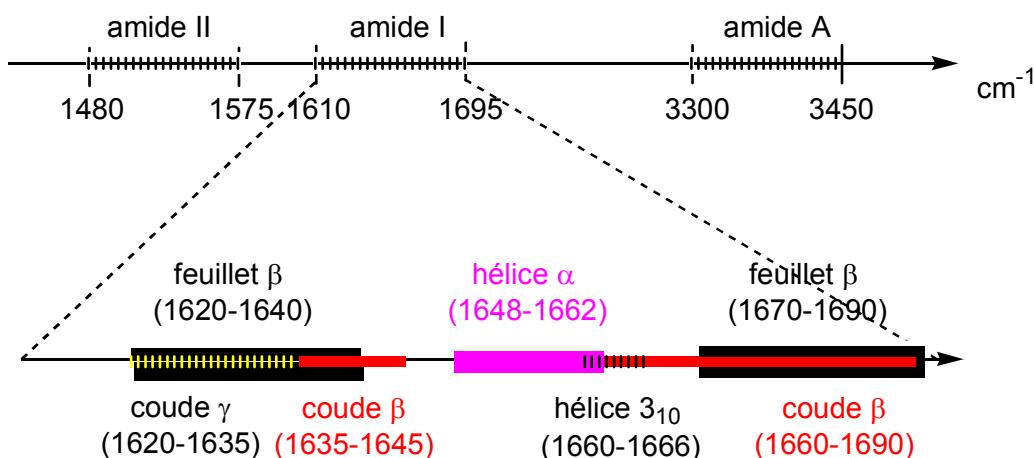


Figure II-3 : Echelle de fréquences des vibrations de la fonction peptidique.

L'interprétation des spectres peut cependant être assez ardue, notamment à cause des recouvrements de bandes amide I correspondant à plusieurs motifs. Les signaux des coudes, souvent moins intenses, peuvent être difficiles à détecter.

L'application des techniques de spectroscopie infrarouge de biomolécules en phase condensée s'est étendue depuis peu à la phase gazeuse.

## II.B. Spectres IR de neutres en matrice inerte

Une expérience alternative consiste à enregistrer les spectres de vibration de molécules isolées en matrice inerte (Ne, Ar ou Kr) à basse température (5.5 et 13K). D'une part, le travail à basse température permet de limiter les interconversions conformationnelles, d'autre part la matrice étant inerte, les interactions entre la molécule et son environnement sont faibles. Ainsi Reva et al. ont étudié le spectre infrarouge de la glycine et de la proline en matrice inerte.<sup>30, 31</sup> Dans le cas de la glycine, ils ont mis en évidence la présence de trois conformères en étudiant les

fréquences observées dans la zone des élongations C=O, les intensités de ces bandes permettant de remonter jusqu'à la population et l'énergie relative de chacun des trois conformères. Ces études expérimentales couplées à des calculs *ab initio* permettent d'identifier la structure des trois conformères formés.

## II.C. Spectres IR de neutres en phase gazeuse

Aujourd'hui, il existe plusieurs groupes, dont ceux de Michel Mons en France, de John Simons en Angleterre ou de Markus Gerhards en Allemagne, qui ont développé des montages basés sur la double résonance à deux photons IR/UV. Les molécules sont produites en phase gazeuse par désorption laser, puis refroidies par expansion dans une détente supersonique. Cette technique permet d'une part de déterminer le nombre de conformères que l'on a formé en phase gazeuse, et d'obtenir la spectroscopie infrarouge de chacun des conformères dans la gamme des vibrations amide A. Ces expériences sont de type pompe-sonde (IR-UV). Un laser infrarouge (souvent un OPO) induit une transition vibrationnelle, ce qui provoque une diminution du nombre de molécules dans l'état fondamental (hole burning), puis ces neutres sont irradiés par un laser UV (la sonde) dont la longueur d'onde est ajustée de manière à ne provoquer que l'ionisation des neutres dans l'état fondamental. Une comparaison entre les spectres UV avec et sans irradiation infrarouge permet d'obtenir le spectre infrarouge du conformère neutre.

L'équipe de Michel Mons s'est récemment intéressée à la mise en évidence de liaisons intramoléculaires C<sub>10</sub> (coudes  $\beta$ ), C<sub>7</sub> (coudes  $\gamma$ ) et C<sub>5</sub> dans les dipeptides Ac-Phe-Gly-NH<sub>2</sub> et Ac-Gly-Phe-NH<sub>2</sub> par spectroscopie infrarouge en sondant la zone amide A. Ils ont mis en évidence la présence de deux conformères pour chacun des peptides. Les conformères prédominants de Ac-Phe-Gly-NH<sub>2</sub> et Ac-Gly-Phe-NH<sub>2</sub> présentent un motif C<sub>5</sub>-C<sub>7</sub> et C<sub>5</sub>-C<sub>5</sub>, respectivement ; quant aux conformères minoritaires ils présentent tous deux la signature d'un coude  $\beta$ .<sup>32</sup> Ils ont récemment étendu cette étude au tripeptide Ac-Phe-Gly-Gly-NH<sub>2</sub>.<sup>33</sup>

## II.D. Spectres IR d'ions en phase gazeuse

### II.D.1. Introduction

Il est aujourd'hui possible d'enregistrer des spectres infrarouges d'ions en phase gazeuse, cependant cette technique est soumise à des contraintes particulières. A cause de la répulsion de charge d'espace, le nombre d'ions que l'on peut placer dans un volume précis est restreint, le faisceau infrarouge ne peut donc interagir qu'avec un nombre limité d'ions et la fraction de rayonnement absorbé est alors trop faible pour être détectée avec précision. C'est pourquoi il

est impossible d'enregistrer un spectre d'absorption IR classique par mesure directe d'absorption de photons. Il est alors nécessaire que le processus d'absorption induise un phénomène physique mesurable, par exemple une fragmentation. L'observation d'un ion fragment sera alors la signature de l'absorption infrarouge. Ceci est d'autant plus intéressant que les techniques de spectrométrie de masse sont, en général, très sensibles. Cependant l'énergie nécessaire à la rupture d'une liaison cation métallique-molécule biologique (typiquement de 80 à 300 kJ/mol dans le cas de nos complexes) est beaucoup plus grande que l'énergie d'un photon infrarouge absorbé (4 à 50 kJ/mol). Il existe deux méthodes pour pallier ce problème.

La première est le « tagging » : au lieu d'étudier l'ion seul, on s'intéresse au complexe de van der Waals formé par l'ion et un atome de gaz rare. Le complexe étant faiblement lié, l'atome de gaz rare ne perturbe pas fortement les modes propres de l'ion. De plus, l'énergie de liaison du complexe est suffisamment faible pour que l'absorption d'un photon provoque sa dissociation. Le spectre IR de l'ion est donc obtenu en suivant le rendement de fragmentation du complexe en fonction de la longueur d'onde du laser. Cette technique a l'avantage d'être une spectroscopie à un photon, donc sans limitation sur la puissance du laser. Cependant la formation de l'adduit ion-gaz rare nécessite l'utilisation d'une détente supersonique et peut dans certains cas modifier le spectre de l'ion.

La seconde technique est aussi une spectroscopie de dissociation, mais on considère, à présent, l'ion isolé. Les énergies de liaison étant beaucoup plus élevées, il est alors nécessaire que l'ion absorbe plusieurs photons pour se dissocier. On parle alors de spectroscopie IRMPD (InfraRed MultiPhoton Dissociation). Elle a l'avantage d'imposer peu de contraintes sur la formation des ions. Cependant elle nécessite l'utilisation d'un laser puissant (avec une grande fluence) et accordable. Il n'existe, aujourd'hui, dans le monde, que trois lasers infrarouges pouvant remplir ces conditions, dont deux en Europe qui ont produit l'essentiel des résultats publiés jusqu'ici, ce sont deux lasers à électrons libres LEL avec une structure temporelle adéquate. L'un se trouve en Hollande et se nomme FELIX, le deuxième, que nous avons utilisé, s'appelle CLIO (Centre Laser Infrarouge d'Orsay) et se trouve en France. Ces deux LEL ont notamment pour caractéristique de permettre de sonder la zone des vibrations amide I et amide II. Nous décrirons les caractéristiques de CLIO dans le chapitre III.

## II.D.2. Principe de l'IRMPD

Nous allons présenter le modèle le plus simple pour décrire le processus d'IRMPD. Au départ la molécule est dans son état fondamental vibrationnel (figure II-4 (a)). Considérons que le laser est accordé sur la longueur d'onde d'un mode particulier. Alors la molécule va absorber un photon, elle va donc passer de l'état vibrationnel  $v=0$  à  $v=1$  et son énergie interne va augmenter

d'un quantum (figure II-4 (b)). Ensuite par redistribution vibrationnelle intramoléculaire (IVR), l'énergie absorbée dans le mode en résonance va se relaxer dans tous les autres modes de vibration (figure II-4 (c)). L'énergie interne de la molécule aura donc augmenté de l'énergie d'un photon, cette énergie se trouve délocalisée dans tous les modes de vibration donc peu probablement dans celui qui est en résonance, qu'on peut donc considérer être retourné dans l'état  $v=0$ . Un second photon peut alors être absorbé, l'énergie interne de la molécule augmente, puis par IVR cette énergie est redistribuée dans tous les autres modes de vibration (figure II-4 (e)). La molécule peut alors absorber un troisième photon et ainsi de suite, jusqu'à ce que l'énergie interne de la molécule atteigne la limite de dissociation. Par IRMPD, la molécule est chauffée vibrationnellement, cela conduit aux dissociations de plus basse énergie. Dans le cas des complexes avec le sodium, cette dissociation de plus basse énergie correspond à la rupture de la liaison métal-ligand, ce qui correspond à une énergie de liaison de l'ordre de 200 kJ/mol. Regardons quelques ordres de grandeur : pour une énergie de photon de  $1750 \text{ cm}^{-1}$  (élongation C=O), il faut l'absorption d'au moins 10 photons pour provoquer la dissociation.

Pour que la fragmentation de l'ion ait lieu, il est nécessaire que l'IVR soit efficace. Or la vitesse de l'IVR augmente avec la densité d'états vibrationnels, qui elle-même croît avec l'énergie interne de l'ion et donc avec le nombre de photons absorbés. Le processus d'IVR sera donc moins efficace pour les premiers photons absorbés. Par ailleurs la densité d'états vibrationnels sera d'autant plus importante que la molécule sera grande, l'IVR sera donc plus rapide pour de grandes molécules. De plus ce processus sera d'autant plus rapide que les modes de vibration seront plus fortement couplés.<sup>34</sup>

Ce modèle est le plus simple ; pour rendre compte des observations expérimentales, il faut également prendre en compte l'anharmonicité croisée des modes de vibration. Cette anharmonicité croisée est due au couplage entre les différents modes et l'une de ces caractéristiques est que plus l'énergie interne de la molécule croît, plus l'énergie de la transition  $v=0 \rightarrow v=1$  diminue, on parle alors de décalage vers le rouge (ou de redshift) des bandes en fonction de l'énergie interne.<sup>35</sup> On peut comprendre analytiquement l'origine de cette anharmonicité en faisant un développement de Taylor à l'ordre 2 de l'énergie vibrationnelle :

$$E(v_i) = \underbrace{\sum_i \omega_i (v_i + 1/2)}_{\text{Harmonicité d'un mode}} + \underbrace{\sum_i x_{ii} (v_i + 1/2)^2}_{\text{Anharmonicité d'un mode}} + \underbrace{\sum_{i < j} x_{ij} (v_i + 1/2)(v_j + 1/2)}_{\text{Anharmonicité croisée entre deux modes}}$$

Le premier terme correspond au potentiel harmonique, le second à l'anharmonicité d'un mode ( $x_{ii}$ , constante d'anharmonicité du mode  $i$ , est négative), le dernier terme correspond à l'anharmonicité croisée entre les mode  $i$  et  $j$  ( $x_{ij}$ , constante d'anharmonicité croisée entre les modes  $i$  et  $j$ , est négative).

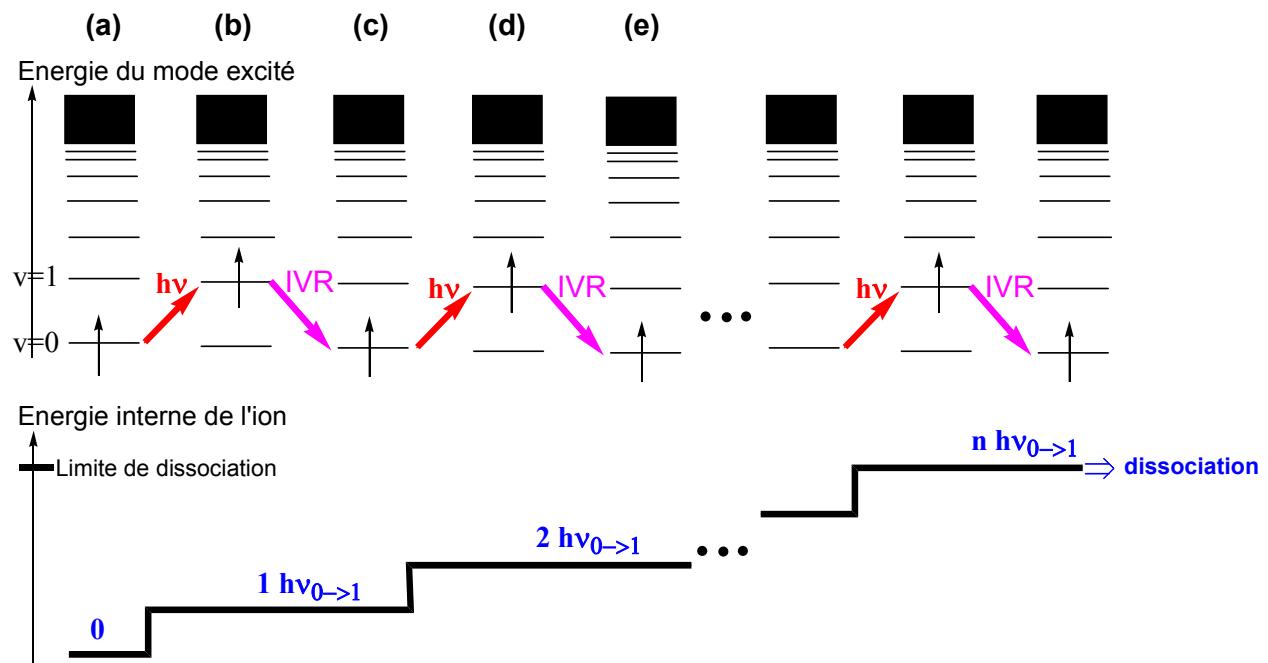


Figure II-4 : Principe de la spectroscopie IRMPD.

### II.D.3. Tracé d'un spectre IRMPD

Comme nous l'avons vu précédemment, le processus d'IRMPD provoque la dissociation de l'ion irradié. Expérimentalement, on piège l'ion dont on veut obtenir le spectre infrarouge, dans notre cas dans la cellule du spectromètre de masse FT-ICR nommé MICRA. Puis on irradie avec le LEL cet ion, à un nombre d'onde précis, pendant un temps pouvant varier de 500 ms à plusieurs secondes. On observe ensuite le nombre d'ions fragments formés normalisé au nombre total d'ions, ou la décroissance du signal de l'ion parent normalisée au signal de l'ion parent non irradié. Enfin il faut réitérer ces différentes étapes en faisant varier le nombre d'onde des photons du LEL, puis on trace le spectre IRMPD point par point. Une fois le spectre IRMPD tracé, on l'interprète en le comparant aux spectres théoriques de chaque isomère.

## II.D.4. Interprétation d'un spectre IRMPD : comparaisons aux spectres IR théoriques

### II.D.4.a. Spectres IR théoriques

Pour interpréter les bandes du spectre IRMPD, il est nécessaire de les comparer à celles des spectres infrarouges théoriques des différents isomères ou conformères les plus bas en énergie de l'ion considéré. Aujourd'hui, on peut aisément simuler des spectres d'absorption à 1 photon, dans l'approximation harmonique, par les méthodes de chimie quantique habituelles (HF, MP2, B3LYP ou BLYP entre autres).

De nombreuses études ont comparé la précision des différentes méthodes quantiques pour le calcul des fréquences et des intensités des absorptions à 1 photon. De manière générale, les fréquences calculées sont surestimées, ceci étant essentiellement du à l'approximation harmonique. Cette surestimation peut être corrigée à l'aide de facteurs d'échelles ou scaling factors. Un scaling factor standard déterminé sur un grand échantillon de petites molécules organiques a été calculé selon la méthode et la base. Le tableau II-2 présente les valeurs des scaling factors, l'erreur quadratique moyenne sur les fréquences calculées par rapport aux valeurs expérimentales<sup>36</sup> et l'incertitude moyenne absolue sur les intensités calculées par rapport à celles calculées avec le plus haut niveau de calcul *ab initio* (QCIDS/6-311+G(3df,3pd))<sup>37</sup>, selon différentes méthodes. Il en ressort que la méthode B3LYP/6-31G\* est la plus précise pour simuler un spectre IR à un photon, puis MP2/6-31G\* et enfin BLYP/6-31G\*.<sup>38</sup>

	B3LYP/6-31G*	BLYP/6-31G*	MP2(full)/6-31G*
Scaling factor standard <sup>36</sup>	0.9614	0.9945	0.9427
Erreur quadratique moyenne sur les fréquences (cm <sup>-1</sup> ) <sup>36</sup>	34	45	61
Incertitude moyenne absolue sur les intensités (km.mol <sup>-1</sup> ) <sup>37</sup>	20-30	20-50	intermédiaire

**Tableau II-1** : scaling factors et incertitude absolue des fréquences selon la méthode et la base utilisée.

D'autres méthodes plus sophistiquées prennent en compte l'anharmonicité plus grande des modes de vibration de fréquences supérieures à 1800 cm<sup>-1</sup> (modes d'elongation X-H), en donnant deux scaling factors selon la gamme de longueur d'onde.<sup>38, 39</sup>

Enfin de nombreux articles sur la détermination structurale d'ions par IRMPD montrent une variété de scaling factors allant de 0.96 à 0.98<sup>40</sup> pour la méthode B3LYP/6-31G\*.

Par ailleurs, afin de faciliter la comparaison entre spectres théoriques et expérimentaux, chaque

bande théorique a été convoluée par un profil lorentzien :  $I(\tilde{\nu}) = \frac{I(\tilde{\nu}_0)}{1 + \left( \frac{\tilde{\nu} - \tilde{\nu}_0}{\Delta\sigma/2} \right)^2}$  avec une largeur

à mi-hauteur de  $\Delta\sigma=50$  cm<sup>-1</sup>, correspondant à l'ordre de grandeur de la largeur des pics expérimentaux.

#### II.D.4.b. Limite

Actuellement, il est impossible de simuler un spectre IRMPD. Il est possible d'accéder aux fréquences d'absorption en simulant un spectre d'absorption infrarouge à 1 photon, dans l'approximation harmonique, par les méthodes de chimie quantique habituelles (MP2, B3LYP ou BLYP). Cependant il peut y avoir des déviations entre les intensités relatives des pics théoriques d'absorption à 1 photon et celles des bandes d'absorption multiphotoniques expérimentales. Notamment l'effet de l'anharmonicité croisée et l'efficacité du processus IVR peuvent modifier les intensités relatives de deux absorptions proches. Une modélisation des spectres IRMPD est en cours et fait l'objet de la thèse de F. Boyrie, au Laboratoire de Chimie Physique d'Orsay.

### II.D.5. Quelques résultats

#### II.D.5.a. Complexes du benzène et benzène protoné

Plusieurs groupes se sont intéressés à la structure du benzène protoné<sup>41</sup> ou complexé par les cations métalliques V<sup>+</sup><sup>42</sup> et Al<sup>+</sup><sup>43</sup>. La coordination peut être soit de type  $\sigma$ , soit de type  $\pi$ . De plus dans le cas du vanadium, élément de transition, il existe plusieurs états de spin de basse énergie (quintet ou triplet). L'étude par spectroscopie IRMPD a mis en évidence que la structure du benzène protoné est de type  $\sigma$ , alors que les complexes C<sub>6</sub>H<sub>6</sub>-M<sup>+</sup> mettent en jeu une complexation de type  $\pi$ . Par contre la spectroscopie IRMPD n'a pas permis d'identifier, avec certitude, l'état de spin du vanadium dans le complexe C<sub>6</sub>H<sub>6</sub>-V<sup>+</sup>.

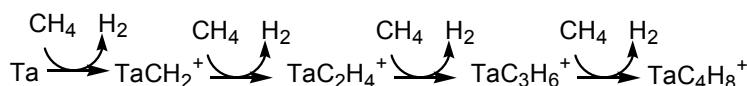
#### II.D.5.b. Différenciation des modes de coordination d'éthers sur Fe<sup>+</sup><sup>44</sup>

S. Le Caër et al. ont déterminé la structure de Fe(MeOMe)<sub>2</sub><sup>+</sup>, Fe(MeOEtOMe)<sup>+</sup> et Fe(MeOEtOMe)<sub>2</sub><sup>+</sup>, formés par une réaction d'échange de ligand sur Fe(CO)<sub>n</sub><sup>+</sup>, par spectroscopie IRMPD. Le spectre IRMPD enregistré dans le cas de Fe(MeOMe)<sub>2</sub><sup>+</sup> montre un accord excellent,

autant pour les fréquences que les intensités d'absorption, avec le spectre simulé au niveau B3LYP/6-31G\*, pour le conformère où les deux ligands sont en position trans (symétrie  $D_{2d}$ ) et le fer dans un état de spin quartet. Le spectre de  $\text{Fe}(\text{MeOEtOMe})^+$  présente les caractéristiques d'une chélation bidentate ( $\eta_2$ ) du ligand, cependant l'état de spin quartet ou sextet du fer n'est pas discernable. Enfin dans le cas du complexe  $\text{Fe}(\text{MeOEtOMe})_2^+$ , le spectre IRMPD, comparé aux spectres IR calculés pour les trois complexes ( $\eta_1, \eta_1$ ), ( $\eta_1, \eta_2$ ) et ( $\eta_2, \eta_2$ ), montre que chacun des deux ligands présente une structure bidentate ( $\eta_2, \eta_2$ ).

### II.D.5.c. Caractérisation d'intermédiaires de réaction.

Dans le cadre de sa thèse, A. Simon a étudié le mécanisme séquentiel de déshydrogénération de quatre molécules de méthane activée par le cation  $\text{Ta}^+$ .<sup>45</sup>



Tous les intermédiaires réactionnels ont notamment été caractérisés par IRMPD.<sup>46</sup> L'étude de l'intermédiaire  $\text{TaC}_2\text{H}_4^+$  a mis en évidence la structure de type  $\text{Ta}(\text{H})_2(\text{C}_2\text{H}_2)^+$  mettant en jeu deux liaisons métal-hydrique. Ceci permet d'expliquer que cet ion réagit successivement avec deux molécules de méthane par un mécanisme faisant intervenir deux réactions consécutives de métathèse mettant en jeu les liaisons  $\text{Ta}^+-\text{H}$  et  $\text{C}-\text{H}$ .

### II.D.6. Conclusion

La technique de spectroscopie IRMPD couplée à des simulations de spectres infrarouges a montré qu'elle était très performante dans la caractérisation de complexes organométalliques. Nous avons décidé d'utiliser cette technique, en un premier temps, à la caractérisation structurale de complexes sodiés d'acides aminés. Nous avons ainsi obtenu les premiers spectres infrarouges d'ions biologiques en phase gazeuse (chapitre IV). Cette technique ayant été efficace pour déterminer la structure de complexes acide aminé- $\text{Na}^+$ , nous l'avons étendu à des complexes de di- et tri-peptides sodiés. Ces résultats seront présentés dans le chapitre V.

## Bibliographie

- [1] Campbell, S.; Rodgers, M. T.; Marzluff, E. M.; Beauchamp, J. L., Deuterium exchange reactions as a probe of biomolecule structure. Fundamental studies of gas phase H/D exchange reactions of protonated glycine oligomers with D<sub>2</sub>O, CD<sub>3</sub>OD, CD<sub>3</sub>CO<sub>2</sub>D, and ND<sub>3</sub>. *J. Am. Chem. Soc.* **1995**, 117, (51), 12840-12854.
- [2] He, F.; Marshall, A. G.; Freitas, M. A., Assignment of gas-phase dipeptide amide hydrogen exchange rate constants by site-specific substitution: GlyGly. *J. Phys. Chem. B* **2001**, 105, (11), 2244-2249.
- [3] Evans, S. E.; Lueck, N.; Marzluff, E. M., Gas phase hydrogen/deuterium exchange of proteins in an ion trap mass spectrometer. *Int. J. Mass Spectrom.* **2003**, 222, (1-3), 175-187.
- [4] Cox, H. A.; Julian, R. R.; Lee, S. W.; Beauchamp, J. L., Gas-phase H/D exchange of sodiated glycine oligomers with ND<sub>3</sub>: Exchange kinetics do not reflect parent ion structures. *J. Am. Chem. Soc.* **2004**, 126, (20), 6485-6490.
- [5] Wyttenbach, T.; Paizs, B.; Barran, P.; Breci, L.; Liu, D. F.; Suhai, S.; Wysocki, V. H.; Bowers, M. T., The effect of the initial water of hydration on the energetics, structures, and H/D exchange mechanism of a family of pentapeptides: An experimental and theoretical study. *J. Am. Chem. Soc.* **2003**, 125, (45), 13768-13775.
- [6] Katta, V.; Chait, B. T., Hydrogen-Deuterium Exchange Electrospray-Ionization Mass-Spectrometry - a Method for Probing Protein Conformational-Changes in Solution. *J. Am. Chem. Soc.* **1993**, 115, (14), 6317-6321.
- [7] Suckau, D.; Shi, Y.; Beu, S. C.; Senko, M. W.; Quinn, J. P.; Wampler, F. M.; McLafferty, F. W., Coexisting Stable Conformations of Gaseous Protein Ions. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, 90, (3), 790-793.
- [8] McLafferty, F. W.; Guan, Z. Q.; Haupt, U.; Wood, T. D.; Kelleher, N. L., Gaseous conformational structures of cytochrome c. *J. Am. Chem. Soc.* **1998**, 120, (19), 4732-4740.
- [9] Jarrold, M. F., Drift-Tube Studies of Atomic Clusters. *J. Phys. Chem.* **1995**, 99, (1), 11-21.
- [10] Wyttenbach, T.; vonHelden, G.; Batka, J. J.; Carlat, D.; Bowers, M. T., Effect of the long-range potential on ion mobility measurements. *J. Am. Soc. Mass Spectrom.* **1997**, 8, (3), 275-282.
- [11] Wyttenbach, T.; Witt, M.; Bowers, M. T., On the stability of amino acid zwitterions in the gas phase: The influence of derivatization, proton affinity, and alkali ion addition. *J. Am. Chem. Soc.* **2000**, 122, (14), 3458-3464.

- [12] Hudgins, R. R.; Jarrold, M. F., Helix formation in unsolvated alanine-based peptides: Helical monomers and helical dimers. *J. Am. Chem. Soc.* **1999**, 121, (14), 3494-3501.
- [13] Hudgins, R. R.; Mao, Y.; Ratner, M. A.; Jarrold, M. F., Conformations of Gly<sub>n</sub>H<sup>+</sup> and Ala<sub>n</sub>H<sup>+</sup> peptides in the gas phase. *Biophys. J.* **1999**, 76, (3), 1591-1597.
- [14] Ervin, K. M., Experimental techniques in gas-phase ion thermochemistry. *Chem. Rev.* **2001**, 101, (2), 391-444.
- [15] McMahon, T. B., Thermochemical ladders: Scaling the ramparts of gaseous ion energetics. *Int. J. Mass Spectrom.* **2000**, 200, (1-3), 187-199.
- [16] Liu, D. F.; Wyttenbach, T.; Barran, P. E.; Bowers, M. T., Sequential hydration of small protonated peptides. *J. Am. Chem. Soc.* **2003**, 125, (28), 8458-8464.
- [17] Kohtani, M.; Jarrold, M. F., The initial steps in the hydration of unsolvated peptides: Water molecule adsorption on alanine-based helices and globules. *J. Am. Chem. Soc.* **2002**, 124, (37), 11148-11158.
- [18] Bogdanov, B.; McMahon, T. B., Structures, thermochemistry, and infrared spectra of chloride ion-fluorinated acetone complexes and neutral fluorinated acetones in the gas phase: experiment and theory. *Int. J. Mass Spectrom.* **2002**, 219, (3), 593-613.
- [19] McMahon, T. B.; Ohanessian, G., An experimental and ab initio study of the nature of the binding in gas-phase complexes of sodium ions. *Chem.-Eur. J.* **2000**, 6, (16), 2931-2941.
- [20] Kish, M. M.; Wesdemiotis, C.; Ohanessian, G., The sodium ion affinity of glycylglycine. *J. Phys. Chem. B* **2004**, 108, (9), 3086-3091.
- [21] Kish, M. M.; Ohanessian, G.; Wesdemiotis, C., The Na<sup>+</sup> affinities of alpha-amino acids: side-chain substituent effects. *Int. J. Mass Spectrom.* **2003**, 227, (3), 509-524.
- [22] Dunbar, R. C.; McMahon, T. B., Activation of unimolecular reactions by ambient blackbody radiation. *Science* **1998**, 279, (5348), 194-197.
- [23] Lemoff, A. S.; Williams, E. R., Binding energies of water to lithiated valine: Formation of solution-phase structure in vacuo. *J. Am. Soc. Mass Spectrom.* **2004**, 15, (7), 1014-1024.
- [24] Lemoff, A. S.; Bush, M. F.; Williams, E. R., Binding energies of water to sodiated valine and structural isomers in the gas phase: The effect of proton affinity on zwitterion stability. *J. Am. Chem. Soc.* **2003**, 125, (44), 13576-13584.
- [25] Jockusch, R. A.; Lemoff, A. S.; Williams, E. R., Effect of metal ion and water coordination on the structure of a gas-phase amino acid. *J. Am. Chem. Soc.* **2001**, 123, (49), 12255-12265.
- [26] Freitas, M. A.; Hendrickson, C. L.; Marshall, A. G., Gas phase activation energy for unimolecular dissociation of biomolecular ions determined by focused RAdiation for gaseous multiphoton ENergy transfer (FRAGMENT). *Rapid Commun. Mass Spectrom.* **1999**, 13, (15), 1639-1642.
- [27] Armentrout, P. B., Kinetic energy dependence of ion-molecule reactions: Guided ion beams and threshold measurements. *Int. J. Mass Spectrom.* **2000**, 200, (1-3), 219-241.

- [28] Ye, S. J.; Moision, R. M.; Armentrout, P. B., Sequential bond energies of water to sodium glycine cation. *Int. J. Mass Spectrom.* **2005**, 240, (3), 233-248.
- [29] Vass, E.; Hollosi, M.; Besson, F.; Buchet, R., Vibrational spectroscopic detection of beta- and gamma-turns in synthetic and natural peptides and proteins. *Chem. Rev.* **2003**, 103, (5), 1917-1954.
- [30] Stepanian, S. G.; Reva, I. D.; Radchenko, E. D.; Rosado, M. T. S.; Duarte, M.; Fausto, R.; Adamowicz, L., Matrix-isolation infrared and theoretical studies of the glycine conformers. *J. Phys. Chem. A* **1998**, 102, (6), 1041-1054.
- [31] Stepanian, S. G.; Reva, I. D.; Radchenko, E. D.; Adamowicz, L., Conformers of nonionized proline. Matrix-isolation infrared and post-Hartree-Fock ab initio study. *J. Phys. Chem. A* **2001**, 105, (47), 10664-10672.
- [32] Chin, W.; Dognon, J. P.; Piuzzi, F.; Tardivel, B.; Dimicoli, I.; Mons, M., Intrinsic folding of small peptide chains: Spectroscopic evidence for the formation of beta-turns in the gas phase. *J. Am. Chem. Soc.* **2005**, 127, (2), 707-712.
- [33] Chin, W.; Compagnon, I.; Dognon, J. P.; Canuel, C.; Piuzzi, F.; Dimicoli, I.; von Helden, G.; Meijer, G.; Mons, M., Spectroscopic evidence for gas-phase formation of successive beta-turns in a three-residue peptide chain. *J. Am. Chem. Soc.* **2005**, 127, (5), 1388-1389.
- [34] Nesbitt, D. J.; Field, R. W., Vibrational energy flow in highly excited molecules: Role of intramolecular vibrational redistribution. *J. Phys. Chem.* **1996**, 100, (31), 12735-12756.
- [35] Herzberg, G., Molecular Spectra and Molecular Structure: Infrared and Raman of Polyatomic Molecules, *Hardcover*. **1991**, 200-209.
- [36] Scott, A. P.; Radom, L., Harmonic vibrational frequencies: An evaluation of Hartree-Fock, Moller-Plesset, quadratic configuration interaction, density functional theory, and semiempirical scale factors. *J. Phys. Chem.* **1996**, 100, (41), 16502-16513.
- [37] Halls, M. D.; Schlegel, H. B., Comparison of the performance of local, gradient-corrected, and hybrid density functional models in predicting infrared intensities. *J. Chem. Phys.* **1998**, 109, (24), 10587-10593.
- [38] Halls, M. D.; Velkovski, J.; Schlegel, H. B., Harmonic frequency scaling factors for Hartree-Fock, S-VWN, B-LYP, B3-LYP, B3-PW91 and MP2 with the Sadlej pVTZ electric property basis set. *Theor. Chem. Acc.* **2001**, 105, (6), 413-421.
- [39] Bauschlicher, C. W.; Langhoff, S. R., The calculation of accurate harmonic frequencies of large molecules: The polycyclic aromatic hydrocarbons, a case study. *Spectroc. Acta Pt. A-Molec. Biomolec. Spectr.* **1997**, 53, (8), 1225-1240.
- [40] Maitre, P.; Le Caer, S.; Simon, A.; Jones, W.; Lemaire, J.; Mestdagh, H. N.; Heninger, M.; Mauclaire, G.; Boissel, P.; Prazeres, R.; Glotin, F.; Ortega, J. M., Ultrasensitive spectroscopy of ionic reactive intermediates in the gas phase performed with the first coupling of

- an IR FEL with an FTICR-MS. *Nucl. Instrum. Methods Phys. Res. Sect. A-Accel. Spectrom. Dect. Assoc. Equip.* **2003**, 507, (1-2), 541-546.
- [41] Jones, W.; Boissel, P.; Chiavarino, B.; Crestoni, M. E.; Fornarini, S.; Lemaire, J.; Maitre, P., Infrared fingerprint of protonated benzene in the gas phase. *Angew. Chem.-Int. Edit.* **2003**, 42, (18), 2057-2059.
- [42] van Heijnsbergen, D.; von Helden, G.; Meijer, G.; Maitre, P.; Duncan, M. A., Infrared spectra of gas-phase V<sup>+</sup>-(benzene) and V<sup>+</sup>-(benzene)<sub>2</sub> complexes. *J. Am. Chem. Soc.* **2002**, 124, (8), 1562-1563.
- [43] van Heijnsbergen, D.; Jaeger, T. D.; von Helden, G.; Meijer, G.; Duncan, M. A., The infrared spectrum of Al<sup>+</sup>-benzene in the gas phase. *Chem. Phys. Lett.* **2002**, 364, (3-4), 345-351.
- [44] Le Caer, S.; Heninger, M.; Lemaire, J.; Boissel, P.; Maitre, P.; Mestdagh, H., Structural characterization of selectively prepared cationic iron complexes bearing monodentate and bidentate ether ligands using infrared photodissociation spectroscopy. *Chem. Phys. Lett.* **2004**, 385, (3-4), 273-279.
- [45] Simon, A.; MacAleese, L.; Boissel, P.; Maitre, P., Towards the characterization of the mechanism of the sequential activation of four methane molecules by Ta<sup>+</sup>. *Int. J. Mass Spectrom.* **2002**, 219, (3), 457-473.
- [46] Simon, A.; MacAleese, L.; Maitre, P., Résultats non publiés. **2002**.

# **Chapitre III :**

## **Méthodes expérimentales et**

## **théoriques**

## Introduction

Ce chapitre présente les différentes méthodes expérimentales et théoriques qui ont été utilisées. Il permet notamment de donner les différentes caractéristiques des montages CLIO : Laser à Electrons Libres (LEL), MICRA (spectromètre de masse FT-ICR) et le couplage de ces deux machines, puis de décrire les caractéristiques originales d'un des codes de calcul que nous avons utilisé, appelé TURBOMOLE.

## I. Le Laser à électrons libres CLIO<sup>1, 2</sup>

### I.A. Introduction

Le Laser à Electrons Libre (LEL) a été mis au point en 1978. Contrairement aux lasers habituels, le LEL peut, en principe, amplifier n'importe quelle longueur d'onde. Cependant, il nécessite l'emploi d'un accélérateur d'électrons qui produit des radiations ionisantes dont il faut se protéger, ce qui rend son installation relativement lourde.

Le LEL est une source de lumière cohérente capable de fournir de fortes puissances crêtes dans un large domaine spectral (des ondes millimétriques à l'ultraviolet). Cependant les contraintes sur l'énergie des électrons, la qualité du faisceau et les miroirs deviennent de plus en plus sévères quand la longueur d'onde diminue. Dans l'infrarouge, la technique du LEL est bien au point.

### I.B. Principe du LEL

Son fonctionnement est basé sur le rayonnement synchrotron délivré par une particule chargée, ici des électrons de haute énergie, en mouvement non rectiligne uniforme. Expérimentalement, des électrons sont produits par un canon (source thermoélectronique) puis accélérés à une énergie comprise entre 16 et 50 MeV. Ils pénètrent dans la cavité optique du laser où règne un champ magnétique périodique modulé spatialement par un onduleur (cf. Figure III-1). Les électrons en interaction avec l'onde magnétique acquièrent une vitesse transversale sinusoïdale, et émettent un rayonnement synchrotron dont la longueur d'onde est fonction du champ magnétique. Ce rayonnement est stocké dans la cavité optique puis amplifié à chaque nouvelle interaction avec le faisceau d'électrons. Un réglage de la longueur de la cavité est nécessaire afin de mettre en phase l'onde stockée et l'onde émise par le faisceau d'électrons. Ceci permet d'atteindre une puissance crête très importante.

A la différence d'un laser traditionnel où la longueur d'onde est fixée par les niveaux d'énergie du milieu amplificateur, la longueur d'onde d'un LEL varie continûment avec l'énergie des électrons et avec l'amplitude du champ magnétique. Ainsi le LEL possède une grande accordabilité de 100 à 3000 cm<sup>-1</sup> pour CLIO et une forte puissance crête de l'ordre de 80 MW à CLIO.

En pratique, l'énergie des électrons est d'abord fixée, définissant une gamme d'énergie de photon accessible. On fait ensuite varier la longueur d'onde en faisant varier

l'entrefer de l'onduleur et donc le champ magnétique, et la puissance laser en optimisant la longueur de la cavité.

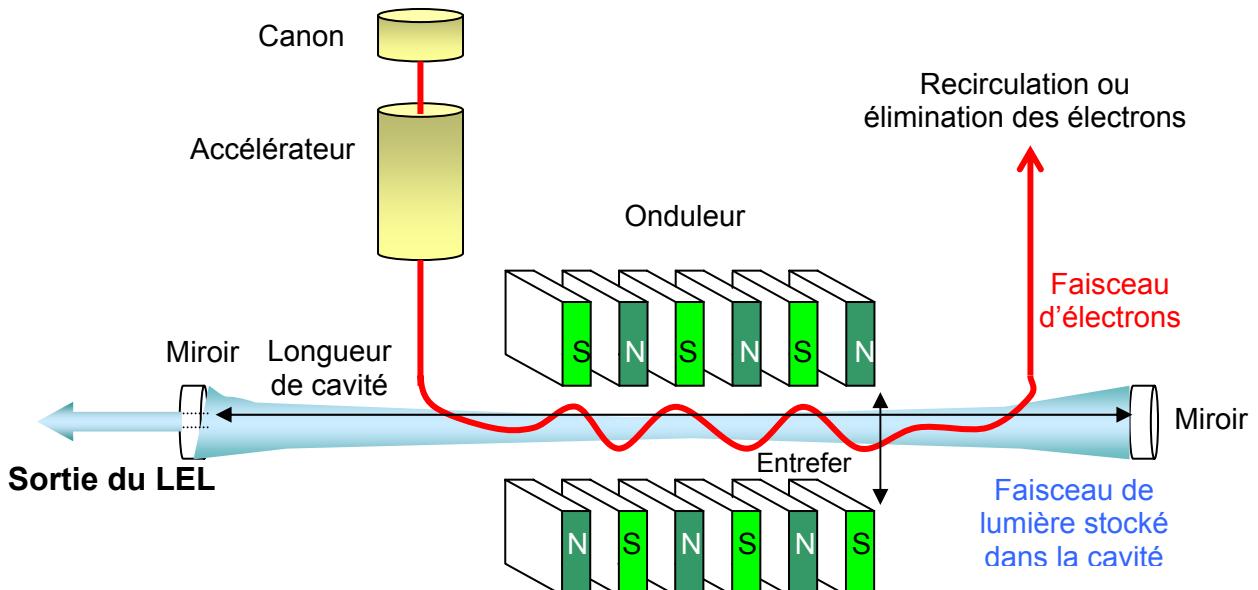


Figure III-1 : Schéma de principe du LEL CLIO.

### I.C. Structure temporelle du LEL

Les lasers à électrons libres sont des lasers pulsés dont la structure temporelle est intimement liée au mode d'émission des électrons à la sortie de l'accélérateur. Les électrons sont accélérés par un champ radiofréquence (RF) qui les regroupe en paquets. La fréquence du champ RF détermine l'espacement entre les paquets d'électrons, et donc l'écart entre deux impulsions émises par paquet d'électrons. La Figure III-2 donne la structure temporelle de CLIO. Elle est constituée de macro-impulsions d'une durée de 10  $\mu$ s. Le temps entre deux macro-impulsions est de 40 ms. Chaque macro-impulsion est constituée d'environ 600 micro-impulsions qui durent environ 1 ps chacune et sont espacées de 16 ns.

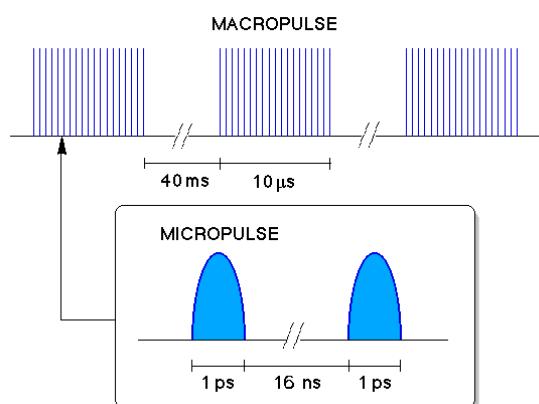


Figure III-2 : Structure temporelle de CLIO.

## I.D. Principales caractéristiques du LEL CLIO

Le tableau III-1 donne un résumé des principales caractéristiques de CLIO.

Caractéristiques	Valeurs
Puissance (énergie) d'une micro-impulsion	50 MW (50μJ)
Puissance (énergie) d'une macro-impulsion	3 kW (25 mJ)
Puissance crête	80 MW
Puissance moyenne disponible	1 W
Nombres d'onde accessibles	100-3000 cm <sup>-1</sup>
Résolution : largeur à mi-hauteur	$\Delta\lambda/\lambda \approx 0.5\%$ soit 5-10 cm <sup>-1</sup> dans la gamme 1000-2000 cm <sup>-1</sup>
Puissance stable à ± 10%	$\Delta\lambda/\lambda = 2.5$

Tableau III-1 : Principales caractéristiques de CLIO.

## II.Le spectromètre de masse MICRA<sup>3</sup>

### II.A. Introduction

Un laser à électrons libres a une architecture assez imposante : la longueur de l'accélérateur est de 20 m et 4.8 m pour la cavité optique. L'utilisation d'un montage mobile facilite fortement le couplage avec le LEL. Dans le cadre de nos expériences, nous avons été amenés à utiliser MICRA (*Mobile ICR Analyser*), un spectromètre de masse de résonance cyclotronique ionique à transformée de Fourier, construit au Laboratoire de Chimie Physique (LCP) d'Orsay par l'équipe « Ions en Phase Gazeuse ». Dans cette partie, nous décrirons le principe de fonctionnement de MICRA, ses caractéristiques et les aménagements spécifiques donnant accès à la spectroscopie infrarouge d'ions biologiques en phase gazeuse.

### II.B. Bref rappel sur la spectrométrie de masse FT-ICR<sup>4, 5</sup>

#### II.B.1. Mouvement cyclotron

Une particule chargée, plongée dans un champ magnétique uniforme  $\vec{B}$ , décrit un mouvement hélicoïdal. Ce mouvement résulte de la combinaison d'un mouvement circulaire, dit mouvement cyclotron de fréquence  $\omega_c = q * B / m$ , où  $q$  et  $m$  désignent respectivement la charge et la masse de la particule, et d'une translation le long de l'axe de  $\vec{B}$  due à la vitesse initiale des ions dans la zone soumise au champ. Chaque ion plongé dans un champ  $\vec{B}$

statique a donc une fréquence cyclotron caractéristique qui ne dépend que de son rapport  $q/m$ .

## **II.B.2. Principe de la FTICR**

La mesure de  $\omega_c$  permet de déterminer le rapport  $q/m$  ; en connaissant  $q$ , on peut alors déduire  $m$  et donc identifier un composé. Le principe de la spectrométrie de masse ICR repose sur l'excitation résonnante de cette fréquence. Du point de vue expérimental, il faut donc mesurer  $\omega_c$ . On se heurte à plusieurs problèmes. D'une part, il faut arriver à confiner les ions dans la cellule de mesure, d'autre part détecter un signal électrique, qui comporte la fréquence  $\omega_c$ .

### **II.B.2.a. Confinement des ions**

Les ions sont introduits dans une cellule placée dans un aimant, qui peut être supraconducteur, un électro-aimant classique ou un aimant permanent. Il s'agit d'un aimant permanent structuré dans le cas de l'expérience MICRA. Cette cellule étant plongée dans un champ magnétique  $\vec{B}$ , les ions acquièrent un mouvement hélicoïdal qui permet de les confiner dans les deux directions perpendiculaires au champ. Cependant pour maintenir ces ions dans un volume restreint, il faut limiter la translation selon l'axe de  $\vec{B}$ . La solution est d'appliquer un champ électrostatique dans cette direction, de façon à créer un puits de potentiel dans la direction de  $\vec{B}$ , grâce à deux électrodes de piégeage polarisées à un potentiel identique, du signe de la charge de l'ion que l'on veut confiner.

La trajectoire des ions est alors assez complexe, elle se décompose en trois modes d'oscillation, dont le plus caractéristique du rapport  $m/z$  est le mouvement cyclotronique. Les ions ont un mouvement circulaire dans le plan perpendiculaire à  $\vec{B}$  et ils oscillent dans le puits de potentiel créé par les électrodes de piégeage dans la cellule. Le rayon de la trajectoire cyclotron dépend notamment du rapport  $q/m$ , de la vitesse des ions et du champ magnétique. Faisons une application numérique. Dans le cadre des expériences sur MICRA, le champ magnétique créé par l'aimant permanent est de 1.25 T dans la cellule, nous étudions des ions dont la masse varie de 18 à 300 unités de masse atomique (uma) et considérons une température de l'ordre de 298 K. Les rayons moyens varient alors de 0.09 à 0.37 mm dans une cellule d'environ 2 cm de côté. Ainsi tous les ions de même rapport  $q/m$  ont une trajectoire de même rayon, mais leur phase est distribuée aléatoirement. Pour pouvoir les détecter, il est alors nécessaire de rendre leurs mouvements cohérents. Ceci est possible grâce à une excitation résonnante, qui par ailleurs augmente le rayon de leur orbite.

### II.B.2.b. Excitation résonnante et détection

Pour exciter le mouvement cyclotron d'un type d'ions (de rapport  $m/z$  donné), on doit appliquer un champ électrique résonnant, c'est à dire un champ électrique radiofréquence de la forme  $E=E_0\sin(\omega t)$  tel que  $\omega=\omega_c$ . Lorsque la cellule contient des espèces ioniques  $i$  de masses et donc de fréquences cyclotrons différentes  $\omega_c^i$ , il faut exciter tous les ions, c'est à dire appliquer un champ d'excitation qui contient toutes les fréquences  $\omega_c^i$ . Dans le cas de l'expérience MICRA, le signal d'excitation est de type « SWIFT » (Stored Waveform Inverse Fourier Transform). Le signal correspondant à toutes les fréquences que l'on veut exciter est calculé par transformée de Fourier inverse.

Ainsi tous les ions excités voient leur énergie cinétique augmenter pour former des paquets d'ions cohérents associés chacun à une fréquence précise. Le rayon de l'orbite cyclotron augmente alors proportionnellement au temps d'application et à l'amplitude du champ radiofréquence  $E_0$ , suivant la relation  $r=E_0t/2B$  ( $r$  ne dépend plus de  $q/m$  et donc de la nature de l'ion). L'application de ce champ permet ainsi de placer les ions sur des trajectoires de grands rayons (Figure III-3-a)) et de créer un mouvement cohérent des ions de même  $m/z$  qui sont regroupés en paquet.

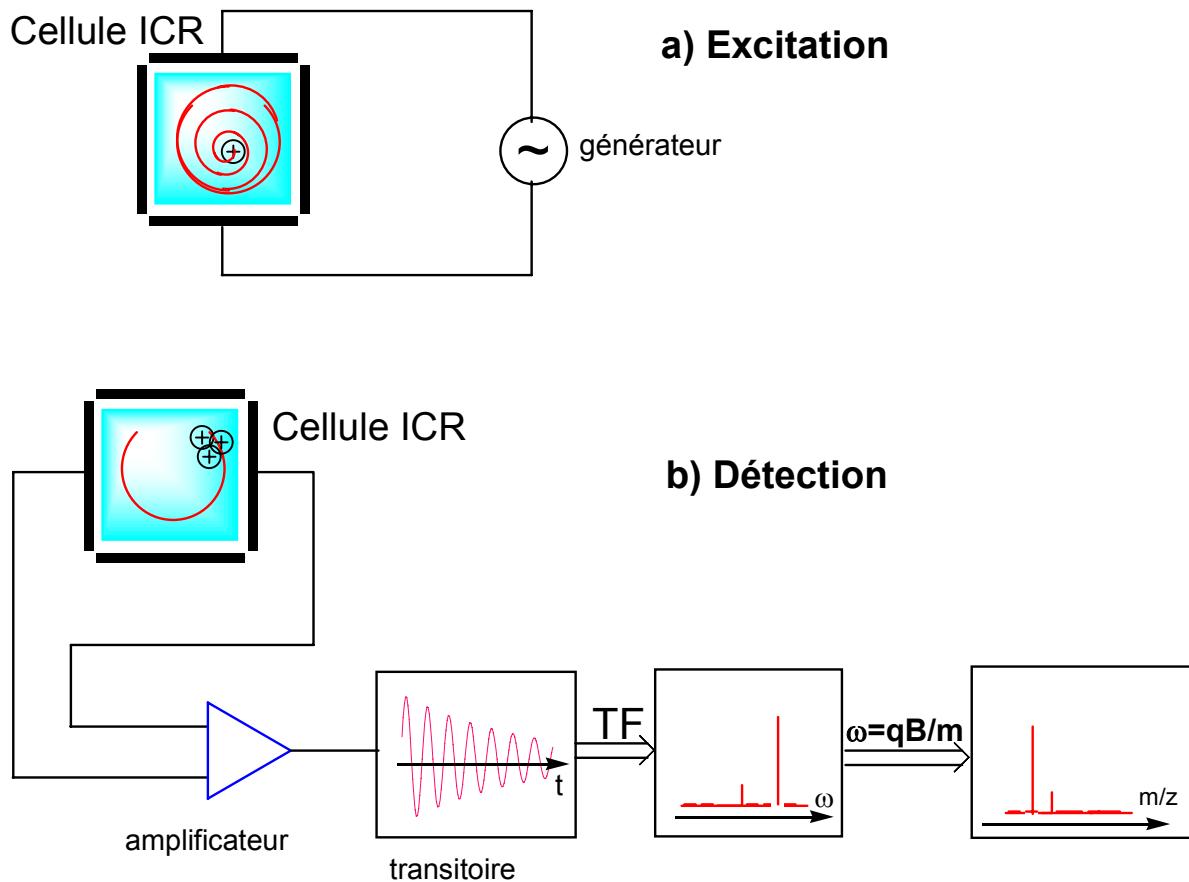
L'étape d'excitation est importante lorsque l'on a plusieurs espèces ioniques piégées simultanément. En effet, il peut être intéressant de sélectionner un type d'ions, pour étudier son comportement spécifique. Ceci est possible en éjectant tous les autres ions présents. Pour ce faire, on excite assez longtemps le mouvement cyclotron des ions à éjecter, leur orbite augmente alors jusqu'à rencontrer une électrode qui les neutralise.

La dernière étape est la détection du signal induit par les ions pour avoir accès aux  $\omega_c^i$ . Le mouvement cyclotron cohérent des paquets d'ions entre les deux électrodes de détection induit un courant alternatif image, appelé signal transitoire. Ce signal est de la forme  $F(t)=\sum_i N^i \cos(\omega_c^i t)$ , où  $N^i$  est proportionnel au nombre d'ions  $i$  dans la cellule.

Cependant ce courant est très faible. Il est donc nécessaire de l'amplifier pour pouvoir l'exploiter. Ensuite une transformée de Fourier du signal transitoire donne accès aux  $\omega_c^i$  (Figure III-3-b)). Puis la transformation  $\omega_c^i=q^iB/m^i$  donne accès à  $m^i/q^i$ . En pratique le transitoire n'est pas sinusoïdal, mais amorti à cause des collisions des ions avec le gaz résiduel présent dans la cellule. Ceci limite la détection des ions, en effet les collisions provoquent non seulement une perte d'ions qui peuvent être éjectés sur les électrodes, mais aussi un déphasage du mouvement des ions qui ne seront plus alors détectés. Il est donc essentiel de travailler à très basse pression, dans ce type d'instrument. Les pressions

utilisées sont le plus souvent dans la gamme  $10^{-7}$ - $10^{-8}$  torr pour les gaz de thermalisation et de réaction, avec une pression de base en général voisine de  $10^{-9}$  torr.

Lorsqu'un processus chimique ou physique a lieu, on peut faire varier le temps de détection pour en déterminer la cinétique.



**Figure III-3 :** Principe de l'excitation et de la détection dans un spectromètre de masse FT-ICR.

### II.B.3. Optimisation de la détection et résolution

Pour améliorer le rapport signal/bruit d'un spectre de masse, on peut accumuler les signaux transitoires en répétant  $N$  fois la séquence. En effet le bruit stochastique diminuant en  $\sqrt{N}$ , l'accumulation de  $N$  transitoires multiplie le rapport signal/bruit par  $\sqrt{N}$ . De plus cela permet de s'affranchir des parasites (dus par exemple à des émetteurs radiofréquence) car leurs phases varient aléatoirement d'une séquence à une autre.

Enfin pour améliorer la résolution en masse, il faut « écouter » le transitoire le plus longtemps possible. On peut alors augmenter le nombre de points et diminuer la fréquence de digitalisation du transitoire. Cependant le défaut de la spectrométrie de masse FTICR est que la résolution en masse décroît pour les grandes masses.

#### II.B.4. Cellules

De nombreuses géométries sont possibles pour les cellules ICR. Elles doivent cependant avoir au moins six électrodes dont deux de piégeage, deux d'excitation et deux de détection. La plus simple est une cellule cubique, fermée. Une autre géométrie couramment utilisée est un cylindre. Les différentes géométries sont optimisées pour améliorer la détection ou l'excitation ou le piégeage.

Après avoir vu le principe de la FTICR, nous allons décrire le montage MICRA utilisé dans le cadre de cette thèse.

### II.C. Originalités du montage MICRA en 2002:

Comme nous l'avons mentionné, MICRA est le prototype de spectromètre de masse FT-ICR compact et transportable développé au Laboratoire de Chimie Physique d'Orsay qui a fait l'objet du brevet AU2003214294. On peut souligner quelques originalités de MICRA, notamment plusieurs modes d'ionisation, les introductions d'espèces neutres disponibles, la géométrie de la cellule ICR qui permet le couplage avec un laser et enfin sa petite dimension et son poids limité pour un spectromètre de masse FT-ICR.

#### II.C.1. Montage expérimental

La figure III-4 présente une coupe du spectromètre de masse MICRA. Le champ magnétique de 1.25 T est produit par un aimant permanent structuré qui offre un gain considérable en encombrement et en masse (40 kg). Le poids global de MICRA est d'environ 200 kg. Rappelons que quasiment tous les spectromètres de masse FT-ICR commerciaux utilisent un aimant supraconducteur, dont le poids varie de une à deux tonnes. La cellule de MICRA de dimension  $2 \times 2 \times 2 \text{ cm}^3$  est située à l'intérieur de l'aimant. Les deux électrodes d'excitation habituellement placées à deux extrémités de la cellule ont été remplacées par un ensemble de quatre électrodes formant un tunnel. La forme du champ d'excitation au centre de la cellule reste proche d'un champ dipolaire. Cette géométrie lui confère une de ses originalités, puisque l'on a une cellule ouverte, qui permet l'accès optique à son centre. L'enceinte à vide possède deux fenêtres en sélénium de zinc transparentes dans l'infrarouge à partir de  $500 \text{ cm}^{-1}$ . Cette enceinte est pompée par une pompe turbomoléculaire, le pompage primaire étant assuré par une pompe à diaphragme (non représentées sur la figure III-4), ce qui permet d'atteindre un vide de  $10^{-9}$  torr (soit  $10^{-12}$  Bar).

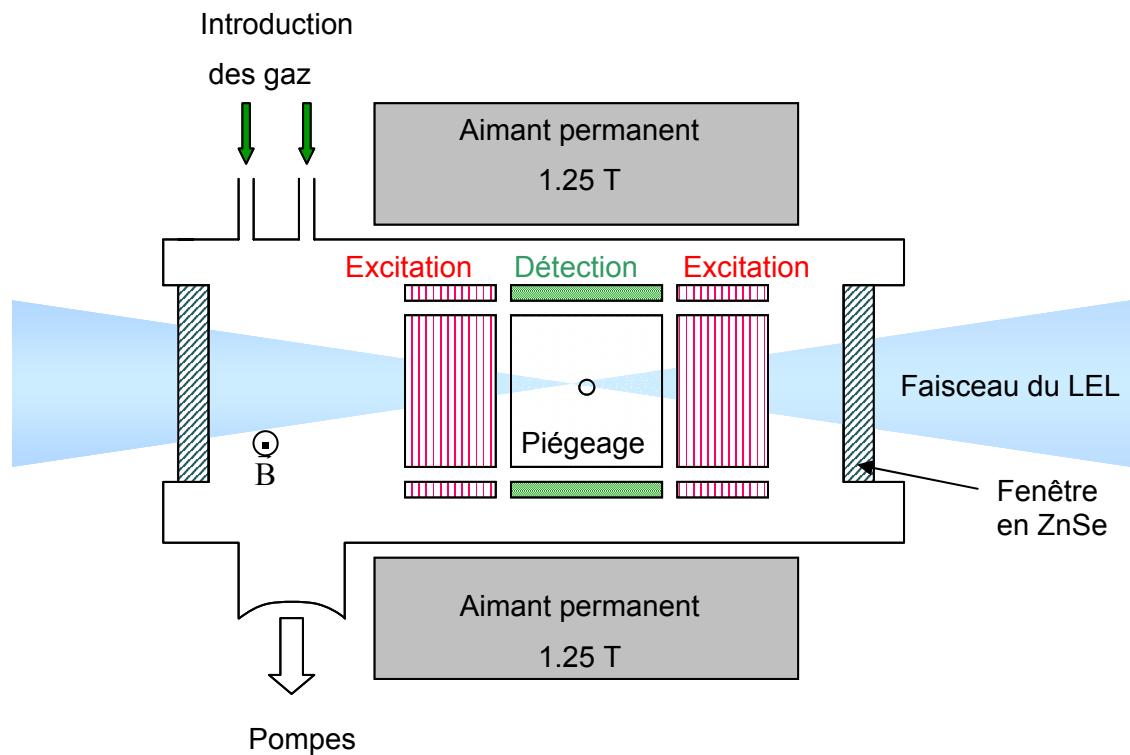


Figure III-4 : Coupe de MICRA.

## II.C.2. Formation des ions et introduction des espèces neutres

### II.C.2.a. Formation des ions

Pour former les cations métalliques (par exemple  $\text{Au}^+$ ,  $\text{Cu}^+$ ), on utilise la méthode d'ablation-ionisation laser. Le faisceau d'un laser Nd-YAG à la fréquence triplée (355 nm, puissance de 1 à 50 mJ par impulsion) est focalisé sur une pellicule métallique solide située au voisinage immédiat de la cellule de réaction. Par conversion thermique sur le métal, un plasma est formé contenant des ions métalliques. Grâce au champ magnétique, ces ions sont guidés à l'intérieur de la cellule. Ils sont suffisamment nombreux pour qu'une partie d'entre eux soit stabilisée dans le puits de potentiel de piégeage.

On peut aussi former des ions directement dans la cellule par impact électronique, par chauffage d'un filament de tungstène.

Cependant, MICRA ne possède pas de système pour refroidir les ions ; on peut seulement les thermaliser par collision avec un gaz neutre ou par équilibrage radiatif avec leur environnement. Les ions produits sont donc chauds, ce qui aura une influence importante sur la largeur des raies d'absorption infrarouge.

### II.C.2.b. Introduction d'espèces neutres

Pour introduire des espèces neutres sous forme gazeuse dans la cellule, on dispose de plusieurs entrées de gaz avec des vannes de fuite ou une vanne pulsée.

## II.D. Développement apporté lors de ces travaux

### II.D.1. Introduction

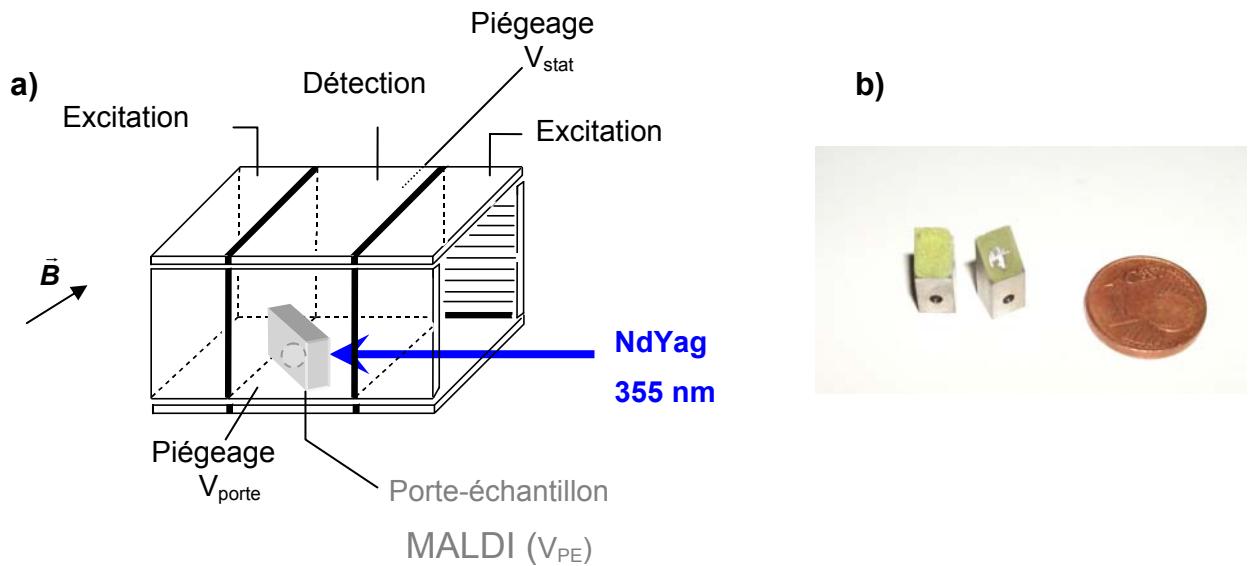
Comme nous venons de le voir, il existait initialement deux modes d'ionisation dans cette expérience : par impact électronique et par ablation-ionisation laser. Or nous souhaitions produire des complexes de molécules biologiques en interaction avec  $\text{Na}^+$ , en phase gazeuse, pour lesquels ces deux méthodes ne sont pas adaptées. Habituellement deux techniques peuvent être utilisées pour former ce type d'ions : l'ionisation electrospray (ESI) et la désorption-ionisation assistée par matrice (MALDI). Cependant il n'est pas possible d'injecter des ions depuis une source externe dans la cellule de MICRA en raison de la configuration du champ magnétique. Un dispositif basé sur l'utilisation d'un aimant permanent cylindrique produisant un champ axial est en cours de réalisation au Laboratoire de Chimie Physique (Orsay). Toutefois le couplage avec une source Electrospray n'est pas un développement simple, notamment en raison du nombre d'étages différentiels pour passer de la pression atmosphérique au vide nécessaire dans la cellule FTICR ( $p \approx 10^{-9}$  torr). Nous avons donc opté pour la production d'ions par MALDI en adaptant le système d'ablation-ionisation laser.

### II.D.2. Aménagement de MICRA : le MALDI

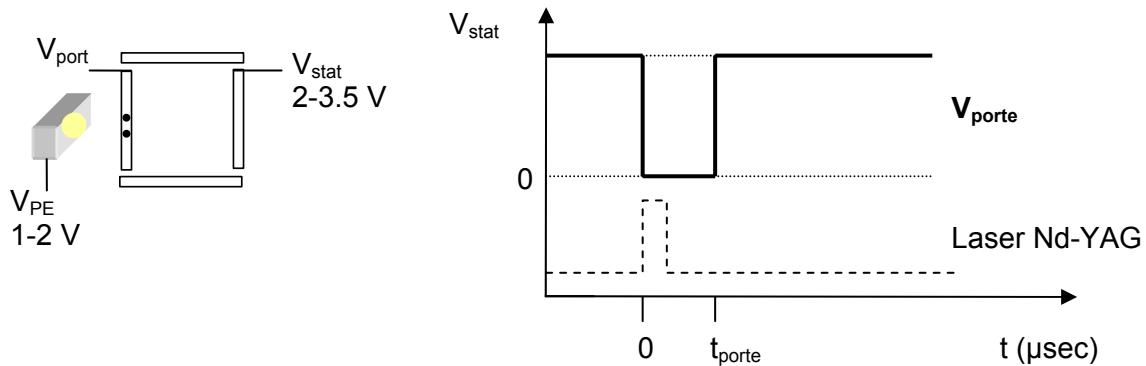
Afin de produire les ions par MALDI, un porte-échantillon métallique est placé au voisinage de la cellule (figure III-5). L'irradiation qui permet de désorber et d'ioniser le mélange matrice/échantillon co-cristallisé est assurée par un laser Nd-YAG en mode triplé (355 nm), dont le faisceau peut être focalisé sur l'échantillon. Les ions sont ensuite guidés à l'intérieur de la cellule grâce au champ magnétique, en passant par un trou de 5 mm percé dans une des plaques de piégeage. Cette technique peut être qualifiée de « MALDI interne », les ions étant produits dans le champ magnétique. Elle permet de former des ions biologiques en phase gazeuse. Cependant, un seul échantillon peut être introduit, de plus le changement d'échantillon nécessite d'interrompre le pompage, ce qui impose un temps d'attente d'environ 4h à 6h entre l'étude de deux échantillons, pour atteindre à nouveau un vide correct.

Afin d'améliorer et de régulariser l'admission des ions dans la cellule, nous avons mis en place une porte de potentiel ( $V_{\text{porte}}$ ) sur la plaque de piégeage adjacente au

porte-échantillon, synchronisée par rapport à l'impact laser, la seconde plaque de piégeage étant à un potentiel statique ( $V_{stat}$ ). Le porte-échantillon est maintenu à un potentiel positif ( $V_{PE}$ ), constant et inférieur au potentiel statique. La structure temporelle de la porte est donnée par la figure III-6. Le potentiel de la porte est nul au moment de la production des ions, puis il est porté à  $V_{stat}$  au temps  $t_{porte}$  (environ 20 à 40  $\mu$ s), ce qui permet d'attirer les ions à l'intérieur de la cellule puis de les piéger.  $V_{stat}$ ,  $V_{PE}$  et  $t_{porte}$  sont des paramètres à optimiser selon l'ion que l'on cherche à produire (figure III-6).



**Figure III-5 :** a) Cellule ICR de MICRA adaptée à la formation d'ions par MALDI ; b) photo du porte échantillon avec l'échantillon avant et après utilisation, à côté d'une pièce d'un centime d'euro.



**Figure III-6 :** Potentiel d'extraction.

### II.D.3. Les échantillons MALDI

L'inconvénient de cette technique de MALDI "interne" est qu'à chaque changement d'échantillon, il faut mettre la cellule à pression atmosphérique. Ceci interdit de tester systématiquement la formation des ions dans MICRA. Pour cette raison, il nous a semblé nécessaire d'effectuer des tests préliminaires sur un autre appareil, pour optimiser les

conditions (matrice, concentration et solvant) pour former les complexes M-Na<sup>+</sup>. Des tests ont été effectués avec Sophie Bourcier en collaboration avec Martine Caroff à l’Institut de Biochimie et de Biophysique Moléculaire et Cellulaire d’Orsay, sur un spectromètre de masse à temps de vol équipé d’une source MALDI utilisant un laser à azote. Nous avons effectué des tests pour la formation de complexes acides aminés-Na<sup>+</sup>.

#### II.D.3.a. Choix préliminaire de 4 matrices

Le mécanisme de formation des ions par MALDI n'est pas encore totalement élucidé, il reste assez difficile de prédire quelle matrice sera la plus adaptée à la formation d'une espèce précise. Cependant on peut faire un raisonnement simple de thermochimie. En effet pour maximiser la formation de AA-Na<sup>+</sup>, il faut minimiser la formation des ions AA-H<sup>+</sup> et Matrice-Na<sup>+</sup>. Il faut donc que PA(AA) < PA(Matrice) et EI(Matrice-Na<sup>+</sup>) < EI(AA-Na<sup>+</sup>).

Afin de choisir parmi le grand nombre de matrices utilisables, nous avons référencé les affinités protoniques et pour le sodium des quatre matrices DHB, 4-HCCA, PA et NA avec lesquelles nous avons effectué les essais préliminaires (tableau III-2).

Composé		ΔH <sub>complexation(Na<sup>+</sup>-R)</sub> à 298 K kJ/mol	PA kJ/mol
Pro		184 <sup>6</sup>	920 <sup>7</sup>
Gly		155 <sup>6</sup>	887 <sup>7</sup>
Gly <sub>3</sub>		237	924-950 <sup>8</sup>
Acide 2,5 dihydroxy- benzoïque	DHB	144 <sup>9</sup>	848-852 <sup>10</sup>
Acide $\alpha$ -cyano-4- hydroxy cinnamique	4HCCA	185 <sup>9</sup>	865 <sup>10</sup>
Acide picolinique	PA	179 <sup>9</sup>	918 <sup>10</sup>
Acide nicotinique	NA	157 <sup>9</sup>	919 <sup>10</sup>

Tableau III-2 : Caractéristiques thermodynamiques de matrices usuelles.

#### II.D.3.b. Choix d'échantillons : test sur MALDI-TOF

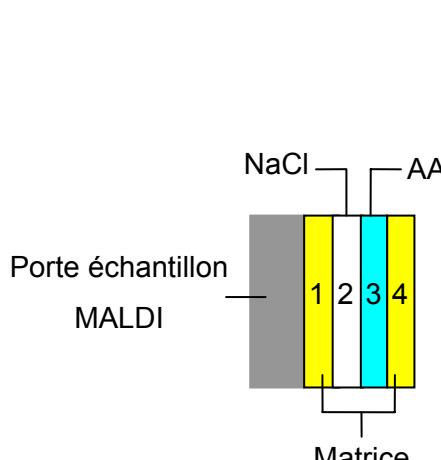
Nous avons testé les quatre matrices HCCA, DHB, NA et PA, 4 solvants eau, méthanol, eau/méthanol (1/1) et eau/acétone (1/3), et des compositions variant de 1 analyte pour 1000 matrices à 1 analyte pour 1 matrice, en faisant varier la composition de l'analyte de 1/1, 1/0.5 et 1/0 (AA/Na<sup>+</sup>). Nous avons ainsi déterminé que l'acide HCCA dans le solvant eau/acétone (1/3), donnait le plus d'ions ProNa<sup>+</sup>, et que la DHB dans de l'eau donnait le meilleur signal d'ions GlyNa<sup>+</sup> toutes deux dans des proportions 9/1/0.5 (Matrice/AA/Na<sup>+</sup>).

#### II.D.3.c. Echantillons finaux

Ces tests préliminaires nous ont permis de déterminer un ordre de grandeur de composition d'échantillons. Cependant les premiers essais sur MICRA ont été peu

concluants, le seul ion visible étant  $\text{Na}^+$ . Il a donc fallut surtout changer le mode de dépôt et réoptimiser la composition.

Le mode de dépôt le plus simple, que nous avions utilisé, est de mélanger directement le sel, l'AA et la matrice dans les proportions souhaitées, de déposer 5  $\mu\text{L}$  de cette solution sur le porte échantillon, puis de laisser sécher à l'air ou en accélérant le séchage (par exposition sous une ampoule électrique alimentée ou un sèche-cheveux). Cette technique ayant été inefficace, nous avons utilisé par la suite le mode de dépôt de type sandwich, c'est-à-dire par couches successives,<sup>11</sup> en déposant tout d'abord une couche de matrice, puis une couche de  $\text{NaCl}$ , puis une couche d'acide aminé et enfin une nouvelle couche de matrice. Nous avons obtenu un meilleur signal des complexes acides aminés- $\text{Na}^+$  avec les proportions atomiques suivantes 9/2/1 (Matrice/AA/ $\text{Na}^+$ ) avec des concentrations d'analytes de l'ordre de  $10^{-2}$  mol/L. Une description détaillée de l'échantillon et de chaque couche est donnée dans le tableau III-3.



		1	2	3	4
		Matrice	$\text{NaCl}$	AA	matrice
Porte échantillon MALDI	<b>Pro-<math>\text{Na}^+</math></b> solvant $c$ (mol/L) $V$ ( $\mu\text{L}$ ) équivalent molaire	Eau/Acétone (1/3)	eau	eau	Eau/Acétone (1/3)
		$3.75 \cdot 10^{-2}$	$2.5 \cdot 10^{-2}$	$2.5 \cdot 10^{-2}$	$3.75 \cdot 10^{-2}$
		15	5	10	15
<b>Gly-<math>\text{Na}^+</math></b> solvant $c$ (mol/L) $V$ ( $\mu\text{L}$ ) équivalent molaire	4.5	1	2	4.5	
	eau	eau	eau	eau	
	$3.75 \cdot 10^{-2}$	$2.5 \cdot 10^{-2}$	$2.5 \cdot 10^{-2}$	$3.75 \cdot 10^{-2}$	
	solvant $c$ (mol/L) $V$ ( $\mu\text{L}$ ) équivalent molaire	15	5	10	15
		4.5	1	2	4.5

**Tableau III-3:** Composition des échantillons MALDI de type sandwich pour former des complexes acides aminés- $\text{Na}^+$ .

Ces conditions peu habituelles sont dues notamment à l'utilisation d'un laser et d'une configuration d'expérience différents. Les principales caractéristiques du laser Nd :YAG de MICRA et du laser à azote utilisé dans les expériences de MALDI classique sont présentées dans le tableau III-4. On peut voir que l'énergie apportée par chaque impulsion laser est beaucoup plus importante avec le laser Nd :YAG qu'avec le laser à azote ; cependant la

fréquence entre deux pulses est beaucoup plus grande avec le laser à azote. On peut donc penser que l'étape d'absorption d'énergie électronique puis le transfert de cette énergie en énergie vibrationnelle seront différents selon le laser utilisé.

caractéristiques	Laser Nd :YAG de MICRA	Laser à azote classique
Longueur d'onde	355 nm	337 nm
Fréquence	1 Hz	10-20 Hz
Durée de l'impulsion	5 ns	5 ns
Énergie ou puissance	1-10 mJ	100-300 µJ

**Tableau III-4 :** Différentes caractéristiques du laser Nd :YAG et du laser à azote utilisés respectivement dans MICRA et dans des expériences MALDI classiques.

### III. Méthodes théoriques

#### III.A. Introduction

##### III.A.1. Introduction

Il existe de nombreuses méthodes de chimie théorique ayant pour but de déterminer des propriétés physiques ou chimiques de molécules isolées, que ce soient des propriétés thermodynamiques comme des enthalpies de liaison, des énergies relatives de différents conformères, ou des simulations de spectres infrarouge, Raman ou électronique. On peut différencier deux grandes classes de méthodes de simulation : d'une part les méthodes de chimie quantique qui permettent de déterminer avec précision les propriétés électroniques des molécules, d'autre part les méthodes de mécanique moléculaire qui sont basées sur des paramètres empiriques qui permettent notamment de faire des études conformationnelles, y compris de molécules de grande taille. Chacune de ces méthodes présente, bien sur, des avantages et des inconvénients qui sont rassemblés dans le tableau III-5.

	Chimie quantique	Mécanique moléculaire
Fondement	Mécanique quantique Equation de Schrödinger	Mécanique classique Champ de force Méthode empirique
Coût CPU N= nombre d'atome N'=nombre de fonctions de base	$N^4$ pour HF	N
Précision énergie relative	Bonne - très bonne selon la méthode	Moyenne – bonne dans les limites de la paramétrisation
Grandeur accessible	Propriétés électroniques : Energie électronique Energies relatives de conformères Enthalpie de formation Spectre vibrationnel Polarisabilité Réactivité...	Energie stérique Energies relatives de conformères Enthalpie de formation Fréquences de vibration
Système	Relativement petit < 200 atomes	Grand

**Tableau III-5 :** Principales caractéristiques des méthodes de chimie quantique et de mécanique moléculaire.

### III.A.2. Stratégies selon la taille et la nature des systèmes

Notre sujet étant la structuration d'acides aminés et de peptides par  $\text{Na}^+$ , nous avons été amenés à effectuer des études conformationnelles de dipeptides à des décapeptides sodiés. Nous avons vu, dans le chapitre I, que le nombre de conformations devient de plus en plus important avec le nombre de résidu. Nous avons donc effectué une recherche systématique des isomères ou conformères de plus basse énergie des peptides sodiés. Pour ce faire, la première étape est une ou plusieurs recherches conformationnelles de type Monte-Carlo, basées sur des calculs de champ de forces de type AMBER. Celles-ci ont permis l'exploration de la surface d'énergie potentielle et l'identification de structures accessibles. En particulier, la méthode AMBER est couramment utilisée pour les études de protéines car elle décrit bien, à la fois, les liaisons peptidiques et les liaisons hydrogène intramoléculaires, cependant les interactions métal-peptide sont moins bien décrites. Ensuite les structures engendrées ont été réoptimisées par des calculs *ab initio* ou de type DFT, qui elles, permettent de mieux décrire l'interaction du peptide et du métal. Dans le cas des grands peptides, nous avons été amenés à utiliser le logiciel TURBOMOLE, qui permet notamment de faire des calculs avec la fonctionnelle BLYP d'optimisation de géométrie et de fréquences de vibration, dans l'approximation de la « Résolution de l'Identité » ou ri-BLYP, qui offre un gain notable en temps de calcul. Dans ce qui suit nous allons développer d'une part la méthode AMBER et ses performances, puis le principe et les performances de l'approximation ri.

## III.B. Le champ de force AMBER

### III.B.1. Les champs de forces

Les méthodes de mécanique moléculaire s'appuient sur une formulation de mécanique classique. Les atomes constituant la molécule sont considérés comme des points matériels maintenus les uns à proximité des autres par différents types de force. Ces forces d'interaction sont représentées par des fonctions d'énergie potentielle, l'ensemble de ces fonctions est appelé le champ de force. La formule générale des champs de force est la suivante :

$$E(MM) = E_l + E_\theta + E_\phi + E_{vdw} + E_{elec} + \dots$$

où  $E_l$  est l'énergie d'élongation des liaisons,  $E_\theta$  l'énergie de déformation des angles de valence,  $E_\phi$  l'énergie de torsion autour des liaisons,  $E_{vdw}$  l'énergie des interactions de van der Waals et  $E_{elec}$  l'énergie d'interaction électrostatique. Cette énergie  $E(MM)$ , appelée énergie

stérique, n'a aucun sens physique ; seule la différence d'énergie entre deux conformères peut être interprétée.

Il existe différentes expressions des champs de forces, cependant deux sont spécialisés dans les macromolécules biologiques : CharmM et AMBER. Nous avons choisi d'utiliser le champ de force AMBER.

### III.B.2. AMBER

AMBER (*Assisted Model Building with Energy Refinement*) a été développé par P. A. Kollman *et al.* pour notamment étudier les protéines.<sup>12</sup> Son champ de force a la forme analytique suivante :

$$E(\text{AMBER}) = \sum_l k_l (l - l_0)^2 + \sum_\theta k_\theta (\theta - \theta_0)^2 + \sum_n \frac{V_n}{2} (1 + \cos(n\phi - \gamma)) + \sum_{i < j} \frac{A_{ij}}{d_{ij}^{12}} - \frac{B_{ij}}{d_{ij}^6} + \frac{q_i q_j}{\epsilon_r d_{ij}}$$

Les énergies d'elongation de liaison et de pliage sont représentées par une simple expression harmonique. L'énergie de van der Waals est représentée par un potentiel 6-12 et l'énergie électrostatique est modélisée par des interactions coulombiennes de charges centrées sur les atomes. Pour calculer  $E(\text{AMBER})$ , il est alors nécessaire de connaître les paramètres  $k_l$ ,  $k_\theta$ ,  $V_n$ ,  $A_{ij}$ ,  $B_{ij}$  et  $q_i$  (les  $q_i$  sont appelés les charges RESP : *Restrained ElectroStatic Potential*). Pour effectuer ces calculs nous avons choisi d'utiliser le logiciel Hyperchem<sup>13</sup> dans lequel est implémentée la version Amber94. Les différents paramètres pour chaque résidu et pour le sodium (charge +1) ont été déterminés par l'équipe de P. A. Kollman<sup>14</sup>. Reste alors à déterminer les paramètres pour le groupement hydroxyle de l'acide terminal et l'amine terminale. Pour le groupement hydroxyle, nous avons choisi les charges -0.661 sur l'oxygène et +0.429 pour H qui correspondent à une moyenne des charges du groupement hydroxyle de la serine et de la thréonine. Enfin pour le groupement amine terminal nous avons utilisé les charges -0.378 pour l'azote et +0.233 pour les hydrogènes. Il faut par ailleurs définir toutes les liaisons du système, notamment les liaisons  $\text{Na}^+$ -ligand, ce qui constitue une contrainte pour les études conformationnelles sur les grands peptides où il existe de nombreux sites de coordination possibles (les oxygènes et l'amine terminale). Nous avons donc été amenés à effectuer au moins deux études conformationnelles pour chaque complexe en contraignant le métal à être coordonné sur différents sites du peptide. Nous avons choisi, d'une part de lier  $\text{Na}^+$  aux trois oxygènes du côté N-terminal, puis dans une recherche indépendante aux trois oxygènes côté C-terminal. Ces choix n'entraînent pas de contraintes drastiques sur les sites de coordination de  $\text{Na}^+$ , lors des calcul DFT qui suivent ces explorations conformationnelles préliminaires.

### III.C. Méthodes de chimie quantique

#### III.C.1. Introduction

Comme dans le cas des méthodes de mécanique moléculaire, il existe une grande diversité de méthodes de chimie quantique. Le but de chacune de ces méthodes est de résoudre l'équation de Schrödinger indépendante du temps :  $\mathbf{H}\Psi_j^{\text{tot}}=E_j^{\text{tot}}\Psi_j^{\text{tot}}$  où  $\mathbf{H}$  est l'opérateur Hamiltonien d'un système composé de  $n$  électrons (de masse  $m_i$  et de position  $r_i$ ) et de  $N$  noyaux (de masse  $M_k$ , de position  $R_k$  et de charge  $Z_k$ ). Les états propres stationnaires du système sont caractérisés par les couples  $(E_j^{\text{tot}}, \Psi_j^{\text{tot}})$ . Cependant, la résolution analytique de cette équation est impossible en général et l'on procède donc à différentes approximations. Dans toutes les méthodes, on procède d'une part à l'approximation de Born-Oppenheimer qui considère les noyaux fixes. En effet les noyaux étant beaucoup plus lents que les électrons, le principe est de séparer les variables et de rechercher  $\Psi_j^{\text{tot}}$  comme le produit d'une fonction nucléaire  $\Xi(R_K)$  et d'une fonction électronique  $\Psi_j^{\text{éI}}(R_K, r_i)$ . La fonction électronique est obtenue en résolvant une équation dite électronique, où la position des noyaux est fixe. D'autre part, pour déterminer la fonction d'onde, il est nécessaire de trouver les OM  $\Phi_i$  (orbitales moléculaires) sur lesquelles sont bâties les  $\Psi_j^{\text{éI}}(R_K, r_i)$ . On procède alors à l'approximation CLOA (*Combinaison Linéaire d'Orbitales Atomiques*) où chaque OM est une combinaison linéaire de fonctions atomiques ( $\chi_p$ ). Pour des raisons pratiques, leur parties radiales sont définies comme des combinaisons de gaussiennes.

On doit alors résoudre l'équation :  $\mathbf{H}^{\text{éI}}\Psi_j^{\text{éI}}=E_j^{\text{éI}}\Psi_j^{\text{éI}}$  où  $\mathbf{H}^{\text{éI}}=\underbrace{\sum_i \left( -\frac{\Delta_i}{2} - \sum_{\mu} \frac{Z_{\mu}}{R_{i,\mu}} \right)}_{\text{Terme mono-électronique}} + \underbrace{\sum_i \sum_{j < i} \frac{1}{r_{i,j}}}_{\text{Terme bi-électronique}}$

$$= \sum_i h_i$$

La partie mono-électronique ne pose pas de grandes difficultés pour être calculée, contrairement au terme bi-électronique. On est alors obligé de faire de nouvelles approximations. Dans la suite, nous n'avons utilisé que des méthodes basées sur une description mono-déterminantale de la fonction d'onde électronique.

### III.C.2. Méthodes Hartree-Fock (HF) et post-Hartree-Fock

#### III.C.2.a. HF

*Principe :*

Dans le cadre de la méthode Hartree-Fock, la fonction d'onde du système est considérée comme le produit antisymétrisé et orthonormé des fonctions d'onde mono-électroniques  $\Phi_i$ , donc comme le déterminant de Slater de ces  $\Phi_i$ . Les  $\Phi_i$  sont les fonctions propres de

l'hamiltonien approché de Fock :  $F(i) = h(i) + \sum_j^{\text{occupées}} (ii|jj) - (ij|ij) = h(i) + \sum_j^{\text{occupées}} J_{ij} - K_{ij}$  où les

intégrales bi-électroniques  $J_{ij}$  et  $K_{ij}$  sont nommées intégrales coulombiennes et intégrales d'échange, respectivement.

$$(ii|jj) = J_{ij} = \iint \phi_i^*(1) \phi_i(1) \frac{1}{r_{1,2}} \phi_j^*(2) \phi_j(2) d\tau_1 d\tau_2 \quad (ij|ij) = K_{ij} = \iint \phi_i^*(1) \phi_j(1) \frac{1}{r_{1,2}} \phi_i^*(2) \phi_j(2) d\tau_1 d\tau_2$$

Dans l'approximation CLOA :  $\phi_i = \sum_{p=1}^N c_{ip} \chi_p$ , la décomposition de  $F(i)$  sur les  $\chi_p$  conduit aux

$$\text{équations de Roothaan : } \sum_p c_{ip} (F_{pq} - \varepsilon_i S_{pq}) = 0 \text{ où } S_{pq} = \langle \chi_p | \chi_q \rangle \text{ et } F_{pq} = \langle \chi_p | F | \chi_q \rangle.$$

Les inconnues du système sont les  $c_{ip}$ . Cependant pour résoudre les équations de Roothaan, il est nécessaire de connaître ces  $c_{ip}$  qui interviennent dans le terme  $F_{pq}$ . La solution est de résoudre ce problème de manière itérative, à partir d'un jeu d'orbitales d'essai ( $\Phi_i^1, \varepsilon_i^1$ ). La résolution des équations de Roothaan conduit à un nouveau jeu d'orbitales ( $\Phi_i^2, \varepsilon_i^2$ ). Le processus itératif est stoppé lorsque l'énergie ou les  $c_{pq}^n$  sont stables, selon un critère de convergence fixé.

*Conséquences :*

D'une part, la résolution des équations de Roothaan fait apparaître des termes bi-électroniques à quatre centres :  $\langle pq | rs \rangle$ , dont le nombre croît comme la puissance 4 du nombre d'orbitales atomiques. Cela pose donc des problèmes techniques de résolution de ces équations.

D'autre part, l'expression mono-déterminantale de la fonction d'onde implique que l'on considère les électrons comme des particules indépendantes, c'est à dire que le mouvement d'un électron n'a qu'un effet moyen sur le mouvement des autres électrons. On néglige donc la corrélation électronique. Il existe plusieurs solutions pour pallier cette approximation, dont la méthode de perturbation de Möller-Plesset à l'ordre 2 (MP2).

### III.C.2.b. Une méthode Post-HF : MP2

Dans la méthode de perturbation Möller-Plesset d'ordre 2 (MP2), la corrélation électronique est traitée en introduisant des configurations électroniques mono et di-excitées en plus de la configuration électronique fondamentale (Hartree-Fock). Pour minimiser la lourdeur des calculs, on traite le problème par perturbation au second ordre, l'hamiltonien non perturbé étant l'hamiltonien HF.

### III.C.2.c. Détermination des énergies de liaison L-Na<sup>+</sup> pour de petits systèmes

Des calculs précédents effectués au laboratoire<sup>6, 15</sup>, ont montré que l'énergie de liaison de complexes acide aminé-Na<sup>+</sup>, pouvait être déterminée par un calcul d'énergie MP2 avec la base 6-311+G(2d,2p) sur une structure optimisée au niveau HF/6-31G\*. La comparaison avec des résultats de calcul à des niveaux très élevés et avec des valeurs expérimentales a montré que ce niveau permet de s'affranchir du calcul d'erreur de superposition de base (BSSE), grâce à une compensation de petites erreurs. Cette méthode permet d'obtenir des résultats fiables et précis, si on considère des molécules de relativement petites tailles. Cependant, pour étudier des molécules biologiques plus grosses, comme des peptides, des méthodes quantiques moins exigeantes en moyens informatiques doivent être utilisées, notamment si on s'intéresse aux calculs de fréquences de vibration. Les méthodes basées sur la théorie de la fonctionnelle de la densité ont atteint un niveau de fiabilité compétitif avec des méthodes *ab initio* incluant la corrélation électronique, tout en permettant de réduire notablement la charge de calculs.

## III.C.3. Méthodes de la DFT

### Principe

La théorie de la fonctionnelle de la densité (DFT) a été développée par P. Hohenberg, W. Kohn et L. J. Sham.<sup>16, 17</sup> Ils ont démontré que toutes les propriétés d'un système dans un état fondamental non dégénéré sont complètement déterminées par sa densité électronique  $\rho(\vec{r})$ . L'énergie totale du système apparaît alors comme une fonctionnelle de cette densité électronique. De plus ils ont étendu le principe variationnel à la densité :  $\frac{\partial E(\rho(\vec{r}))}{\partial \rho(\vec{r})} = 0$ , ce qui permet de trouver la densité électronique, puisque la fonction

$\rho(\vec{r})$  exacte correspond au minimum d'énergie. Le problème cependant est que la fonctionnelle exacte de l'énergie est inconnue. On est alors amené à faire des approximations.

L'énergie peut être décomposée en quatre termes : le terme d'énergie cinétique des électrons, un terme d'énergie potentielle d'attraction noyau-électron, un terme de répulsion

coulombienne entre électrons et un terme d'énergie de corrélation et d'échange. La fonctionnelle d'énergie cinétique a été résolue par Kohn et Sham par l'introduction de fonctions mono-électroniques. Le problème se ramène alors à la résolution d'une équation mono-électronique très analogue à celle de Hartree-Fock :

$$E = \sum_{i=1}^N \left\langle \phi_i \left| -\frac{1}{2} \Delta \right| \phi_i \right\rangle + \int \rho(\vec{r}) V_{Ne}(\vec{r}) d\vec{r} + \frac{1}{2} \iint \frac{\rho(\vec{r}_1)\rho(\vec{r}_2)}{r_{1,2}} d\vec{r}_1 d\vec{r}_2 + E_{xc}[\rho]$$

Le terme de répulsion électron-électron développé sur la base des orbitales de Kohn et Sham fait apparaître l'opérateur  $J$  précédemment décrit dans le principe de la méthode HF :

$$J = \iint \frac{\rho(\vec{r}_1)\rho(\vec{r}_2)}{r_{1,2}} d\vec{r}_1 d\vec{r}_2 = \sum_i \sum_j (ii|jj) = \sum_i \sum_j J_{ij} = \sum_i \sum_j \iint \phi_i^*(1)\phi_i(1) \frac{1}{r_{1,2}} \phi_j^*(2)\phi_j(2) d\tau_1 d\tau_2$$

#### *Les différentes générations de fonctionnelles*

Il existe trois générations de fonctionnelles qui présentent une sophistication croissante du terme d'échange et de corrélation. La première correspond aux fonctionnelles de type LDA (*Approximation de la Densité Locale*) où la partie d'échange est basée sur le calcul du gaz uniforme d'électrons. La forme du potentiel d'échange-corrélation  $\left( \frac{\partial E_{xc}[\rho]}{\partial \rho(\vec{r})} \right)$

est alors un opérateur multiplicatif, il est donc plus simple que l'opérateur d'échange  $K$  qui est intégral. La seconde génération sont les fonctionnelles GGA (*Approximation du Gradient Généralisé*) qui dépendent de la densité et de son gradient (par exemple : BLYP, BP86...). Enfin la dernière génération, que constituent les fonctionnelles hybrides, prend en compte explicitement le terme d'échange Hartree-Fock, calculé au moyen des orbitales de Kohn et Sham ; la fonctionnelle la plus utilisée est B3LYP. Les fonctionnelles hybrides donnent d'excellents résultats sur les structures électroniques et les propriétés associées. Ainsi nous avons montré dans le chapitre II partie II-D, que les spectres infrarouges simulés par la méthode B3LYP/6-31G\* sont de bonne qualité. Nous avons donc systématiquement utilisé ce niveau de calcul pour la construction des spectres infrarouges théoriques.

Même si les fonctionnelles hybrides ont su prouver qu'elles étaient très efficaces pour les déterminations structurales, les fonctionnelles de type GGA sont aussi de bonnes alternatives, surtout pour de gros systèmes, puisqu'elles permettent un gain non négligeable en temps de calcul.

### III.D. La méthode ri-BLYP : principe et calibration

Pour les études conformationnelles de gros peptides sodiés (du penta-peptide au déca-peptide), nous avons choisi d'utiliser la fonctionnelle BLYP, couplée à la méthode

d'approximation de « résolution de l'identité ». La méthode ri-BLYP est, comme nous allons le voir très performante : d'une part en temps de calcul, ce qui est important car l'espace conformationnel accessible étant très grand, nous avons été contraints d'étudier un grand nombre de conformations pour chaque peptide ; d'autre part en précision, puisque la calibration de cette méthode montre un très bon accord entre les énergies B3LYP et ri-BLYP avec une même qualité de base.

### III.D.1. Principe de l'approximation ri

#### Principe

Dans le cas des fonctionnelles GGA, le terme coulombien est le plus difficile à calculer, car comme dans le cas de la méthode HF, il fait intervenir des termes bi-électroniques à quatre centres. L'approximation ri « résolution de l'identité » permet de réduire notablement le temps de calcul en se limitant à des intégrales à trois centres au maximum. Cette technique a été développée par R. Ahlrichs *et al.*<sup>18-21</sup> et implémentée dans le logiciel TURBOMOLE<sup>22</sup>, que nous avons utilisé. Cette approximation consiste à développer la densité électronique  $\rho$  sur un jeu de bases auxiliaires  $\alpha$  étendues et centrées sur les atomes. L'intérêt de cette méthode est que la dimension de la base auxiliaire  $N^{\text{aux}}$  est inférieure à  $N^2$ . On a alors

$\rho(\vec{r}) = \sum_{k,l}^{N,N} D_{kl} k(\vec{r}) l(\vec{r}) \approx \tilde{\rho}(\vec{r}) = \sum_{\alpha}^{N^{\text{aux}}} c_{\alpha} \alpha(\vec{r})$ , où  $D_{kl}$  sont les éléments de la matrice densité. Les coefficients  $c_{\alpha}$  sont déterminés de manière à minimiser le terme  $(\rho - \tilde{\rho}|\rho - \tilde{\rho}) = (\rho - \tilde{\rho}|J|\rho - \tilde{\rho})$ . On trouve alors :

$$\frac{\partial(\rho - \tilde{\rho}|\rho - \tilde{\rho})}{\partial c_{\alpha}} = 0 \Leftrightarrow \sum_{\beta}^{N^{\text{aux}}} c_{\beta} (\alpha|\beta) = \sum_{k,l}^{N,N} D_{kl} (\alpha|kl)$$

La résolution de ce système d'équations linéaires, qui n'impliquent que des intégrales bi-électroniques à deux et trois centres, donne accès aux  $c_{\alpha}$ .

Une fois les coefficients  $c_{\alpha}$  déterminés, on peut calculer l'énergie de coulomb  $J$  et ses éléments de matrice  $J_{kl}$  par les formules :  $J = \frac{1}{2}(\rho|\rho) \approx \frac{1}{2}(\rho|\tilde{\rho}) = \frac{1}{2}(\tilde{\rho}|\tilde{\rho}) = \sum_{\alpha, \beta}^{N^{\text{aux}}, N^{\text{aux}}} c_{\alpha} c_{\beta} (\alpha|\beta)$  et

$$J_{kl} = \sum_{\alpha}^{N^{\text{aux}}} c_{\alpha} (\alpha|kl).$$

Dans l'approximation ri, il faut donc déterminer les  $N^{\text{aux}}$  coefficients  $c_{\alpha}$  par la méthode ci-dessus et les  $N$  coefficients  $c_{ip}$  par le processus itératif identique aux méthodes de DFT classiques.

*Remarque*

Les intégrales bi-électroniques à quatre centres valent :  $(ij|kl) \approx \sum_{\alpha,\beta}^{N^{\text{aux}},N^{\text{aux}}} (ij|\alpha)(\alpha|\beta)^{-1}(\beta|kl)$ .

Ceci correspond mathématiquement à l'insertion d'une résolution de l'identité, d'où le nom de cette méthode. Remarquons que cette relation serait strictement vraie si  $N^{\text{aux}}=N^2$ , car c'est une formule de changement de base.

*Conséquences*

Le calcul des termes bi-électroniques à quatre centres peut être approximé par des calculs de termes bi-électroniques à trois centres et les termes  $(\alpha|\beta)^{-1}$ . Si la base auxiliaire a une taille raisonnable (les bases auxiliaires utilisées dans le programme TURBOMOLE ont une dimension  $N^{\text{aux}} \approx 3N$ ), la matrice  $(\alpha|\beta)^{-1} = (\alpha|J^{-1}|\beta)$  peut être stockée en mémoire. Il ne sera donc pas nécessaire de la recalculer à chaque pas d'itération. Ahlrichs *et al.* ont montré que pour de grands systèmes la détermination des coefficients  $c_\alpha$  varie en  $N^2$ .<sup>21</sup>.

*Coût en temps de calcul*

Nous avons comparé les performances en temps de calcul avec la méthode HF/6-31G\* implémentée dans Gaussian 03<sup>23</sup> et la méthode ri-BLYP/SV(P) implémentée dans le logiciel TURBOMOLE. Les calculs Gaussian ont été effectués avec monoprocesseur, 2.5 GHz, tandis que les calculs TURBOMOLE<sup>22</sup> ont été effectués avec un monoprocesseur, 2.4 GHz.

		Nombre d'atomes	Nombre de fonctions de base (base auxiliaire)	Temps du cycle d'optimisation	Temps de calcul des fréquences
Gly <sub>6</sub> -Na <sup>+</sup>	HF	46	408	33 min	10h30
	ri-BLYP	46	425 (1053)	4 min	6h15
Gly <sub>10</sub> -Na <sup>+</sup>	HF	74	656	50min	
	ri-BLYP	74	648 (1705)	5 min	32h

**Tableau III-6 :** Comparaison du temps de calcul par cycle d'optimisation et par calcul de fréquences de vibration.

### III.D.2. Calibration de ri-BLYP pour le calcul d'énergies

Ce travail ayant pour but de déterminer des énergies absolues de complexation au sodium et des énergies relatives d'isomères pour identifier les structures de plus basse énergie, nous avons effectué une calibration de la méthode ri-BLYP sur ces deux aspects.

### III.D.2.a. Energies de liaisons H<sub>2</sub>O-Na<sup>+</sup>

Le groupe de D. Feller a étudié, à de très hauts niveaux de calcul, l'énergie de liaison du cation Na<sup>+</sup> avec une molécule d'eau. L'étude de cette énergie de liaison au niveau MP2 avec une taille croissante de bases permet d'estimer, par extrapolation, l'énergie de liaison avec une base infinie. La valeur déterminée de l'enthalpie standard de complexation à 298K par Feller est 97.5 kJ/mol.<sup>24</sup> La valeur déterminée au niveau ri-BLYP/TZVPP//ri-BLYP/SV(P) est de 98.5 kJ/mol, en très bon accord avec la limite MP2.

### III.D.2.b. Energie relative des isomères de Gly-Na<sup>+</sup>

Les énergies relatives des trois isomères de Gly-Na<sup>+</sup> les plus bas en énergie, aux niveaux MP2/GB1//MP2/PB1, B3LYP/GB1//B3LYP/PB1, BLYP/GB1//BLYP/PB1 et ri-BLYP/GB2//ri-BLYP/PB2 où PB1 est la base 6-31G\*, PB2 est la base SV(P) de même qualité que PB1, GB1 est la base 6-311+G(2d,2p) et GB2 la base TZVPP qui présente une extension identique à GB1, sont présentées dans le tableau III-7. A priori, la fonctionnelle d'échange corrélation B3LYP est plus fiable que celle de BLYP. Le niveau MP2 a largement fait les preuves de sa précision pour ce type de quantités. MP2 et B3LYP peuvent donc être utilisés comme niveaux de référence. Les résultats à tous les niveaux de calcul montrent que l'isomère Gly-Na<sup>+</sup> **a** est le plus bas en énergie, et que les isomères **b** et **Z** sont iso-énergétiques (leur différence d'énergie n'excède pas 1.5 kJ/mol sauf avec la méthode BLYP). Les énergies relatives au niveau ri-BLYP/GB2//ri-BLYP/PB2 sont en très bon accord avec les valeurs MP2/GB1//MP2/PB1, elles diffèrent de moins de 1 kJ/mol. Elles sont aussi en relativement bon accord avec les valeurs B3LYP/GB1//B3LYP/PB1 bien qu'elles soient sous-estimées pour les isomères **b** et **Z**, cette tendance est accentuée pour les énergies relatives BLYP/GB1//BLYP/PB1. Paradoxalement, les plus grandes différences de ri-BLYP sont avec BLYP. Ce test ne permet évidemment pas d'affirmer que la tendance de ri-BLYP à donner des résultats plus fiables que BLYP est générale.

---

Nom	MP2/GB1// MP2/PB1	B3LYP/GB1// B3LYP/PB1	BLYP/GB1// BLYP/PB1	ri-BLYP/GB2// ri-BLYP/PB2
GlyNa <sup>+</sup> <b>a</b>	0	0	0	<b>0</b>
GlyNa <sup>+</sup> <b>b</b>	9.5	10.7	6.2	<b>8.8</b>
GlyNa <sup>+</sup> <b>Z</b>	8.3	11.1	3.2	<b>7.3</b>

---

**Tableau III-7 :** Energies relatives à 298 K (kJ/mol) des isomères de Gly-Na<sup>+</sup>, selon différentes méthodes.

En conclusion, les énergies relatives au niveau ri-BLYP/GB2//ri-BLYP/PB2 montrent un très bon accord avec les énergies relatives aux niveaux MP2 et B3LYP. Une calibration sur des systèmes plus grands sera présentée en annexe du chapitre VI.

### III.D.3. Calibration de ri-BLYP pour le calcul des géométries

#### III.D.3.a. Géométrie de Gly-Na<sup>+</sup> a et b

La structure générale et quelques paramètres géométriques des isomères **a** et **b** de Gly-Na<sup>+</sup> sont présentés dans les figure III-7 et tableau III-8. La structure de l'isomère **a** présente une coordination de Na<sup>+</sup> avec l'oxygène du carbonyle et l'amine terminale. Les paramètres géométriques de l'isomère **a** montrent un excellent accord pour les différentes méthodes, la déviation maximale étant de 0.04 Å pour les liaisons Métal-Ligand et 1.5° pour les angles Ligand-Métal-Ligand. Dans le cas de l'isomère Gly-Na<sup>+</sup> **b**, d'une part, le métal est coordonné avec les deux oxygènes de la fonction acide, d'autre part l'hydrogène de la fonction acide trans interagit avec l'amine terminale par liaison hydrogène. Les méthodes corrélées sont en très bon accord pour l'interaction Métal-Ligand, alors que la méthode HF a tendance à privilégier l'interaction du métal avec l'oxygène du carbonyle plutôt que celui de l'hydroxyle. Cependant on peut remarquer que la liaison hydrogène intramoléculaire et a fortiori la liaison C-OH sont moins bien décrites avec les méthodes basées sur la fonctionnelle BLYP et la méthode HF comparées aux méthodes MP2 et B3LYP. Dans le cas des méthodes ri-BLYP et BLYP, la liaison hydrogène est beaucoup plus courte, cet effet étant inversé avec la méthode HF. L'équipe de D. Dixon a rencontré un désaccord similaire. Dans le cadre de la recherche conformationnelle du composé CH<sub>3</sub>-CO-Ala-NH-CH<sub>3</sub>, le conformère comportant un C<sub>10</sub> présente une liaison hydrogène NH---O anormalement longue au niveau BLYP/aug-cc-pVDZ, comparée à la valeur équivalente en MP2/aug-cc-pVDZ. Ce désaccord est absent avec les conformères présentant un C<sub>7</sub> ou un C<sub>5</sub>.<sup>25</sup> Les bonnes énergies relatives des isomères de Gly-Na<sup>+</sup> obtenues par calculs ri-BLYP pourraient être dues en partie à une compensation d'erreurs.

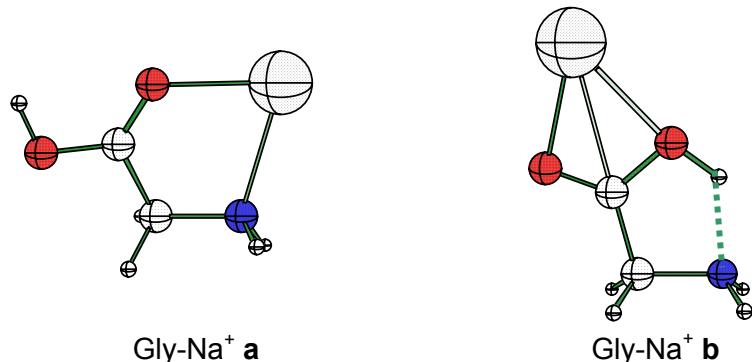


Figure III-7 : Structures des isomères Gly-Na<sup>+</sup> **a** et **b**

Structure Gly-Na <sup>+</sup> a	C=O	C-OH	O-H	N-H	C-N	Na <sup>+</sup> -O	Na <sup>+</sup> -N	∠CNHH	∠COH	∠CONa <sup>+</sup>	∠CNNa <sup>+</sup>
<b>ri-BLYP</b>	1.23	1.33	0.99	1.03	1.48	2.23	2.46	116.8	109.5	118.3	107.5
<b>BLYP</b>	1.24	1.34	0.99	1.03	1.49	2.24	2.43	115.9	108.7	116.8	107.7
<b>B3LYP</b>	1.23	1.32	0.98	1.02	1.48	2.22	2.42	116.7	109.4	117.5	107.9
<b>MP2</b>	1.23	1.33	0.98	1.02	1.47	2.26	2.44	116.2	108.8	117.2	108.4
<b>HF</b>	1.20	1.30	0.96	1.00	1.46	2.22	2.45	117.9	111.1	119.8	108.3

Structure Gly-Na <sup>+</sup> b	C=O	C-OH	O-H	H...N	N-H	Na <sup>+</sup> -O	Na <sup>+</sup> -OH	∠CNHH	∠COH	∠CONa <sup>+</sup>	∠CO <sub>H</sub> Na <sup>+</sup>
<b>ri-BLYP</b>	1.23	1.34	<b>1.08</b>	<b>1.61</b>	1.03	2.29	2.46	126.1	100.7	96.2	85.7
<b>BLYP</b>	1.24	1.35	<b>1.05</b>	<b>1.68</b>	1.03	2.28	2.44	124.1	101.1	96.3	86.0
<b>B3LYP</b>	1.22	1.34	1.02	1.75	1.01	2.26	2.44	124.5	103.2	97.5	86.0
<b>MP2</b>	1.23	1.34	1.02	1.75	1.01	2.32	2.43	123.8	102.9	95.7	87.9
<b>HF</b>	1.20	1.31	<b>0.97</b>	<b>1.98</b>	1.00	<b>2.23</b>	<b>2.61</b>	124.9	108.6	104.8	83.7

**Tableau III-8 :** Paramètres structuraux pour les isomères **a** et **b** de Gly-Na<sup>+</sup> (longueurs de liaison en Å et angles en °) optimisés avec la base 6-31G\*.

### III.D.4. Calibration de ri-BLYP pour le calcul de fréquences de vibration

Nous avons déjà vu que les fréquences calculées au niveau B3LYP/6-31G\* et multipliées par le scaling factor standard donnent un bon accord avec les résultats expérimentaux. Ce niveau de calcul sera donc considéré comme la référence dans ce qui va suivre. Les valeurs et les assignations des différentes fréquences obtenues aux niveaux B3LYP/6-31G\* (scaling factor standard : 0.9613), MP2/6-31G\* (0.9427), BLYP/6-31G\* (0.9940) et ri-BLYP/SV(P) (nous avons utilisé le scaling factor standard de la méthode BLYP/6-31G\* : 0.9940) sont présentées dans le tableau III-9. La comparaison des fréquences obtenues au niveau BLYP/6-31G\* avec les fréquences de référence montre un résultat globalement acceptable puisque la déviation moyenne sur les valeurs absolues est de 7.6 cm<sup>-1</sup>, avec une déviation maximale de 33 cm<sup>-1</sup> pour la vibration d'élongation O-H. Les déviations sont plus importantes dans la région des élongations X-H. Ce comportement est fortement accentué dans le cas des fréquences ri-BLYP où la déviation maximale est de 116 cm<sup>-1</sup> pour la vibration d'élongation O-H, pour une déviation moyenne de 17.6 cm<sup>-1</sup>. Ceci pose la question du scaling factor, en effet il n'existe pas actuellement dans la littérature de valeur standard au niveau ri-BLYP/SV(P). Nous avons donc choisi d'utiliser celui de la méthode BLYP/6-31G\*. Par ailleurs, il semble assez clair qu'un simple scaling factor ne permettra pas d'avoir un réel

accord entre les valeurs ri-BLYP et B3LYP. Il serait donc vraisemblablement plus judicieux d'utiliser deux scaling factors, pour les fréquences inférieures à  $2000\text{ cm}^{-1}$  d'une part et pour les fréquences supérieures à  $2000\text{ cm}^{-1}$  d'autre part. Ces dernières concernent les elongations X-H plus anharmoniques. En effet, il n'existe aucune raison de principe pour utiliser un facteur d'échelle commun à tous les types de modes.

En conclusion, la méthode ri-BLYP/SV(P) en l'absence de plusieurs scaling factors standard ne semble pas une solution adaptée pour le calcul des fréquences de vibration, avec une précision permettant de les confronter à l'expérience.

### **III.D.5. Conclusion**

Les valeurs énergétiques obtenues par calculs ri-BLYP/TZVPP//ri-BLYP/SV(P) sont en très bon accord avec les valeurs de références, que ce soit au niveau B3LYP/GB1//B3LYP/PB1 ou MP2/GB1//MP2/PB1. Les géométries ri-BLYP/SV(P) sont aussi en bon accord avec les géométries B3LYP/PB1 et MP2/PB1, excepté pour les liaisons hydrogène intramoléculaires dont la longueur semble être sous-estimée. Quant à la simulation de spectres infrarouges, il s'avère que ri-BLYP n'est pas une méthode adaptée. Il apparaît donc, au vu des performances en temps de calcul, que ri-BLYP est une méthode adaptée pour l'énergétique et la géométrie de gros systèmes. Une calibration complémentaire sera présentée au chapitre VI pour de plus grands systèmes.

modes Gly-Na <sup>+</sup> a	B3LYP <i>*0.9613</i>	MP2 <i>*0.9427</i>	BLYP <i>*0.9940</i>	ri-BLYP <i>*0.9940</i>	$\Delta f$ BLYP/ B3LYP	$\Delta f$ ri-BLYP/ B3LYP	attribution
1	84	80	83	95	0	11	Twist COOH
2	99	105	102	109	2	9	Twist CCNH <sub>2</sub>
3	<b>208</b>	<b>210</b>	<b>211</b>	<b>206</b>	<b>2</b>	<b>-2</b>	<i>ν</i> Na-O + <i>ν</i> Na-N
4	<b>242</b>	<b>228</b>	<b>241</b>	<b>247</b>	<b>-2</b>	<b>5</b>	<i>ν</i> Na-O + <i>ν</i> Na-N
5	324	319	325	324	1	0	$\delta$ NCC+ $\delta$ CCO
6	<b>410</b>	<b>407</b>	<b>413</b>	<b>427</b>	<b>3</b>	<b>18</b>	$\tau$ NH <sub>2</sub>
7	478	476	477	487	-1	9	
8	513	513	513	514	0	1	$\omega$ COOH
9	<b>637</b>	<b>629</b>	<b>634</b>	<b>643</b>	<b>-3</b>	<b>6</b>	$\delta$ NCC+ $\delta$ CCO
10	<b>682</b>	<b>677</b>	<b>690</b>	<b>694</b>	<b>8</b>	<b>12</b>	$\omega$ O-H
11	826	834	817	823	<b>-10</b>	-3	$\nu$ C-C
12	911	914	911	911	0	0	$\tau$ NH <sub>2</sub> + $\tau$ CH <sub>2</sub>
13	<b>1012</b>	<b>1027</b>	<b>991</b>	<b>996</b>	<b>-21</b>	<b>-16</b>	<i>ν</i> C-N + $\omega$ NH <sub>2</sub>
14	<b>1076</b>	<b>1080</b>	<b>1083</b>	<b>1072</b>	<b>8</b>	<b>-4</b>	$\omega$ NH <sub>2</sub>
15	<b>1166</b>	<b>1153</b>	<b>1158</b>	<b>1156</b>	<b>-8</b>	<b>-11</b>	$\tau$ NH <sub>2</sub> + $\tau$ CH <sub>2</sub>
16	1172	1162	1171	1163	-1	-9	$\nu$ C-OH
17	1293	1285	1294	1280	1	<b>-13</b>	$\tau$ NH <sub>2</sub> + $\tau$ CH <sub>2</sub>
18	1333	1326	1337	1324	4	-9	$\tau$ NH <sub>2</sub> + $\tau$ CH <sub>2</sub>
19	<b>1420</b>	<b>1426</b>	<b>1406</b>	<b>1391</b>	<b>-14</b>	<b>-29</b>	$\delta$ COH+ <i>ν</i> C-O + $\omega$ CH <sub>2</sub>
20	1439	1445	1447	1415	8	<b>-24</b>	$\delta$ CH <sub>2</sub>
21	<b>1641</b>	<b>1631</b>	<b>1652</b>	<b>1608</b>	<b>11</b>	<b>-33</b>	$\delta$ NH <sub>2</sub>
22	<b>1725</b>	<b>1710</b>	<b>1700</b>	<b>1725</b>	<b>-25</b>	<b>0</b>	<i>ν</i> C=O
23	2956	2946	2976	2942	<b>20</b>	<b>-15</b>	$\nu^s$ CH <sub>2</sub>
24	2997	2998	3015	2988	<b>18</b>	-8	$\nu^{as}$ CH <sub>2</sub>
25	3340	3308	3339	3285	-1	<b>-55</b>	$\nu^s$ NH <sub>2</sub>
26	<b>3409</b>	<b>3386</b>	<b>3408</b>	<b>3351</b>	0	<b>-57</b>	$\nu^{as}$ NH <sub>2</sub>
27	<b>3517</b>	<b>3441</b>	<b>3484</b>	<b>3401</b>	<b>-33</b>	<b>-116</b>	<i>ν</i> O-H

Tableau III-9 : Fréquences calculées pour différentes méthodes avec les bases 6-31G\* ou SV(P) pour Gly-Na<sup>+</sup> a (les valeurs en italique correspondent aux bandes d'intensité supérieure à 20 km/mol).

## Bibliographie

- [1] CLIO, <http://www.lure.u-psud.fr/CLIO.HTM>.
- [2] Ortega, J. M.; Berset, J. M.; Chaput, R.; Glotin, F.; Humbert, G.; Jaroszynski, D.; Joly, P.; Kergosien, B.; Lesrel, J.; Marcouille, O.; Peremans, A.; Prazeres, R.; Tadjeddine, A., Activities of the CLIO infrared facility. *Nucl. Instrum. Methods Phys. Res. Sect. A-Accel. Spectrom. Dect. Assoc. Equip.* **1996**, 375, (1-3), 618-625.
- [3] Maitre, P.; Le Caer, S.; Simon, A.; Jones, W.; Lemaire, J.; Mestdagh, H. N.; Heninger, M.; Mauclaire, G.; Boissel, P.; Prazeres, R.; Glotin, F.; Ortega, J. M., Ultrasensitive spectroscopy of ionic reactive intermediates in the gas phase performed with the first coupling of an IR FEL with an FTICR-MS. *Nucl. Instrum. Methods Phys. Res. Sect. A-Accel. Spectrom. Dect. Assoc. Equip.* **2003**, 507, (1-2), 541-546.
- [4] Marshall, A. G., Milestones in Fourier transform ion cyclotron resonance mass spectrometry technique development. *Int. J. Mass Spectrom.* **2000**, 200, (1-3), 331-356.
- [5] Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S., Fourier transform ion cyclotron resonance mass spectrometry: A primer. *Mass Spectrom. Rev.* **1998**, 17, (1), 1-35.
- [6] Hoyau, S.; Norrman, K.; McMahon, T. B.; Ohanessian, G., A quantitative basis for a scale of  $\text{Na}^+$  affinities of organic and small biological molecules in the gas phase. *J. Am. Chem. Soc.* **1999**, 121, (38), 8864-8875.
- [7] NIST Chemistry WebBook, <http://webbook.nist.gov/chemistry/>.
- [8] Strittmatter, E. F.; Williams, E. R., Computational approach to the proton affinities of Gly<sub>n</sub> (n = 1-10). *Int. J. Mass Spectrom.* **1999**, 187, 935-948.
- [9] Ohanessian, G., Interaction of MALDI matrix molecules with  $\text{Na}^+$  in the gas phase. *Int. J. Mass Spectrom.* **2002**, 219, (3), 577-592.
- [10] Bourcier, S.; Hoppilliard, Y., B3LYP DFT molecular orbital approach, an efficient method to evaluate the thermochemical properties of MALDI matrices. *Int. J. Mass Spectrom.* **2002**, 217, (1-3), 231-244.
- [11] Kussmann, M.; Nordhoff, E.; RahbekNielsen, H.; Haebel, S.; RosselLarsen, M.; Jakobsen, L.; Gobom, J.; Mirgorodskaya, E.; KrollKristensen, A.; Palm, L.; Roepstorff, P., Matrix-assisted laser desorption/ionization mass spectrometry sample preparation techniques designed for various peptide and protein analytes. *J. Mass Spectrom.* **1997**, 32, (6), 593-601.
- [12] Weiner, P. K.; Kollman, P. A., Amber - Assisted Model-Building with Energy Refinement - a General Program for Modeling Molecules and Their Interactions. *J. Comput. Chem.* **1981**, 2, (3), 287-303.

- [13] HyperChem 6.0 : Hypercube, Inc., 1115 NW 4th Street, Gainesville, FL 32601 (USA).
- [14] Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A., A 2nd Generation Force-Field for the Simulation of Proteins, Nucleic-Acids, and Organic-Molecules. *J. Am. Chem. Soc.* **1995**, 117, (19), 5179-5197.
- [15] McMahon, T. B.; Ohanessian, G., An experimental and ab initio study of the nature of the binding in gas-phase complexes of sodium ions. *Chem.-Eur. J.* **2000**, 6, (16), 2931-2941.
- [16] Hohenberg, P.; Kohn, W., Inhomogeneous Electron Gas. *Phys. Rev. B* **1964**, 136, (3B), B864-&.
- [17] Kohn, W.; Sham, L. J., Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, 140, (4A), 1133-&.
- [18] Eichkorn, K.; Treutler, O.; Ohm, H.; Haser, M.; Ahlrichs, R., Auxiliary Basis-Sets to Approximate Coulomb Potentials. *Chem. Phys. Lett.* **1995**, 240, (4), 283-289.
- [19] Eichkorn, K.; Treutler, O.; Ohm, H.; Haser, M.; Ahlrichs, R., Auxiliary Basis-Sets to Approximate Coulomb Potentials (Vol 240, Pg 283, 1995). *Chem. Phys. Lett.* **1995**, 242, (6), 652-660.
- [20] M. Von Arnim, R. A., Performance of parallel TURBOMOLE for Density Functional Calculations. *J. Comput. Chem.* **1998**, 19, (15), 1746-1757.
- [21] Sierka, M.; Hogekamp, A.; Ahlrichs, R., Fast evaluation of the Coulomb potential for electron densities using multipole accelerated resolution of identity approximation. *Journal of Chemical Physics* **2003**, 118, (20), 9136-9148.
- [22] TURBOMOLE 5.6 , Ahlrichs, R. et al, Quantum Chemistry Group, University of Karlsruhe (Germany).
- [23] Gaussian03, revision B.05, Frisch, M. J. et al., Gaussian, Inc., Pittsburgh, PA, 2003.
- [24] Feller, D.; Glendening, E. D.; Woon, D. E.; Feyereisen, M. W., An Extended Basis-Set Ab-Initio Study of Alkali-Metal Cation-Water Clusters. *J. Chem. Phys.* **1995**, 103, (9), 3526-3542.
- [25] Vargas, R.; Garza, J.; Hay, B. P.; Dixon, D. A., Conformational study of the alanine dipeptide at the MP2 and DFT levels. *J. Phys. Chem. A* **2002**, 106, (13), 3213-3218.

## **Chapitre IV :**

**Rôle structurant de  $\text{Na}^+$  sur les  
acides aminés**

**Glycine et Proline**

## I. Introduction

### I.A. Structures des acides aminés isolés

#### I.A.1. Acides aminés isolés en solution aqueuse

Nous avons vu dans le chapitre I que les acides aminés sont sous forme zwitterionique en solution aqueuse dans une large gamme de pH. Ils possèdent donc deux charges opposées, qui peuvent être localisées sur la chaîne principale : un ion ammonium du côté N terminal et un ion carboxylate du côté C terminal. L'une de ces charges peut aussi être localisée sur la chaîne latérale, si celle-ci comporte le site le plus acide ou le plus basique : c'est par exemple le cas de l'arginine. Dans tous les cas, ces deux charges sont stabilisées par interaction avec des molécules de solvant.

#### I.A.2. Acides aminés isolés en phase gazeuse

L'absence de molécules de solvant en phase gazeuse ne permet plus de stabiliser les charges et donc la forme zwitterionique par rapport à ses isomères sans séparation de charge. Cependant cette déstabilisation peut être compensée si la réaction de transfert de proton de l'acide vers l'amine est exothermique. La fonction acide étant commune à tous les acides aminés, il faut donc que l'affinité protonique et la basicité en phase gazeuse de l'acide aminé soient les plus grandes possibles. Cependant pour les vingt acides aminés naturels, même ceux qui comportent les chaînes latérales les plus basiques comme l'arginine, les études expérimentales et théoriques existantes indiquent que le bilan de la réaction acido-basique est défavorable : ils présentent tous une structure sans charge formelle ou non-zwitterionique.

##### I.A.2.a. Cas de la glycine

La glycine, qui est le plus petit acide aminé, a été le plus étudié expérimentalement et théoriquement. Tous les résultats convergent pour montrer que la glycine est sous forme non-zwitterionique en phase gazeuse.

La surface d'énergie potentielle (SEP) de la glycine a largement été explorée dans les années 90, que ce soient par des méthodes *ab initio* ou de la fonctionnelle de la densité.<sup>1</sup> Les trois conformères les plus bas en énergie, ainsi que la structure zwitterionique, sont présentés par la Figure III-2. Dans Gly-**a** et Gly-**c**, les deux hydrogènes de l'amine terminale forment des liaisons hydrogène avec l'oxygène du carbonyle et de l'hydroxyle,

respectivement. Gly-**b** présente une fonction acide trans dont l'hydrogène est orienté pour interagir avec l'azote de l'amine terminale. Ces trois conformations se trouvent dans une gamme de 8 kJ/mol. Quant à la forme zwitterionique, elle ne correspond pas à un extremum sur la surface de potentiel.

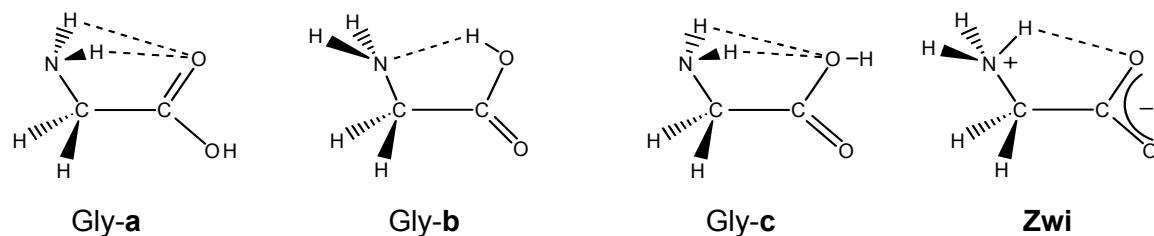


Figure IV-1 : Conformères les plus bas en énergie et zwitterion de la glycine isolée

L'étude des conformations de la glycine a été aussi menée expérimentalement, par plusieurs groupes. Une première série d'expériences de spectroscopie micro-ondes a permis d'identifier les conformères Gly-**b**, qui possède un moment dipolaire important, et Gly-**a**.<sup>3, 4</sup> Grâce à la spectroscopie micro-ondes à jet supersonique couplée à des calculs *ab initio* de constantes de rotation, les trois conformères Gly-**a**, **b**, **c** ont pu être identifiés.<sup>5</sup> De même, l'équipe de Reva a enregistré le spectre infrarouge de la glycine isolée en matrice.<sup>6</sup> L'étude de la région des élongations C=O, autour de  $1700 \text{ cm}^{-1}$ , montre la présence de trois conformères. La comparaison du spectre obtenu expérimentalement avec les spectres infrarouges simulés par des calculs *ab initio* des conformères de plus basse énergie a permis l'identification des trois conformères Gly-**a**, **b**, **c**.

### I.A.2.b. Cas de la proline

L'étude conformationnelle *ab initio* de la proline a mis en évidence quatre conformères de plus basse énergie dans une gamme de 10 kJ/mol, qui diffèrent d'une part par la conformation du cycle à cinq chaînons et par l'orientation de la fonction carboxylique (figure IV-2).<sup>7-10</sup> Les deux structures les plus basses en énergie Pro-**b1** et Pro-**b2** sont équivalentes à la structure Gly-**b** pour l'orientation du groupement carboxylique et diffèrent par la conformation du cycle. Les structures Pro-**a1** et Pro-**a2** (environ 5 kJ/mol moins stables) sont analogues à Gly-**a**; comme Pro-**b1** et Pro-**b2**, elles diffèrent par la conformation du cycle.

Ces conformations ont été confirmées expérimentalement, notamment par spectroscopie infrarouge en matrice par l'équipe de Reva.<sup>8</sup>

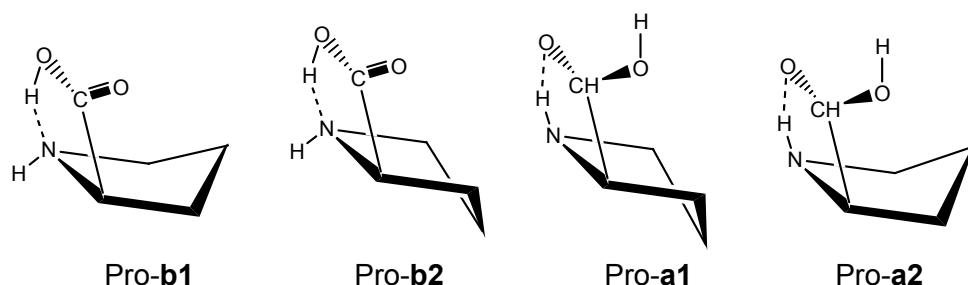


Figure IV-2 : Conformères les plus bas en énergie de la proline isolée

### I.A.2.c. Cas de l'arginine

L'arginine est le plus basique des acides aminés, de par la fonction guanidine de sa chaîne latérale. Williams *et al.*<sup>11</sup> ont émis l'hypothèse que cette basicité pourrait suffire pour que l'isomère le plus stable soit le zwitterion. Des expériences de spectroscopie<sup>12</sup> semblent contradictoires avec cette hypothèse. Des calculs à des niveaux élevés,<sup>13, 14</sup> en accord avec ces expériences, indiquent que le conformère zwitterionique le plus bas en énergie est moins stable d'environ 12 kJ/mol que le meilleur conformère sans charge formelle. L'importance de la basicité de la chaîne latérale a été exploitée par Julian et Jarrold<sup>15</sup>, qui ont montré par calculs que des méthylations successives de l'arginine sur le groupement guanidine, en augmentant son affinité protonique, stabilisent la forme zwitterionique jusqu'à en faire la plus stable.

Des travaux expérimentaux par spectroscopie de photoélectron sur l'anion de l'arginine ont également permis de caractériser le zwitterion.<sup>16</sup> La méthode consiste à former les anions  $\text{Arg}^-$  en phase gazeuse, et à mesurer l'énergie de cohésion de l'électron excédentaire. La plupart des molécules usuelles n'ont pas d'affinité électronique significative, mais lorsqu'un moment dipolaire supérieur à 2.5 D existe, des états spécifiques, appelés « dipole-bound », peuvent être formés. Le moment dipolaire très important des zwitterions s'y prête particulièrement. Le spectre de photodétachement de  $\text{Arg}^-$  est cohérent avec la co-existence de structures zwitterionique et non-zwitterionique, et les énergies de détachement de l'électron sont en bon accord avec les valeurs calculées.<sup>17</sup> Bien que l'électron interagisse peu avec la molécule dans ce type d'état, il stabilise légèrement les structures ayant le plus fort moment dipolaire ; ces observations n'indiquent donc pas que le zwitterion et son isomère sans charge isolé ont la même énergie, mais que cette perturbation faible suffit à les rendre quasi dégénérés.

## I.B. Stabilisation du zwitterion en phase gazeuse

### I.B.1. Microsolvatation

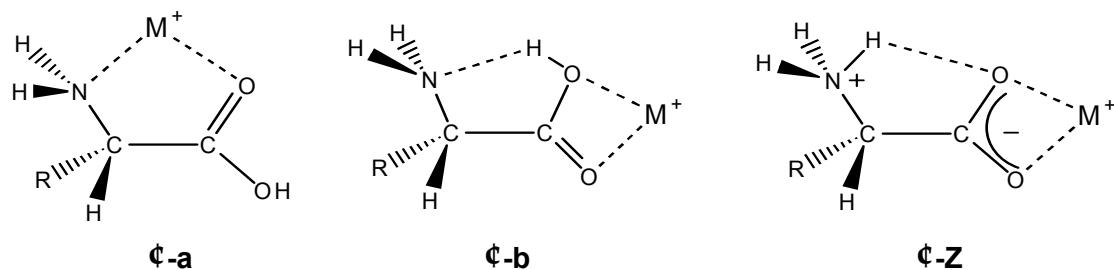
La différence de structure des acides aminés en phase gazeuse et en solution aqueuse suggère la question : combien de molécules d'eau sont-elles nécessaires pour que le zwitterion devienne la structure la plus stable ? Ceci suggère de mimer la situation en solution aqueuse en micro-solvantant l'acide aminé. Cette approche a fait l'objet de nombreux travaux théoriques, dans le cas de la glycine,<sup>18-25</sup> qui indiquent qu'une seule molécule d'eau suffit pour que la structure zwitterionique soit un minimum local sur la surface d'énergie potentielle, que deux molécules ne suffisent pas pour en faire l'isomère le plus stable, et qu'avec trois molécules les deux types de structure sont d'énergies très proches. La difficulté d'obtenir une valeur définitive tient au grand nombre de structures envisageables, avec une compétition entre liaisons hydrogène de molécules d'eau avec la glycine (y compris pontant entre l'acide et l'amine) et de molécules d'eau entre elles. Cette question a également été abordée expérimentalement, par spectroscopie de photoélectron sur les anions  $\text{AA}(\text{H}_2\text{O})_n^-$ , comme pour Arg<sup>-</sup> ci-dessus.<sup>26</sup> Pour la glycine, un signal attribuable à l'état « dipole-bound » du zwitterion est détecté à partir de  $n=5$ . Cette valeur est de 4 pour la phénylalanine et le tryptophane. Comme pour Arg, la présence de l'électron supplémentaire stabilise probablement le zwitterion ; ces valeurs doivent donc être considérées comme des limites inférieures du nombre de molécules d'eau nécessaires pour que le zwitterion soit la structure la plus stable pour le système neutre.

### I.B.2. Complexation par un alcalin

Une autre façon de favoriser la forme zwitterionique par rapport à ses isomères sans charge formelle est de complexer l'acide aminé par un cation. L'interaction électrostatique qui en résulte, formant un pont salin entre le cation, le groupe carboxylate et le groupe ammonium (ou la chaîne latérale protonée), est particulièrement stabilisante.

De nombreuses équipes ont étudié cette stabilisation par des méthodes de chimie quantique. Il apparaît que le complexe zwitterionique devient un minimum sur la surface d'énergie potentielle, qui peut être global ou local. Le rôle du cation métallique a notamment été étudié par les équipes de J. Bertran et G. Ohanessian dans le cas de la glycine.<sup>27-29</sup> Trois types de structures de basse énergie peuvent être envisagés (Figure IV-3). Dans **a**, le cation est complexé à l'azote de l'amine et à l'oxygène du carbonyle, l'acide aminé ayant une structure proche de celle de Gly-**a** (Figure IV-1) avec l'amine inversée. Dans **b**,  $\text{Na}^+$  est coordonné aux deux oxygènes de l'acide, le squelette de l'acide aminé étant identique à

celui de Gly-**b**. Enfin dans **c-Z**,  $\text{Na}^+$  est coordonné au groupement carboxylate de la forme zwitterionique.



**Figure IV-3 :** Structures de basses énergies des complexes acide aminé - cation métallique

Pour la glycine, les calculs indiquent que les petits alcalins ont tendance à favoriser la structure **c-a**, les plus gros le complexe **c-b**. Les métaux ayant un potentiel d'ionisation élevé, comme  $\text{Cu}^+$  ou  $\text{Au}^+$ , favorisent la structure **c-a** pour maximiser l'interaction de transfert de charge, alors que les métaux divalents favorisent le zwitterion.<sup>27</sup> Cependant si ces structures dépendent du cation métallique en interaction, elles dépendent bien sur de la nature de l'acide aminé. Le premier cas de complexe ayant le zwitterion pour structure la plus stable, Pro- $\text{Na}^+$ , a été identifié par Hoyau *et al.*<sup>30</sup> L'équipe de Bowers a étudié l'influence de l'acide aminé sur la structure des complexes au sodium et au rubidium. Ils ont montré une dépendance quasi linéaire de la stabilité du zwitterion avec l'affinité protonique de l'acide aminé.<sup>31</sup>

### I.B.3.Cas de Gly- $\text{Na}^+$

Parmi d'autres, S. Hoyau et G. Ohanessian ont exploré la surface d'énergie potentielle de Gly- $\text{Na}^+$  par calculs *ab initio*.<sup>27, 28</sup> Les trois isomères **c-a**, **c-b** et **c-Z** sont relativement bas en énergie, dans une gamme de 10 kJ/mol. La forme zwitterionique **c-Z** n'est pas la plus stable, elle se trouve environ 10 kJ/mol au dessus du complexe **c-a**.

Plusieurs types d'expériences ont été menés pour tenter d'identifier la structure la plus stable. Cependant les expériences d'échange H/D et de mobilité ionique se sont révélées inefficaces pour la différenciation de la forme zwitterionique de la forme sans charge formelle. Comme nous l'avons montré dans le chapitre II, la technique d'échange H/D présente certaines limites, qui ont été illustrées dans le cas des oligoglycines sodiés  $\text{Gly}_n\text{Na}^+$  ( $n=1-5$ ) par l'équipe de Beauchamp.<sup>32</sup> Dans le cas de Gly- $\text{Na}^+$ , les trois échanges rapides de deutérium observés semblaient indiquer une forme zwitterionique ; cependant l'étude du profil réactionnel d'échange H/D de Gly- $\text{Na}^+$  **c-a** montre qu'à l'intérieur du complexe formé avec la molécule deutérée, Gly- $\text{Na}^+$  peut facilement s'isomériser en une forme zwitterionique qui échange très rapidement ses hydrogènes ammonium. Ainsi les résultats d'échange H/D

ne permettent pas de conclure quant à la structure du complexe. Dans le cas des expériences de mobilité ionique, la forme zwitterionique et la forme non-zwitterionique ont des sections efficaces de collision trop proches pour pouvoir attribuer la valeur expérimentale avec certitude.<sup>33</sup>

La comparaison des affinités pour  $\text{Na}^+$  des acides aminés et de leurs esters a également été étudiée en utilisant la méthode cinétique. Les esters sont un peu plus riches en électrons et polarisables que les acides correspondants, leurs affinités pour  $\text{Na}^+$  doivent donc être supérieures, si les modes d'attachement sont identiques. Si l'acide forme un zwitterion, structure impossible pour l'ester correspondant, l'ordre d'affinité est inversé. Cette méthode permet effectivement de « repérer » les zwitterions probables, en particulier Pro- $\text{Na}^+$ , mais cette approche très indirecte est entachée d'exceptions.<sup>34</sup> Enfin, l'équipe de P. Armentrout a déterminé l'enthalpie standard de complexation de la glycine avec  $\text{Na}^+$  en mesurant l'énergie du seuil de dissociation.<sup>35</sup> La valeur obtenue de  $164 \pm 6$  kJ/mol se trouve dans la gamme des valeurs calculées pour les différents isomères du complexe. Elle ne permet donc pas d'identifier la structure zwitterionique ou non de la glycine dans le complexe formé.

On peut enfin combiner les deux façons de tenter de stabiliser le zwitterion, en considérant la micro-solvatation de complexes tels que Gly- $\text{Na}^+$ . Des mesures de seuil de dissociation d'une molécule d'eau de  $\text{GlyNa}^+(\text{H}_2\text{O})_n$  ( $n=1-4$ ), interprétées à l'aide de calculs quantiques, indiquent que la glycine n'est pas zwitterionique, même si cet isomère est toujours proche du plus stable.<sup>36</sup> Des résultats similaires ont été obtenus par calculs pour  $\text{GlyLi}^+(\text{H}_2\text{O})_n$  ( $n=1-3$ ).<sup>37</sup>

#### I.B.4.Cas de Pro- $\text{Na}^+$

Si des mesures d'affinité relative de Pro et d'autres acides aminés pour  $\text{Na}^+$  existent,<sup>38, 39</sup> la structure du complexe Proline- $\text{Na}^+$  a uniquement été étudiée par des méthodes de chimie quantique.<sup>30, 34</sup> La proline possède une affinité protonique supérieure de 33 kJ/mol à celle de la glycine, parce que son amine est secondaire alors que celle de la glycine est primaire. Cette différence permet d'inverser l'ordre énergétique des complexes. L'isomère de plus basse énergie est alors le complexe zwitterionique **c-Z**. Le second est **c-a**, moins stable d'environ 20 kJ/mol que **c-Z**. Quant à **c-b**, c'est un minimum local très peu profond sur la surface d'énergie potentielle. Ce point sera décrit plus en détail dans la partie suivante.

En 2003 il n'existait aucune tentative d'identification expérimentale de la structure zwitterionique de Pro- $\text{Na}^+$ . Récemment, l'équipe de E. Williams a étudié, par des expériences de BIRD, la structure des complexes entre  $\text{Na}^+$  et des dérivés méthylés de la proline :  $\alpha$ -Me-Pro, N-Me-Pro et ProOMe, par mesure de l'énergie de seuil de dissociation

d'une molécule d'eau.<sup>40</sup> Les affinités protoniques et les basicités en phase gazeuse indiquent que la tendance à former un zwitterion est meilleure pour N-Me-Pro puis pour Pro et enfin pour  $\alpha$ -Me-Pro, l'ester méthylé ne pouvant pas former de zwitterion. Ces auteurs ont montré que les enthalpies de dissociation de la molécule d'eau sont proches dans le cas de la  $\alpha$ -Me-Pro et N-Me-Pro (différence d'énergie de 2 kJ/mol) et plus faibles de 3 à 5 kJ/mol par rapport à ProOMe. Ceci indique que les structures de  $\alpha$ -Me-Pro et N-Me-Pro sont proches et donc certainement zwitterioniques, contrairement à ProOMe. Cependant les valeurs d'enthalpies expérimentales sont en désaccord avec les valeurs issues de calculs de DFT, qui sont ~15 kJ/mol plus grandes.

## I.C. Conclusion

Dans le cas de Gly- $\text{Na}^+$  et Pro- $\text{Na}^+$ , les calculs montrent que la structure de l'acide aminé est non-zwitterionique et zwitterionique, respectivement. Cependant, la plupart des expériences de détermination structurale que nous avons décrites au chapitre II et ci-dessus sont indirectes, et se sont avérées inefficaces quant à la détermination de la structure de ces complexes.

## II. Spectroscopie infrarouge des complexes Gly- $\text{Na}^+$ et Pro- $\text{Na}^+$

### II.A. Résumé

Comme nous l'avons vu dans le chapitre III, le couplage du spectromètre de masse FT-ICR MICRA avec le LEL CLIO a permis de déterminer la structure de divers types d'ions gazeux, parmi lesquels des ions organiques<sup>41, 42</sup> ou des complexes organométalliques du fer<sup>43, 44</sup>. La question toujours ouverte de l'isométrie dans les complexes d'acides aminés avec  $\text{Na}^+$  en faisait donc des candidats intéressants pour l'identification par spectroscopie infrarouge grâce au couplage MICRA-CLIO.

Dans un premier temps, nous avons développé un nouveau mode d'introduction d'ions biologiques, basé sur la technique MALDI et adapté à la géométrie du spectromètre de masse MICRA (cf. chapitre III). Ceci a permis l'enregistrement des spectres IRMPD expérimentaux de Pro- $\text{Na}^+$  et Gly- $\text{Na}^+$ . La comparaison du spectre IRMPD expérimental pour le complexe Pro- $\text{Na}^+$  avec les différents spectres infrarouge des complexes  $\ddot{\text{c}}\text{-a}$ ,  $\ddot{\text{c}}\text{-b}$  et  $\ddot{\text{c}}\text{-Z}$  montre la présence exclusive du complexe zwitterionique. Dans le cas de Gly- $\text{Na}^+$ , nous avons identifié de manière univoque le complexe  $\ddot{\text{c}}\text{-a}$ , cependant la présence de l'isomère zwitterionique ne peut être exclue. Ces résultats, en bon accord avec les calculs, montrent que nous avons formé les complexes les plus bas en énergie.

Ces expériences constituent d'une part les premières caractérisations d'ions biologiques en phase gazeuse par spectroscopie infrarouge, et d'autre part la première preuve de l'existence d'un zwitterion d'acide aminé en phase gazeuse.

## **II.B. Article<sup>45</sup>**

## Vibrational Signature of Charge Solvation vs Salt Bridge Isomers of Sodiated Amino Acids in the Gas Phase

Catherine Kapota,<sup>†</sup> Joël Lemaire,<sup>‡,§</sup> Philippe Maître,<sup>\*,‡</sup> and Gilles Ohanessian<sup>\*,†</sup>

Contribution from the Laboratoire des Mécanismes Réactionnels, UMR 7651 du CNRS, Ecole Polytechnique, 91128 Palaiseau CEDEX, France, Laboratoire de Chimie Physique, UMR 8000 du CNRS, Université de Paris-Sud, 91405 Orsay CEDEX, France, and Laboratoire pour l'Utilisation du Rayonnement Electromagnétique, UMR 130 du CNRS, Université de Paris-Sud, CLIO-BP34, 91405 Orsay CEDEX, France

Received June 27, 2003; E-mail: gilles.ohanessian@polytechnique.fr

**Abstract:** The vibrational spectra of the gaseous sodium complexes of glycine ( $\text{Gly}-\text{Na}^+$ ) and proline ( $\text{Pro}-\text{Na}^+$ ) have been recorded in the spectral range  $1150$ – $2000\text{ cm}^{-1}$ . The complexes were formed by matrix-assisted laser desorption–ionization, introduced in the cell of a Fourier transform ion cyclotron resonance mass spectrometer, and their infrared spectra were recorded using photons of variable energy emitted by a free electron laser. Photon absorption was probed by the diminished intensity of the parent ion, due to its infrared-induced dissociation into bare sodium cation and the free amino acid and the appearance of  $\text{Na}^+$ . The observed absorption bands are assigned using *ab initio* computations of the IR spectra of the lowest energy isomers in each case. They provide the first experimental evidence that the salt bridge isomer is formed in the case of  $\text{Pro}-\text{Na}^+$ . In contrast, charge solvation by chelation of  $\text{Na}^+$  between nitrogen and the carbonyl oxygen seems to be most favorable for  $\text{Gly}-\text{Na}^+$ , but a mixture of isomers cannot be ruled out in this case.

### I. Introduction

$\alpha$ -Amino acids (AA) are known to be in their zwitterionic form in solution over a wide range of pH values.<sup>1</sup> In this isomer, acidic groups are deprotonated and the basic groups protonated. In the gas phase, however, all available evidence points to non-zwitterionic isomers being more favorable. Extensive studies on gaseous Gly have shown that the zwitterion is not even a local minimum on the potential energy surface.<sup>2</sup> There are however strong indications that the arginine zwitterion with a protonated side chain and deprotonated carboxyl group is of similar energy as its neutral counterpart without any charged sites.<sup>3</sup>

Two different ways may be envisioned to induce the neutral to zwitterion transition in the gas phase: either microsolvate the AA until a cluster of solvent molecules mimics the effects of a dilute solution<sup>4</sup> or else favor the zwitterion by attaching a charged species, such as an atomic metal cation, to the AA. In the latter case, the strong electrostatic interaction between the metal charge and the large dipole of the zwitterion may be

enough to offset the intrinsic preference of the AA for the neutral isomer. The second approach is used in this work.

Several studies<sup>5–15</sup> have established that when a metal cation such as  $\text{Na}^+$  is attached to an AA, several low-energy isomers can be formed, which fall into two broad categories: (i) charge solvation (CS), in which the metal charge is stabilized by

- (4) (a) Jensen, J. H.; Gordon, M. S. *J. Am. Chem. Soc.* **1995**, *117*, 8159. (b) Ding, Y.; Krogh-Jespersen, K. *J. Comput. Chem.* **1996**, *17*, 338. (c) Tunon, I.; Silla, E.; Ruiz-Lopez, M. F. *Chem. Phys. Lett.* **2000**, *321*, 433. (d) Bandyopadhyay, P.; Gordon, M. S. *J. Chem. Phys.* **2000**, *113*, 1104. (e) Kassab, E.; Langlet, J.; Eyleth, E.; Akacem, Y. *J. Mol. Struct. (THEOCHEM)* **2000**, *531*, 267. (f) Wang, W.; Pu, X.; Zheng, W.; Wong, N.-B.; Tian, A. *J. Mol. Struct. (THEOCHEM)* **2003**, *626*, 127.
- (5) Jockusch, R. A.; Price, W. D.; Williams, E. R. *J. Phys. Chem. A* **1999**, *103*, 9266.
- (6) Bouchonnet, S.; Hoppilliard, Y. *Org. Mass Spectrom.* **1992**, *27*, 71. (b) Jensen, F. J. *Am. Chem. Soc.* **1992**, *114*, 9533. (c) Hoyau, S.; Ohanessian, G. *Chem. Eur. J.* **1998**, *4*, 1561. (d) Wytenbach, T.; Witt, M.; Bowers, M. T. *Int. J. Mass Spectrom.* **1999**, *183*, 243.
- (7) Hoyau, S.; Norrman, K.; McMahon, T. B.; Ohanessian, G. *J. Am. Chem. Soc.* **1999**, *121*, 8864.
- (8) Marino, T.; Russo, N.; Toscano, M. *J. Phys. Chem. B* **2003**, *107*, 2588.
- (9) Shoeib, T.; Hopkinson, A. C.; Siu, K. W. M. *J. Phys. Chem. B* **2001**, *105*, 12399.
- (10) Hoyau, S.; Ohanessian, G. *Compt. Rend. Acad. Sci. Sér. IIc* **1998**, *1*, 795.
- (11) Kapota, C.; Ohanessian, G. Unpublished results. See also: Hoyau, S.; Pélavier, J.-P.; Rogalewicz, F.; Hoppilliard, Y.; Ohanessian, G. *Eur. J. Mass Spectrom.* **2001**, *7*, 303.
- (12) (a) Ryzhov, V.; Dunbar, R. C.; Cerda, B. A.; Wesdemiotis, C. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 1037. (b) Cerda, B. A.; Wesdemiotis, C. *Analyst* **2000**, *125*, 657. (c) Cerda, B. A.; Tally, J. M.; Wesdemiotis, C.; Ohanessian, G. *Chem. Eur. J.* **2002**, *8*, 1377. (d) Kish, M. M.; Ohanessian, G.; Wesdemiotis, C. *Int. J. Mass Spectrom.* **2003**, *227*, 509.
- (13) (a) Jockusch, R. A.; Lemoff, A. S.; Williams, E. R. *J. Am. Chem. Soc.* **2001**, *123*, 12255. (b) Jockusch, R. A.; Lemoff, A. S.; Williams, E. R. *J. Phys. Chem. A* **2001**, *105*, 10929.
- (14) Wytenbach, T.; Witt, M.; Bowers, M. T. *J. Am. Chem. Soc.* **2000**, *122*, 3458.
- (15) Moision, R. M.; Armentrout, P. B. *J. Phys. Chem. A* **2002**, *106*, 10350.

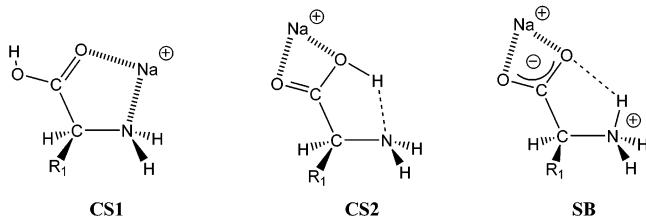
<sup>†</sup> Laboratoire des Mécanismes Réactionnels, Ecole Polytechnique.

<sup>‡</sup> Laboratoire de Chimie Physique, Université de Paris-Sud.

<sup>§</sup> Laboratoire pour l'Utilisation du Rayonnement Electromagnétique, Université de Paris-Sud.

- (1) Voet, D.; Voet, J. G. *Biochemistry*, 2nd ed.; John Wiley and Sons: New York, 1995.
- (2) (a) Ding, Y.; Krogh-Jespersen, K. *Chem. Phys. Lett.* **1992**, *199*, 261. (b) Yu, D.; Armstrong, D. A.; Rauk, A. *Can. J. Chem.* **1992**, *70*, 1762. (c) Nguyen, D. T.; Scheiner, A. C.; Andzelm, J. W.; Sirois, S.; Salahub, D. R.; Hagler, A. T. *J. Comput. Chem.* **1997**, *18*, 1609.
- (3) (a) Price, W. D.; Jockusch, R. A.; Williams, E. R. *J. Am. Chem. Soc.* **1997**, *119*, 11988. (b) Chapo, C. J.; Paul, J. B.; Provencal, R. A.; Roth, K.; Saykally, R. J. *J. Am. Chem. Soc.* **1998**, *120*, 12956. (c) Rak, J.; Skursky, P.; Simons, J.; Gutowski, M. *J. Am. Chem. Soc.* **2001**, *123*, 11695.

Scheme 1



interaction with one, two, or more electron-rich site(s) of the AA in its neutral form, and (ii) salt bridges (SB), in which the metal charge interacts directly with the carboxylate group of the AA in its zwitterionic form (Scheme 1). The specificities of metal–AA interactions will determine whether CS or SB is most favorable in each case.

Several factors may influence the relative energies of CS and SB isomers:

(a) in a rough picture, the neutral  $\rightarrow$  zwitterion transformation in the free AA may be viewed as a proton transfer from the carboxyl to a basic group (either the  $\alpha$ -amino or a more basic side chain group if any, depending on the AA). The more basic the latter, the more favored the zwitterion will be. This reasoning was used to predict that Arg, the most basic of all 20 common AAs of living systems, might be more stable as a zwitterion.<sup>3a</sup> Further studies showed that this is not the case.<sup>3b</sup> However the lowest energy lying neutral and zwitterionic conformers were found to be close in energy.<sup>3c</sup> In such a case, attachment of a cation might well lead to a SB isomer that is more stable than the CS, and there exists indirect but compelling evidence that this is the case for Arg attached to heavy alkali metal ions.<sup>5</sup>

(b) The structural details of the best SB conformer may also be influential. For an optimum electrostatic stabilization, the protonated group should be able to approach the carboxylate in such a way as to make a quasi-linear array of charges with the metal ion. For long side chains with strongly basic groups, such as lysine and arginine, folding of the chain will allow the protonated end to approach the carboxylate but the arrangement of charges relative to the metal ion is not optimum. Computational studies on complexes of glycine with alkali cations reveal that protonation of the  $\alpha$ -amino group is nearly ideal for such a purpose.<sup>6</sup> Therefore, the best candidate for a salt bridge might be the AA bearing the most basic  $\alpha$ -amino group, that is, proline. It is indeed the case that Pro– $\text{Na}^+$  has been predicted, on the basis of *ab initio* calculations, to be most stable as a salt bridge, with a significant energy gap of ca. 6 kcal/mol with the most stable charge solvation isomer.<sup>7</sup> These results were recently confirmed by further, more detailed computations.<sup>8</sup>

(c) The electronic properties of the metal ion are also important. The above results for Pro– $\text{Na}^+$  have been extended to the closely related Pro– $\text{Li}^+$  and Pro– $\text{K}^+$  complexes, and again the SB was found to be most favorable.<sup>8</sup> Other computational studies have provided similar results for Pro– $\text{Ag}^+$ ,<sup>9</sup> however, Pro– $\text{Cu}^+$ <sup>10</sup> and Pro– $\text{Au}^+$ <sup>11</sup> are predicted to be different, with a CS isomer being the most stable.

There have been a number of experimental studies of cationized AAs using a variety of mass spectrometric techniques. Most have aimed at determining the metal ion–AA binding enthalpy and/or free energy. Successful results were obtained using the kinetic method,<sup>12</sup> dissociation kinetics induced by blackbody radiation,<sup>13</sup> ion mobility,<sup>14</sup> guided ion beam,<sup>15</sup> and

metal ion exchange equilibrium<sup>16</sup> experiments. Assignment of the data to specific structures mostly relied on quantum chemical calculations, since none of the above methods can provide direct structural information. Several attempts were made to obtain indirect structural correlations, with some notable success.

(a) The kinetic method was used to compare the metal ion affinity of AAs and their methyl ester.<sup>12b,c</sup> Due to the methyl inductive effect, the ester has a larger metal ion affinity than the corresponding acid if both have similar structures. This is often the case if the acid leads to a CS isomer. However, when it is an SB, the analogous structure cannot be formed with the ester and the affinity order is reversed. This provides valuable clues as to the CS or SB structure of the metal–AA complex; however, it is rather indirect, and it may fail if the acid and the ester bear two different CS structures.

(b) Blackbody infrared radiation dissociation experiments were used to study the dissociation kinetics of metal–AA complexes into the bare metal ion and AA and to compare it to related cases which are known to have either CS or SB structures. Dissociation kinetics were shown to fall into two categories, which could be assigned to CS and SB types of isomers.<sup>13</sup> This indirect approach requires that structurally similar species with known structures exist, to which the species of interest must be compared.

However large the body of information already gathered, it is clear that a direct experimental characterization of the structure of cationized AA complexes is still lacking. It is the purpose of this paper to propose a new approach toward this goal, infrared spectroscopy of gaseous ions over a selected but wide spectral range. This technique has been made possible by the coupling of a radio frequency trap associated to a time-of-flight mass spectrometer and a free electron laser (FEL).<sup>17</sup> The FEL has two advantages over commercial tunable IR lasers: it has a much wider spectral range, and it provides a high peak power. Using this technique has provided a wealth of structural information on a broad variety of ion types, including ionized polycyclic aromatic hydrocarbons (PAHs) and metal clusters.<sup>18</sup> The technique has been made much more flexible and powerful recently by the coupling of a Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer to a FEL.<sup>19</sup> It provides the added value of trapping the ions for long times, which makes cooling possible and allows the gas-phase synthesis of new ions via gas-phase ion–molecule chemistry. This experimental setup has already been used to record the infrared spectra of organometallic complexes such as alkene– $\text{Fe}^+$ <sup>19a</sup> and to characterize the structure of reactive intermediates involved in the activation of methane by  $\text{Ta}^+$ <sup>20</sup> and also of organic ions

- (16) (a) Gapeev, A.; Dunbar, R. C. *J. Am. Chem. Soc.* **2001**, *123*, 8360. (b) Dunbar, R. C. posters presented at the 6th and 7th Gordon Conference on Structures, Energetics and Reaction Dynamics of Gaseous Ions, Ventura, CA, March 2001 and March 2003.
- (17) Oomens, J.; van Roij, A. J. A.; Meijer, G.; von Helden, G. *Astrophys. J.* **2000**, *542*, 404.
- (18) (a) Oomens, J.; Meijer, G.; von Helden, G. *J. Phys. Chem. A* **2001**, *105*, 8302. (b) Oomens, J.; Sartakov, B. G.; Tiemens, A.; Meijer, G.; von Helden, G. *Astrophys. J.* **2001**, *560*, L99. (c) van Heijnsbergen, D.; von Helden, G.; Meijer, G.; Maitre, P.; Duncan, M. *J. Am. Chem. Soc.* **2002**, *124*, 1562.
- (19) (a) Lemaire, J.; Boissel, P.; Heninger, M.; Mauclaire, G.; Bellec, G.; Mestdagh, H.; Simon, A.; LeCaer, S.; Ortega, J. M.; Glotin, F.; Maitre, P. *Phys. Rev. Lett.* **2002**, *89*, 273002–1. (b) Maitre, P.; LeCaer, S.; Simon, A.; Jones, W.; Lemaire, J.; Mestdagh, H.; Heninger, M.; Mauclaire, G.; Boissel, P.; Prazeres, R.; Glotin, F.; Ortega, J. M.; *Nucl. Instrum. Methods A* **2003**, *507*, 541.
- (20) MacAleese, L.; Simon, A.; Lemaire, J.; Boissel, P.; Maitre, P. Manuscript in preparation.

suchas protonated benzene.<sup>21</sup> This experimental setup is used here to probe the structures of gaseous sodium complexes of AAs. As representative cases, we studied Gly- $\text{Na}^+$  and Pro- $\text{Na}^+$ , which are expected to form CS and SB isomers, respectively.

Gas-phase ions are produced in much too small quantities to measure the absorption of IR photons directly. Therefore, the absorption must be probed by a process induced by the absorption. Taking advantage of the capabilities of the mass spectrometer, a readily observable process is ion fragmentation, leading to a decreased intensity of the parent ion signal and the appearance of a fragment ion signal. In the present case, we monitored the fragmentation reaction shown in eq 1 as a function of photon energy:



On the basis of computed AA- $\text{Na}^+$  binding energies, a single IR photon carries too small an energy to induce this fragmentation for any AA: the weakest binder is Gly, for which computations indicate that the binding enthalpy is ca. 37 kcal/mol. Therefore, we probe an InfraRed MultiPhoton Dissociation (IRMPD) process. It has been shown that IR-FEL are particularly well suited for this particular IR spectroscopy, since its fluence is so high that an ion can absorb several photons until it acquires enough internal energy to dissociate, via reaction 1 in the present case.

## II. Experimental and Computational Section

**1. Sample Preparation.** Gaseous AA- $\text{Na}^+$  ions were produced just outside the ICR cell by matrix-assisted laser desorption ionization (MALDI) with a pulsed Nd:YAG laser at tripled frequency, delivering 355 nm photons. The best AA- $\text{Na}^+$  signal was obtained with a layered sample prepared following the sandwich technique.<sup>22</sup> The first layer contained 15  $\mu\text{L}$  of the matrix, dissolved at a  $3.75 \times 10^{-2}$  mol  $\text{L}^{-1}$  concentration (see below). The second layer contained 5  $\mu\text{L}$  of an aqueous solution of NaCl at a concentration of  $2.5 \times 10^{-2}$  mol  $\text{L}^{-1}$ . The AA of interest was deposited as a third layer in the same conditions as NaCl, and a final layer of matrix identical to the bottom layer was added on top. Each new layer was let dry before the next was added; however, it is likely that some dissolution occurred. The matrix:NaCl:AA:matrix relative molar fractions in the sample were 4.5:1:2:4.5. Both 2,5-dihydroxy benzoic acid (DHB) in water and 4-hydroxy- $\alpha$ -cyano cinammic acid (HCCA) in a 1:3 v/v water-acetone solution were used as matrices. This choice was based on the relative sodium affinities of matrix molecules<sup>23</sup> and the AA of interest.<sup>12d</sup> The best spectra were obtained with HCCA for Pro- $\text{Na}^+$  and with DHB for Gly- $\text{Na}^+$ .

**2. Infrared Free Electron Laser Parameters.** The FEL CLIO<sup>26</sup> is based on a 10–50 MeV electron linear accelerator followed by an undulator placed in an optical cavity. The broadly tunable output (3–120  $\mu\text{m}$ ) and the high peak power makes it well suited for IR multiphoton dissociation studies. In our case the long-wavelength cutoff

was ca. 19  $\mu\text{m}$  due to the absorption of the ZnSe entrance window. The FEL output consists of 8  $\mu\text{s}$  long macropulses fired at a repetition rate of 25 Hz. Each macropulse contains 500 micropulses, each a few picoseconds long. The mean IR power was about 500 mW, corresponding to micropulse and macropulse energies of 40  $\mu\text{J}$  and 20 mJ, respectively, and to a peak power of ca. 20 MW.

For a given electron energy, continuous tunability over a large spectral range ( $\Delta\lambda/\lambda = 2.5$ ) is obtained by scanning the undulator gap. The FEL bandwidth can be adjusted by tuning the length of the optical cavity. The laser wavelength profile was monitored while recording the spectra with a monochromator associated with a spiricon. In our case, the relative laser bandwidth was less than 0.5% of the central wavelength.

**3. Operating Conditions.** The experimental setup has been described in detail previously.<sup>19b</sup> It features an easily transportable FT-ICR-MS, MICRA,<sup>28</sup> which owes its compactness to the use of a 1.24 T permanent magnet made of an assemblage of Halbach cylinders.<sup>29</sup> The ICR cell is cubic like ( $2 \times 2 \times 2$  cm), where each excitation electrode has been replaced by a set of four interconnected electrodes to provide optical access to the IR FEL.

The MALDI sample was deposited on a metallic holder (Cu/Ni alloy), which was mounted just outside the ICR cell, 6 mm away from the middle of the nearest trapping plate. The ions formed by the MALDI process enter the ICR cell, collinearly with the magnetic field, through a 5 mm diameter hole. A 3.5 V potential was applied to the trapping plates, and the target holder was maintained to a static potential, tuned to 1.8 V, so as to optimize the AA- $\text{Na}^+$  ion signal. Furthermore, the voltage of the entrance trapping plate was pulsed down to 0 V for 25  $\mu\text{s}$  just after the laser pulse in order to favor ion entrance into the cell and then back to 3.5 V for trapping.

The infrared laser beam, perpendicular to the magnetic field, is focused in the middle of the ICR cell with a 1 m focal length spherical mirror. The laser beam waist at the focal point is probably smaller than the ion cloud dimensions. Nevertheless, provided that an irradiation time of 1 s or more is used, photofragmentation ratios as high as 80% indicate that all ions eventually interact with the IR beam. The irradiation time is set by a fast electromechanical shutter, synchronized with the FEL, that allows for the selection of a given number of macropulses.

In the present study, AA- $\text{Na}^+$  ions of interest were mass selected 600–800 ms after their formation, then irradiated with the IR beam for 3 s (Pro- $\text{Na}^+$ ) or 1 s (Gly- $\text{Na}^+$ ), and the resulting ion signal was detected after the end of IR irradiation. Such a sequence followed by a quench was repeated 25 times, and the mass spectrum was the Fourier transform of the accumulated ion signal. In the case of Pro- $\text{Na}^+$ , the IR laser wavenumber was increased by 10  $\text{cm}^{-1}$  between two consecutive points. Due to difficulties to produce Gly- $\text{Na}^+$  ion, the corresponding IRMPD spectrum is the result of a multiple scans, with a smaller wavenumber spacing (10  $\text{cm}^{-1}$ ) near the IRMPD bands.

The IRMPD spectra presented in this paper are not corrected for variation of the laser fluence (F). The dependence of the IRMPD dissociation rate with the laser fluence is not a simple function. Approximation by a power law  $F^n$  has been used with values of  $n$  ranging between 1<sup>19a</sup> and 2.<sup>18b</sup> However,  $n$  may depend on the range of fluence, the nature of the ion, and probably also upon the IR-active mode which is resonantly excited. Therefore, no corrections were applied here, but one should keep in mind that the IRMPD process is probably favored by a higher fluence on the long wavelength side, considering that the monitored average FEL power was remarkably stable during these experiments.

**4. Computational Details.** The potential energy surfaces of Gly- $\text{Na}^+$  and Pro- $\text{Na}^+$  were computed at the MP2/6-31G\* and

- (21) Jones, W.; Boissel, P.; Chiavarino, B.; Crestoni, M. E.; Fornarini, S.; Lemaire, J.; Maître, P. *Angew. Chem., Int. Ed. Engl.* **2003**, *42*, 2057.
- (22) Kussmann, M.; Nordhoff, E.; Rahbek-Nielsen, H.; Haebel, S.; Rossel-Larsen, M.; Jakobsen, L.; Gobom, J.; Mirgorodskaya, E.; Kroll-Kristensen, A.; Palm, L.; Roepstorff, P. *J. Mass Spectrom.* **1997**, *32*, 593.
- (23) (a) Ohanessian, G. *Int. J. Mass Spectrom.* **2002**, *219*, 577. (b) Zhang, J.; Knochenmuss, R.; Stevenson, E.; Zenobi, R. *Int. J. Mass Spectrom.* **2002**, *213*, 237. (c) Zhang, J.; Ha, T.-K.; Knochenmuss, R.; Zenobi, R. *J. Phys. Chem. A* **2002**, *106*, 6610.
- (24) Gaussian98, revision A.6; Frisch, M. J., *et al.*; Gaussian, Inc.: Pittsburgh, PA, 1998.
- (25) Black, J.; Yablonovitch, E.; Bloembergen, N.; Mukamel, S. *Phys. Rev. Lett.* **1977**, *38*, 1131.
- (26) Prazeres, R.; Glotin, F.; Insa, C.; Jaroszynski, D.; Ortega, J. M. *Eur. Phys. J. D* **1998**, *3*, 87.

- (27) Kerstel, E. R. Th.; Lehmann, K. K.; Mentel, T. F.; Pate, B. H.; Scoles, G. *J. Phys. Chem.* **1991**, *95*, 8282.
- (28) Mauclaire, G.; Lemaire, J.; Boissel, P.; Bellec, G.; Héninger, M. *Eur. J. Mass Spectrom.* **2003**, in press.
- (29) Coey, J. M. D. *J. Magn. Magn. Mater.* **2002**, *248*, 441.

B3LYP/6-31G\* levels. Optimized geometries and relative energies obtained at identical or similar levels for the low-lying energy minima have been reported previously for Gly- $\text{Na}^{+6}$  and Pro- $\text{Na}^{+7,8}$  and are not repeated here. The vibrational spectra obtained at the same levels were used to interpret the recorded spectra. All calculations used the Gaussian98 package.<sup>24</sup> The calculated band intensities were convoluted assuming a Lorentzian profile with a 50  $\text{cm}^{-1}$  full-width at half-maximum.

Hybrid DFT methods such as B3LYP have been shown to outperform local or gradient-corrected DFT, as well as MP2, in predicting the infrared intensities,<sup>30</sup> and an average absolute difference of 20–25  $\text{cm}^{-1}$  with the highest level *ab initio* method can be expected using B3LYP. Calculated frequencies are in general blue-shifted compared to the experimental ones. Provided the use of an appropriate scaling, hybrid DFT methods have been shown to be very efficient in producing the observed fundamental frequencies, and an average absolute difference of 33  $\text{cm}^{-1}$  can be expected using B3LYP.<sup>31</sup> Sophisticated schemes for scaling the force constants rather than the frequencies have been developed, but their use is delicate since they strongly rely on the choice of internal coordinates. To account for the larger anharmonicity of the high-frequency modes, Schlegel and co-workers proposed a “dual-scaling”: one scaling factor for low-energy modes ( $<1800 \text{ cm}^{-1}$ ) and the other for the high-energy ones, essentially involving H stretches. This dual-scaling approach suggests that the low-energy modes are intrinsically better described than the high-energy ones: at the B3LYP level and provided the use of a large basis set, the scaling of the frequencies lower (higher) than 1800  $\text{cm}^{-1}$  is 0.9927 (0.9659) while the uniform scaling factor is 0.9726. Similar conclusions were drawn at other levels of theory including MP2. Considering, in addition, that uniform scaling factors of 0.943 (0.9614) have been proposed for MP2 (B3LYP) using the 6-31G\* basis set,<sup>32</sup> scaling factors of 0.96 and 0.98 were applied in the present work (1000–2000  $\text{cm}^{-1}$ ) to MP2/6-31G\* and B3LYP/6-31G\* frequencies, respectively. After applying these scalings, both sets of frequencies and intensities were very similar. It should be noted that the same scaling factor (0.98) was derived from the comparison of experimental and B3LYP/6-31G\* frequencies in the 800–1800  $\text{cm}^{-1}$  range of the IRMPD spectra of Fe(I) complexes.<sup>19</sup>

**5. Other Accessible Spectral Ranges.** The 1000–2000  $\text{cm}^{-1}$  range has been chosen because it is in this region that calculations predict the most intense and well-separated bands. Indeed, in the 100–1000  $\text{cm}^{-1}$  region there are numerous and less intense bands, making it difficult to interpret the spectrum. The 2000–3300  $\text{cm}^{-1}$  region corresponds to the OH and NH stretching modes, so it seems structurally specific and therefore worth probing. However, there is always coincidence of several bands, for example, in the case of Gly- $\text{Na}^+$  the N–H stretching band of the SB isomer is in resonance with the O–H stretching band of CS2. Thus, the study of these regions would probably not bring any supplementary information.

**6. IR Spectra of Deuterated Species.** The spectra of some deuterated isotopomers of Gly- $\text{Na}^+$  and Pro- $\text{Na}^+$  have been simulated. We have only considered amino acids deuterated at nonexchangeable positions ( $\text{CH}_2$ ), since it would be impossible to produce gaseous ions deuterated at exchangeable positions without the use of deuterated solvents and matrices. Thus, only *d*<sub>2</sub>-Gly- $\text{Na}^+$  and *d*<sub>7</sub>-Pro- $\text{Na}^+$  spectra were computed. The intensities of  $\text{CD}_2$  (or  $\text{CH}_2$ ) bending and stretching modes are weak, because of the small electronegativity difference between carbon and hydrogen. In conclusion, it appears that the study of deuterated species cannot be a complement to the assignment of spectra in these cases.

### III. Results

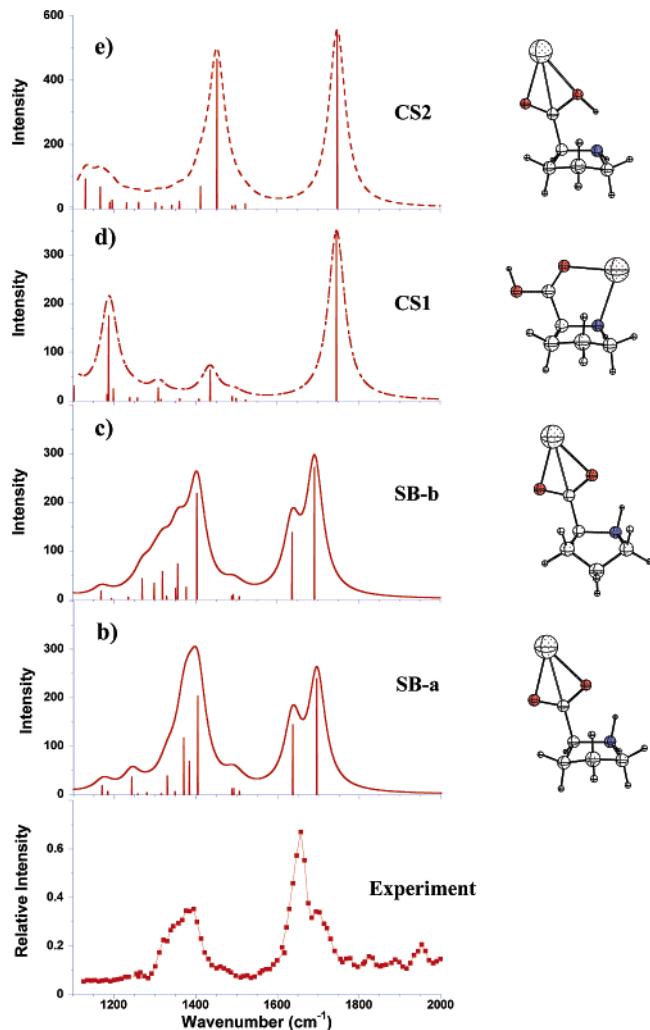
#### 1. Experimental Results.

The absorption of multiple IR photons by the AA- $\text{Na}^+$  complexes is probed by their dis-

(30) Halls, M. D.; Schlegel, H. B. *J. Chem. Phys.* **1998**, *109*, 10587.

(31) Halls, M. D.; Velkovski, J.; Schlegel, H. B. *Theor. Chem. Acc.* **2001**, *105*, 413.

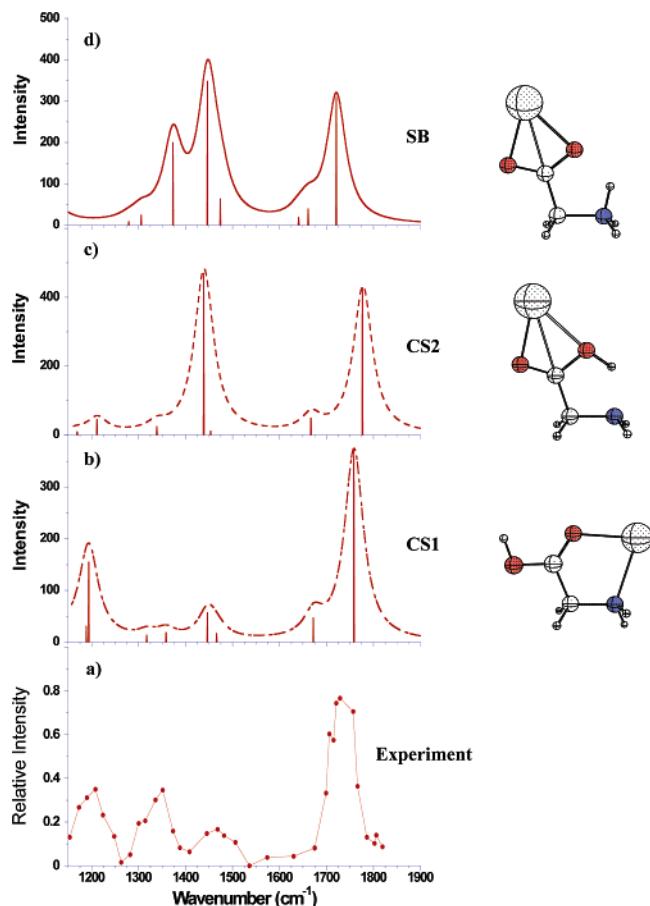
(32) Scott, A. P.; Radom, L. *J. Phys. Chem.* **1996**, *100*, 16502.



**Figure 1.** IRMPD spectrum of Pro- $\text{Na}^+$  (a). IR spectra and structures of the lowest energy Pro- $\text{Na}^+$  isomers, computed at the B3LYP/6-31G\* level (b–e).

pearance as a function of photon energy, due to the fragmentation reaction to form AA and  $\text{Na}^+$  (eq 1). Monitoring the intensity of  $\text{Na}^+$  relative to the total ion signal due to  $\text{Na}^+$  and Pro- $\text{Na}^+$  leads to the IR spectrum shown in Figure 1a. Two main features are apparent: a broad signal between 1300 and 1430  $\text{cm}^{-1}$  and another intense signal, with two partly resolved bands, between 1600 and 1750  $\text{cm}^{-1}$ . Monitoring eq 1 for the Gly- $\text{Na}^+$  complex yielded an IR spectrum of lower quality than for Pro- $\text{Na}^+$  (Figure 2a). It may be seen that an intense feature exists between 1670 and 1800  $\text{cm}^{-1}$ . A series of points showing one or several absorptions between 1150 and 1400  $\text{cm}^{-1}$  is also apparent in the experimental spectrum. There is also a small feature in the 1400–1500  $\text{cm}^{-1}$  range.

**2. Low-Energy Isomers of Gly- $\text{Na}^+$  and Pro- $\text{Na}^+$ .** The low-energy isomers of both Gly- $\text{Na}^+$  and Pro- $\text{Na}^+$  have been studied previously, together with the transition states for some of their interconversions.<sup>6–8</sup> In both cases, the most favorable modes of attachment of the sodium ion to the AA are the same, although their energetic order is different. For Gly- $\text{Na}^+$ , all computational levels indicate that the most stable isomer is CS1, in which the ion chelates Gly between nitrogen and the carbonyl oxygen. The SB and CS2 isomers lie 2–2.5 kcal/mol higher in energy, and other isomers are less stable. Pro- $\text{Na}^+$  differs from



**Figure 2.** IRMPD spectrum of Gly- $\text{Na}^+$  (a). IR spectra and structures of the lowest energy Pro- $\text{Na}^+$  isomers, computed at the B3LYP/6-31G\* level (b-d).

Gly- $\text{Na}^+$  in several significant ways. First, the SB isomer is the most stable, followed by CS1, which lies ca. 5 kcal/mol higher in energy. Second, as shown by Russo *et al.*,<sup>8</sup> each isomer exists as a pair of conformers differing by the puckering orientation of the five-membered ring. For each isomer, the energy difference between the two conformers is less than 1 kcal/mol, and they are connected by transition states with energy barriers of less than 3 kcal/mol. Third, the CS2 isomer exists as a very shallow minimum. It was impossible to locate CS2 on the MP2/6-31G\* potential surface. Russo *et al.*<sup>8</sup> found it at the B3LYP/6-31G\*+G\*\* level; however, attempts to localize the transition state for proton transfer leading to SB were unsuccessful. We were able to locate both CS2 and this TS at the B3LYP/6-31G\* level, and the TS was found to lie 0.3 kcal/mol higher than CS2 at the B3LYP/6-31G\*+G(2d,2p)//B3LYP/6-31G\* level. However, when thermal corrections at 298 K are included, it is found to lie 1.2 kcal/mol *below* CS2. This clearly indicates that CS2 lies on a rather flat portion of the potential surface and corresponds to a very shallow minimum at best. As for Gly- $\text{Na}^+$ , other minima exist at higher energies, typically more than 10 kcal/mol higher than SB.

**3. Assignment of the Computed IR Absorption Spectra of Pro- $\text{Na}^+$  Isomers.** The simulated IR spectra and structures of the low-lying isomers of Pro- $\text{Na}^+$  are presented in Figure 1. Those for the two conformers of SB are shown in Figure 1b and 1c; those for CS1 and CS2 are shown in Figure 1d and 1e, respectively. The spectra of the second conformers of CS1 and CS2, corresponding to puckering inversion of the five-membered

**Table 1:** Measured (left column) and Calculated Vibrational Spectra for Pro- $\text{Na}^+$ <sup>a</sup>

Wavenumber (cm <sup>-1</sup> )	Calculated Intensity (km/mol)				Vibrational Mode
	Experimental	Calculated	CS1	CS2	
1131			92		vC-OH
1167			67		$\tau$ C-OH
1187			174		vC-OH
1306		1319		57	
		1357		73	
		1370		117	Ring deformation <sup>a</sup>
		1384		68	
	1429				
		1403		218	$\nu^s\text{CO}_2$
		1404		203	
		1412		68	$\tau\text{NH}$
		1436	63		$\delta\text{COH}$
		1452	465		
1656		1637		137	$\delta\text{NH}_2$
		1637		144	
1698		1692		271	$\nu^s\text{CO}_2$
		1696		239	
		1745	349		$\nu\text{C=O}$
		1747	550		

<sup>a</sup> Mode mainly localized on the five-membered ring, i.e., combination of  $\tau\text{NH}_2$ ,  $\tau\text{CH}_2$ ,  $\omega\text{CH}_2$ , and  $\tau\text{CH}$ .

ring, are omitted for clarity. As mentioned in the Experimental and Computational Section, these IR absorption spectra are derived from a convolution of the calculated band intensities. The positions and intensities of the main bands are given in Table 1. The spectra of both CS isomers have two common bands. The strongest around 1740 cm<sup>-1</sup> corresponds to the C=O stretch, while the second corresponds to the COH bend near 1450 cm<sup>-1</sup>. The latter is much more intense for CS2 than for CS1. The CS1 spectrum presents another strong band at 1197 cm<sup>-1</sup> due to the C-OH stretch. In the case of CS2, that absorption is much less intense. In comparison with the CS spectra, those for the conformers of SB are very different. In the 1600–1750 cm<sup>-1</sup> range, two bands are predicted for SB instead of one for CS1 and CS2. A normal-mode analysis reveals that the two bands of SB display components on both the NH<sub>2</sub> scissoring ( $\delta\text{NH}_2$ ) and the antisymmetric combination of C=O stretches ( $\nu^s\text{CO}_2$ ), the less intense at 1637 cm<sup>-1</sup> presenting a larger component on  $\delta\text{NH}_2$ , the most intense at 1696 cm<sup>-1</sup> corresponding to  $\nu^s\text{CO}_2$  mode. A group of peaks exists between 1300 and 1450 cm<sup>-1</sup>, the most intense of which is associated with the symmetric combination of C=O stretches with the others being combinations of ring CH<sub>2</sub>, CH, and NH<sub>2</sub> bendings.

**4. Assignment of the Computed IR Absorption Spectra of Gly- $\text{Na}^+$  Isomers.** The simulated IR spectra and structures of the low-lying CS1, CS2, and SB isomers of Gly- $\text{Na}^+$  are displayed in Figure 2b, c, and d, respectively, and the positions of the most intense calculated IR bands are given in Table 2, along with their corresponding intensities.

As for Pro- $\text{Na}^+$ , CS1 and CS2 share some common bands: the C-OH stretch near 1200 cm<sup>-1</sup>, COH bend near 1440 cm<sup>-1</sup>, NH<sub>2</sub> scissoring near 1670 cm<sup>-1</sup>, and C=O stretch near 1760 cm<sup>-1</sup>. The CS1 and CS2 spectra remain easily distinguishable because of the very different relative band intensities, especially

**Table 2:** Measured (left column) and Calculated Vibrational Spectra for Gly- $\text{Na}^+$

wavenumber ( $\text{cm}^{-1}$ )		calculated intensity (Km/mol)			vibrational mode
experimental	calculated	CS1	CS2	SB	
1204	1189	30			$\tau\text{NH}_2 + \tau\text{CH}_2$
	1195	154			$\nu\text{C-OH}$
	1211		44		
	1306			24	$\omega\text{NH}_2 + \omega\text{CH}_2$
	1318	13			$\omega\text{NH}_2 + \omega\text{CH}_2$
	1339		23		$\omega\text{CH}_2$
1351	1359	18			$\tau\text{NH}_2 + \tau\text{CH}_2$
	1373			199	$\nu^s\text{CO}_2$
	1439		469		$\delta\text{C-OH}$
	1446	56			
1466	1446			348	$\text{NH}_3$ umbrella
	1474			63	$\delta\text{CH}_2$
	1662			39	$\text{NH}_3$ bending
	1667		48		$\delta\text{NH}_2$
	1672	46			
1727	1721			309	$\nu^{\text{as}}\text{CO}_2$
	1759	370			$\nu\text{C=O}$
	1778		426		

in the  $1300\text{--}1400\text{ cm}^{-1}$  region. The spectrum of isomer SB presents three characteristic bands, one at  $1720\text{ cm}^{-1}$  corresponding to the antisymmetric combination of C–O stretches and two overlapping bands at  $1446$  and  $1373\text{ cm}^{-1}$  corresponding to the  $\text{NH}_3$  umbrella mode and the symmetric combination of C–O stretches, respectively. Contrary to the case of Pro- $\text{Na}^+$ , distinguishing between SB and CS cannot be based on the  $1650\text{--}1800\text{ cm}^{-1}$  zone since there are now two bands for all isomers, corresponding to  $\text{NH}_2$  scissoring and combination of CO stretches.

#### IV. Discussion

A satisfactory agreement between the IRMPD and IR linear absorption spectra has been observed for a large variety of ions.<sup>17–19,21</sup> Thus, despite the multiple photon character of the IRMPD process, it is generally believed that the relative intensities of the bands in the IRMPD spectrum are governed by the absorption of the few first photons. This early stage of the IRMPD which occurs within the discrete regime<sup>25</sup> is critical, and no fragmentation occurs unless the first photon is resonant with a vibrational transition of the ion. Subsequent excitation occurs within the quasi-continuum regime in which the density of states is so high that the vibrational energy is rapidly randomized among all the modes of the molecule. Nevertheless, deviations from the linear absorption spectrum may be observed depending on the transition between these two regimes. Such deviation is expected for small ions for which the density of vibrational states is too low to allow for a fast energy redistribution. Vibrational couplings are also important, and it has been shown, for instance, that heavy atoms<sup>27</sup> may severely increase the lifetime of excited vibrational states, thereby decreasing the IRMPD efficiency. Furthermore, as cross-anharmonicities increase as a function of the internal energy, each vibrational transition is expected to be more and more red-shifted and broadened. As a result, the position of the peak maximum on the IRMPD spectrum is generally slightly red-shifted with respect to the corresponding maximum of the linear absorption spectrum.

**1. Assignment of the Experimental IR Absorption Spectrum of Pro- $\text{Na}^+$ .** The assignment of the IRMPD spectrum of Pro- $\text{Na}^+$  is based on its comparison to the spectra computed

for the various isomers described above. To make the comparison easier, the absorption cross-section is represented in Figure 1, assuming a Lorentzian profile for each calculated infrared band. As long as one considers only the relative positions of the two IRMPD broad bands, they are in better agreement with the two SB conformers (Figure 1b and c) than with the two CS isomers. Thus, a tentative assignment of the observed bands is proposed in Table 1 based on the calculated IR spectra of the SB isomers. As recalled above, the multiple photon character of the IRMPD process is expected to lead to some deviations from the relative intensities of the linear absorption spectrum (see discussion below), and this might be the case in the  $1600\text{--}1750\text{ cm}^{-1}$  region.

If one considers the IRMPD experimental feature between  $1100$  and  $1500\text{ cm}^{-1}$ , it does not show the principal signature of the CS1 isomer. Indeed, the intense C–OH stretch band at  $1197\text{ cm}^{-1}$  of isomer CS1 (Figure 1d) is not observed experimentally. Isomer CS2 presents a strong absorption at  $1450\text{ cm}^{-1}$ ; however, there is only one mode with significant intensity in this region, which would lead to a much narrower feature than that experimentally observed. Furthermore, it has been found that CS2 is at best a very shallow minimum on the potential surface; therefore, it is unlikely to be observable in significant intensity. On the other hand, the computed spectra for the two conformers of SB are in fair agreement with the broad experimental feature. The broad character of the experimental signal between  $1300$  and  $1450\text{ cm}^{-1}$  could be the result of the excitation of Pro- $\text{Na}^+$  through the various ring deformations of SB structures, and the maximum of the IRMPD rate could be attributed to the contribution of the  $\nu^s\text{CO}_2$  at  $1404\text{ cm}^{-1}$ , which are predicted to be intense for both SB isomers (see Table 1).

The comparison of the position of the experimental and calculated bands between  $1600$  and  $1750\text{ cm}^{-1}$  further supports the preference for the two SB conformers. Indeed, in these two structures,  $\text{NH}_2$  scissoring and the antisymmetric combination of  $\text{C=O}$  stretches lead to two absorption bands whose calculated positions are in very good agreement with experiment. The broad experimental band presents two maxima at  $1656$  and  $1698\text{ cm}^{-1}$ , very close to the calculated position of the  $\text{NH}_2$  scissoring ( $1637$ ) and  $\text{C=O}$  stretches ( $1692\text{--}1696$ ), respectively. On the other hand, both CS isomers only have one IR-active mode in this region (the  $\text{C=O}$  stretch), which should lead to a single maximum in the IRMPD spectrum. Furthermore, the computed position of this mode in CS1 and CS2 is at  $1747\text{ cm}^{-1}$ , a frequency at which the experimental signal tends to vanish.

On the basis of this analysis of the relative positions of the experimental and calculated bands, it seems that the two SB conformers, which are computed to be the lowest energy structures for Pro- $\text{Na}^+$ , are the best candidates for the interpretation of the experimental spectrum. On the other hand, the difference in relative intensities between the experimental and theoretical spectra (parts a, b, and c in Figure 1, respectively) in the  $1600$  and  $1750\text{ cm}^{-1}$  region might seem large. Considering the proximity of the two normal modes calculated at  $1637$  and  $1696\text{ cm}^{-1}$ , in the case of SB-a, for example, and the strong coupling between the  $\text{NH}_2$  scissoring and the antisymmetric  $\text{CO}_2$  stretch ( $\nu^{\text{as}}\text{CO}_2$ ), it is conceivable that the relative contributions of these two local modes to the normal modes are incorrectly described at the present computational level.

It is also conceivable that this apparent discrepancy is due to the multiple photon nature of the IRMPD process. Indeed, as the internal energy increases, the two vibrational transitions are red-shifted to an amount which in this small ion could be of the order of magnitude of the difference between the two frequencies of interest ( $1637$  and  $1696\text{ cm}^{-1}$ ). As a result, the IRMPD band corresponding to the high-energy mode would be simply red-shifted as compared to the calculated value. On the contrary, a more subtle absorption process could be expected in the case of the low-energy mode. Indeed, as the internal energy increases, the red-shift effect could bring the higher energy mode into resonance with the FEL, thereby enhancing the global multiphoton absorption. As a result, this multiple photon effect will increase the intensity of the lowest energy mode with respect to the highest energy one, as compared to the relative intensities of the linear absorption spectrum calculated assuming that the ion has no internal energy. One could also expect a blue shift of the lowest energy mode and therefore two IRMPD bands closer to each other than the two calculated positions. Such an effect can be observed (see Table 1), but its extent is too small to be meaningfully discussed.

Finally, in light of the above discussion on the relative intensities in the  $1600$  and  $1750\text{ cm}^{-1}$  region, it is interesting to consider the possible contribution of the CS1 isomer to the IRMPD spectrum. As said above, DFT calculations predict that this isomer displays only one band in this region, and it is predicted to be higher in energy than the two vibrational transitions of the SB conformers. As a result, any contribution of CS1 isomer to the IRMPD spectrum would enhance the IRMPD intensity on the blue side of the  $1600$ – $1750\text{ cm}^{-1}$  band, which is not observed experimentally.

**2. Assignment of the Experimental IR Absorption Spectrum of  $\text{Gly}-\text{Na}^+$ .** The IR spectrum of  $\text{Gly}-\text{Na}^+$  shown in Figure 2a is of somewhat lower quality than that for  $\text{Pro}-\text{Na}^+$ . Still, comparison with the spectra computed for the low-energy isomers of  $\text{Gly}-\text{Na}^+$  (Figure 2b–d) does provide structural information on the experimentally probed ions. It may be seen that an intense feature exists in the experimental spectrum between  $1680$  and  $1800\text{ cm}^{-1}$  and that all isomers display intense absorption bands in this area, corresponding to  $\text{C}=\text{O}$  stretches. The experimental resolution is too low to permit assignment of this feature to a specific isomer. Absorption between  $1150$  and  $1250\text{ cm}^{-1}$  is also apparent in the experimental spectrum, and the only isomer for which an absorption band exists in this region is CS1 (CS2 only has a weak band, while SB has none). Another clear dichotomy between CS1 on the one hand and CS2 and SB on the other is that absorption intensity in the  $1400$ – $1500\text{ cm}^{-1}$  range is weak in the former, in good agreement with experimental results, while it is stronger in the latter two. On the basis of these observations, an assignment of the experimental bands is proposed in Table 2. The most intense observed band may owe its origin to the intense absorption band of  $\nu\text{C}=\text{O}$  in CS1 at  $1759\text{ cm}^{-1}$ , but the  $\nu^{\text{as}}\text{CO}_2$  mode at  $1721\text{ cm}^{-1}$  of SB cannot be excluded. The

observed band at  $1204\text{ cm}^{-1}$  could be assigned to the  $\nu\text{C}-\text{OH}$  stretch calculated at  $1195\text{ cm}^{-1}$  in CS1. Finally, the experimental band with a maximum at  $1351\text{ cm}^{-1}$  may be attributed to the superposition of  $\tau\text{NH}_2 + \tau\text{CH}_2$  (calculated at  $1359\text{ cm}^{-1}$ ) and of  $\omega\text{NH}_2 + \omega\text{CH}_2$  (at  $1318\text{ cm}^{-1}$ ) of CS1. It is clear overall that the observed bands are only consistent with a dominant population of the CS1 isomer, although a mixture involving SB cannot be ruled out.

## V. Conclusions

The coupling of a free electron laser with an FT-ICR mass spectrometer enables recording the infrared spectra of gaseous ions. Although the ion density in the ICR cell is small, the unusually high brightness of the FEL and the temporal structure of its pulses make possible the absorption of multiple IR photons by a large fraction of trapped ions. The broad range of tunability of the FEL offers the possibility of recording IR spectra, thereby obtaining vibrational fingerprints of gaseous ions selectively prepared in an ICR cell. Photon absorption is probed by unimolecular fragmentation processes which are triggered by the accumulation of many vibrational quanta. In this work, we characterized the structural types of  $\text{Gly}-\text{Na}^+$  and  $\text{Pro}-\text{Na}^+$  for which all other mass spectrometric methods have failed to date to provide direct and unambiguous structural information. The unimolecular process that reveals absorption is the dissociation of each AA– $\text{Na}^+$  complex into its constituents AA and  $\text{Na}^+$ . It is remarkable that for both  $\text{Pro}-\text{Na}^+$  and  $\text{Gly}-\text{Na}^+$  the IRMPD spectrum is in good agreement with the absorption spectrum predicted by DFT calculations, that is, despite the complicated nature of the IRMPD process, the present approach yields an IR spectrum in close agreement with the absorption features. Furthermore, the method appears to provide access to the lowest energy isomer in each case. The positions of the recorded absorption bands show that the experimental spectrum is consistent with the formation of the salt bridge isomer for  $\text{Pro}-\text{Na}^+$ . In contrast, the experimental spectrum of  $\text{Gly}-\text{Na}^+$  is consistent with the charge solvation isomer CS1. In this case, however, some population of another isomer cannot be excluded. These results provide the first direct answer to the long-standing question of charge solvation vs salt bridge structure of cationized amino acids. Extending this work to complexes of metal ions with other amino acids and oligopeptides should be relatively straightforward. Therefore, this technique holds great promise for the structural identification of biological molecules by their vibrational fingerprints.

**Acknowledgment.** We thank Drs. Sophie Bourcier and Martine Caroff for their help in optimizing the MALDI conditions, Luke MacAleese for his help with the MICRA/CLIO setup, and Dr. Jean-Michel Ortega and the CLIO team. We also thank the laser program POLA in Orsay for providing us with an Nd–YAG laser. The reviewers are also acknowledged for their constructive comments.

JA036932V

## Bibliographie

- [1] Bludsky, O.; Chocholousova, J.; Vacek, J.; Huisken, F.; Hobza, P., Anharmonic treatment of the lowest-energy conformers of glycine: A theoretical study. *J. Chem. Phys.* **2000**, 113, (11), 4629-4635.
- [2] Nguyen, D. T.; Scheiner, A. C.; Andzelm, J. W.; Sirois, S.; Salahub, D. R.; Hagler, A. T., A density functional study of the glycine molecule: Comparison with post-Hartree-Fock calculations and experiment. *J. Comput. Chem.* **1997**, 18, (13), 1609-1631.
- [3] Suenram, R. D.; Lovas, F. J., Millimeter Wave Spectrum of Glycine. *J. Mol. Spectrosc.* **1978**, 72, (3), 372-382.
- [4] Suenram, R. D.; Lovas, F. J., Millimeter Wave Spectrum of Glycine - a New Conformer. *J. Am. Chem. Soc.* **1980**, 102, (24), 7180-7184.
- [5] Godfrey, P. D.; Brown, R. D.; Rodgers, F. M., The missing conformers of glycine and alanine: Relaxation in seeded supersonic jets. *J. Mol. Struct.* **1996**, 376, 65-81.
- [6] Stepanian, S. G.; Reva, I. D.; Radchenko, E. D.; Rosado, M. T. S.; Duarte, M.; Fausto, R.; Adamowicz, L., Matrix-isolation infrared and theoretical studies of the glycine conformers. *J. Phys. Chem. A* **1998**, 102, (6), 1041-1054.
- [7] Pecul, M.; Ruud, K.; Rizzo, A.; Helgaker, T., Conformational effects on the optical rotation of alanine and proline. *J. Phys. Chem. A* **2004**, 108, (19), 4269-4276.
- [8] Stepanian, S. G.; Reva, I. D.; Radchenko, E. D.; Adamowicz, L., Conformers of nonionized proline. Matrix-isolation infrared and post-Hartree-Fock ab initio study. *J. Phys. Chem. A* **2001**, 105, (47), 10664-10672.
- [9] Kuntz, A. F.; Boynton, A. W.; David, G. A.; Colyer, K. E.; Poutsma, J. C., The proton affinity of proline analogs using the kinetic method with full entropy analysis. *J. Am. Soc. Mass Spectrom.* **2002**, 13, (1), 72-81.
- [10] Marino, T.; Russo, N.; Toscano, M., Interaction of  $\text{Li}^+$ ,  $\text{Na}^+$ , and  $\text{K}^+$  with the proline amino acid. Complexation modes, potential energy profiles, and metal ion affinities. *J. Phys. Chem. B* **2003**, 107, (11), 2588-2594.
- [11] Price, W. D.; Jockusch, R. A.; Williams, E. R., Is arginine a zwitterion in the gas phase? *J. Am. Chem. Soc.* **1997**, 119, (49), 11988-11989.
- [12] Chapo, C. J.; Paul, J. B.; Provencal, R. A.; Roth, K.; Saykally, R. J., Is arginine zwitterionic or neutral in the gas phase? Results from IR cavity ringdown spectroscopy. *J. Am. Chem. Soc.* **1998**, 120, (49), 12956-12957.
- [13] Rak, J.; Skurski, P.; Simons, J.; Gutowski, M., Low-energy tautomers and conformers of neutral and protonated arginine. *J. Am. Chem. Soc.* **2001**, 123, (47), 11695-11707.

- [14] Gdanitz, R. J.; Cardoen, W.; Windus, T. L.; Simons, J., Very large scale computations of the free energies of eight low-lying structures of arginine in the gas phase. *J. Phys. Chem. A* **2004**, 108, (3), 515-518.
- [15] Julian, R. R.; Jarrold, M. F., Gas-phase zwitterions in the absence of a net charge. *J. Phys. Chem. A* **2004**, 108, (49), 10861-10864.
- [16] Xu, S. J.; Zheng, W. J.; Radisic, D.; Bowen, K. H., The stabilization of arginine's zwitterion by dipole-binding of an excess electron. *J. Chem. Phys.* **2005**, 122, (9).
- [17] Skurski, P.; Rak, J.; Simons, J.; Gutowski, M., Quasidegeneracy of zwitterionic and canonical tautomers of arginine solvated by an excess electron. *J. Am. Chem. Soc.* **2001**, 123, (44), 11073-11074.
- [18] Ding, Y. B.; KroghJespersen, K., The 1:1 glycine zwitterion-water complex: An ab initio electronic structure study. *J. Comput. Chem.* **1996**, 17, (3), 338-349.
- [19] Jensen, J. H.; Gordon, M. S., On the Number of Water-Molecules Necessary to Stabilize the Glycine Zwitterion. *J. Am. Chem. Soc.* **1995**, 117, (31), 8159-8170.
- [20] Kassab, E.; Langlet, J.; Evleth, E.; Akacem, Y., Theoretical study of solvent effect on intramolecular proton transfer of glycine. *Theochem-J. Mol. Struct.* **2000**, 531, 267-282.
- [21] Fernandez-Ramos, A.; Smedarchina, Z.; Siebrand, W.; Zgierski, M. Z., A direct-dynamics study of the zwitterion-to-neutral interconversion of glycine in aqueous solution. *J. Chem. Phys.* **2000**, 113, (21), 9714-9721.
- [22] Wang, W. Z.; Zheng, W. X.; Pu, X. M.; Wong, N. B.; Tian, A. M., The 1 : 1 glycine-water complex: some theoretical observations. *Theochem-J. Mol. Struct.* **2002**, 618, (3), 235-244.
- [23] Wang, W. Z.; Pu, X. M.; Zheng, W. X.; Wong, N. B.; Tian, A. M., Some theoretical observations on the 1 : 1 glycine zwitterion-water complex. *Theochem-J. Mol. Struct.* **2003**, 626, 127-132.
- [24] Chaudhari, A.; Sahu, P. K.; Lee, S. L., Many-body interaction in glycine-(water)<sub>3</sub> complex using density functional theory method. *J. Chem. Phys.* **2004**, 120, (1), 170-174.
- [25] Chaudhari, A.; Lee, S. L., Computational study of glycine-(water)<sub>3</sub> complex by density functional method. *Chem. Phys.* **2005**, 310, (1-3), 281-285.
- [26] Xu, S. J.; Nilles, M.; Bowen, K. H., Zwitterion formation in hydrated amino acid, dipole bound anions: How many water molecules are required? *J. Chem. Phys.* **2003**, 119, (20), 10696-10701.
- [27] Hoyau, S.; Pelicier, J. P.; Rogalewicz, F.; Hoppilliard, Y.; Ohanessian, G., Complexation of glycine by atomic metal cations in the gas phase. *Eur. J. Mass Spectrom.* **2001**, 7, (4-5), 303-311.
- [28] Hoyau, S.; Ohanessian, G., Interaction of alkali metal cations ( $\text{Li}^+$ - $\text{Cs}^+$ ) with glycine in the gas phase: A theoretical study. *Chem.-Eur. J.* **1998**, 4, (8), 1561-1569.

- [29] Pulkkinen, S.; Noguera, M.; Rodriguez-Santiago, L.; Sodupe, M.; Bertran, J., Gas phase intramolecular proton transfer in cationized glycine and chlorine substituted derivatives (M-gly, M =  $\text{Na}^+$ ,  $\text{Mg}^{2+}$ ,  $\text{Cu}^+$ ,  $\text{Ni}^+$ , and  $\text{Cu}^{2+}$ ): Existence of zwitterionic structures? *Chem.-Eur. J.* **2000**, 6, (23), 4393-4399.
- [30] Hoyau, S.; Norrman, K.; McMahon, T. B.; Ohanessian, G., A quantitative basis for a scale of  $\text{Na}^+$  affinities of organic and small biological molecules in the gas phase. *J. Am. Chem. Soc.* **1999**, 121, (38), 8864-8875.
- [31] Wyttenbach, T.; Witt, M.; Bowers, M. T., On the stability of amino acid zwitterions in the gas phase: The influence of derivatization, proton affinity, and alkali ion addition. *J. Am. Chem. Soc.* **2000**, 122, (14), 3458-3464.
- [32] Cox, H. A.; Julian, R. R.; Lee, S. W.; Beauchamp, J. L., Gas-phase H/D exchange of sodiated glycine oligomers with  $\text{ND}_3$ : Exchange kinetics do not reflect parent ion structures. *J. Am. Chem. Soc.* **2004**, 126, (20), 6485-6490.
- [33] Wyttenbach, T.; Bushnell, J. E.; Bowers, M. T., Salt bridge structures in the absence of solvent? The case for the oligoglycines. *J. Am. Chem. Soc.* **1998**, 120, (20), 5098-5103, [http://www.chem.ucsb.edu/~bowersgroup/publications/supporting\\_info/JACS\\_1998\\_120\\_5098.shtml](http://www.chem.ucsb.edu/~bowersgroup/publications/supporting_info/JACS_1998_120_5098.shtml).
- [34] Talley, J. M.; Cerdá, B. A.; Ohanessian, G.; Wesdemiotis, C., Alkali metal ion binding to amino acids versus their methyl esters: Affinity trends and structural changes in the gas phase. *Chem.-Eur. J.* **2002**, 8, (6), 1377-1388.
- [35] Moision, R. M.; Armentrout, P. B., Experimental and theoretical dissection of sodium cation/glycine interactions. *J. Phys. Chem. A* **2002**, 106, (43), 10350-10362.
- [36] Ye, S. J.; Moision, R. M.; Armentrout, P. B., Sequential bond energies of water to sodium glycine cation. *Int. J. Mass Spectrom.* **2005**, 240, (3), 233-248.
- [37] Benzakour, M.; Cartier, A.; McHarfi, M.; Daoudi, A., Interaction of aminoacids with cations. I - the glycine- $\text{Li}^+$ - $n(\text{H}_2\text{O})$  ( $n=0-3$ ) complexes. *Theochem-J. Mol. Struct.* **2004**, 681, (1-3), 99-104.
- [38] Gapeev, A.; Dunbar, R. C.,  $\text{Na}^+$  affinities of gas-phase amino acids by ligand exchange equilibrium. *Int. J. Mass Spectrom.* **2003**, 228, (2-3), 825-839.
- [39] Kish, M. M.; Ohanessian, G.; Wesdemiotis, C., The  $\text{Na}^+$  affinities of alpha-amino acids: side-chain substituent effects. *Int. J. Mass Spectrom.* **2003**, 227, (3), 509-524.
- [40] Lemoff, A. S.; Bush, M. F.; Williams, E. R., Structures of cationized proline analogues: Evidence for the Zwitterionic form. *J. Phys. Chem. A* **2005**, 109, (9), 1903-1910.
- [41] Chiavarino, B.; Crestoni, M. E.; Fornarini, S.; Lemaire, J.; Mac Aleese, L.; Maitre, P., Infrared absorption features of gaseous isopropyl carbocations. *ChemPhysChem* **2004**, 5, (11), 1679-1685.

- [42] Jones, W.; Boissel, P.; Chiavarino, B.; Crestoni, M. E.; Fornarini, S.; Lemaire, J.; Maitre, P., Infrared fingerprint of protonated benzene in the gas phase. *Angew. Chem.-Int. Edit.* **2003**, 42, (18), 2057-2059.
- [43] Simon, A.; Jones, W.; Ortega, J. M.; Boissel, P.; Lemaire, J.; Maitre, P., Infrared multiphoton dissociation spectroscopy of gas-phase mass-selected hydrocarbon-Fe<sup>+</sup> complexes. *J. Am. Chem. Soc.* **2004**, 126, (37), 11666-11674.
- [44] Le Caer, S.; Heninger, M.; Lemaire, J.; Boissel, P.; Maitre, P.; Mestdagh, H., Structural characterization of selectively prepared cationic iron complexes bearing monodentate and bidentate ether ligands using infrared photodissociation spectroscopy. *Chem. Phys. Lett.* **2004**, 385, (3-4), 273-279.
- [45] Kapota, C.; Lemaire, J.; Maitre, P.; Ohanessian, G., Vibrational signature of charge solvation vs salt bridge isomers of sodiated amino acids in the gas phase. *J. Am. Chem. Soc.* **2004**, 126, (6), 1836-1842.

## **Chapitre V :**

**Rôle structurant de  $\text{Na}^+$  sur des  
petits peptides :  
du di- au tétra-peptide.**

## Introduction

Ce chapitre a pour but l'étude de la structuration de petits peptides par le cation sodium. Cette étude est décomposée en deux parties. D'une part, l'étude de la structure des adduits  $\text{Gly}_n\text{-Na}^+$  et  $\text{Ala}_n\text{-Na}^+$   $n < 4$  par spectroscopie IRMPD ainsi que l'étude théorique des structures des di- aux tétra-peptides sodiés. En effet, la technique de spectroscopie IRMPD ayant fait ses preuves pour l'étude des isomères de  $\text{Gly-Na}^+$  et  $\text{Pro-Na}^+$  (décrise dans le chapitre précédent), nous avons souhaité l'étendre à l'étude des conformères de peptides. D'autre part, dans le cadre d'une collaboration avec le Professeur Chrys Wesdemiotis (Akron, Etats-Unis), une échelle d'affinités du cation sodium pour des di- tri- et tétra-peptides a été déterminée par la méthode cinétique de Cooks développée au chapitre II. Nous avons réalisé le pendant théorique de ce projet en calculant les énergies de liaison du cation  $\text{Na}^+$  avec des peptides. Nous verrons que l'accord théorie/expériences est excellent.

## I. Structure des complexes $\text{Na}^+$ - $\text{Gly}_n$ et $\text{Na}^+$ - $\text{Ala}_n$ ( $n \leq 4$ )

Dans une première partie, nous discuterons des différents conformères des complexes  $\text{Gly}_n\text{-Na}^+$  et  $\text{Ala}_n\text{-Na}^+$  ( $n < 5$ ), déterminés par calculs de chimie quantique. Dans la seconde partie nous confronterons le spectre infrarouge obtenu par spectroscopie IRMPD, par couplage du spectromètre de masse MICRA avec le laser à électrons libres CLIO, avec les spectres infrarouge simulés pour les différents conformères (cf. chapitres II et III pour le principe). L'étude de la structure de  $\text{Gly}_n\text{-Na}^+$  a été menée, d'une part, grâce à la technique de cinétique d'échange H/D par l'équipe de J. Beauchamp<sup>1</sup>, où ces auteurs ont montré les limites de cette méthode (cf. Chapitre II I-A), et d'autre part grâce à la technique de mobilité ionique par l'équipe de M. Bowers,<sup>2</sup> où ils ont montré que la structure du peptide dans l'adduit avec le sodium est compacte et non-zwitterionique. Cependant ces deux techniques ne permettent pas de déterminer le mode de cationisation du peptide non-zwitterionique avec le sodium. Il existe plusieurs articles théoriques sur la structure de la diglycine sodiée. Par contre pour les peptides de plus grande taille, seules les structures les plus basses en énergie pour  $\text{Gly}_3\text{-Na}^+$  et  $\text{Gly}_4\text{-Na}^+$  obtenues par l'équipe de Bowers sont disponibles, dans une étude qui ne paraît pas exhaustive.<sup>2</sup>

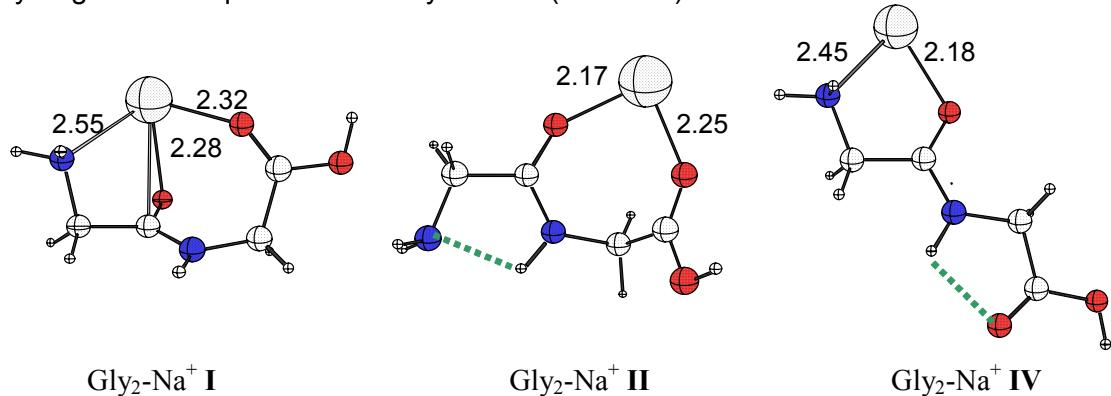
### I.A. Conformères des complexes $\text{Gly}_n\text{-Na}^+$ et $\text{Ala}_n\text{-Na}^+$ ( $n < 5$ )

#### I.A.1. $\text{Gly}_2\text{-Na}^+$ et $\text{Ala}_2\text{-Na}^+$

##### I.A.1.a. $\text{Gly}_2\text{-Na}^+$

L'étude théorique des différents conformères de la diglycine sodiée a déjà été menée par les équipes de G. Ohanessian<sup>3</sup> et A. Daoudi<sup>4</sup>. Le cation sodium, cation dur, privilégie les interactions de nature électrostatique, principalement avec les groupements carbonyles polaires et riches en électrons, puis avec l'amine terminale. Trois conformères de plus basses énergies ont été identifiés, leurs structures sont présentées sur la figure V-1. Nous avons préservé la dénomination des conformères de la référence 3. Le conformère  $\text{Gly}_2\text{-Na}^+$  I présente une coordination du sodium avec les deux oxygènes et l'amine terminale ( $\text{O}^1\text{O}^2\text{N}^1$ ). On peut remarquer que l'environnement du sodium est pyramidal, on peut alors le considérer comme un centre stéréogénique. La plupart des conformères existeront donc sous la forme d'un couple d'énanthiomères. Le sodium est coordonné seulement aux deux oxygènes ( $\text{O}^1\text{O}^2$ ) dans la structure  $\text{Gly}_2\text{-Na}^+$  II ; quant à l'amine terminale, elle reçoit une

liaison hydrogène de l'amide peptidique formant un cycle à cinq chaînons  $\text{C}_5$  ( $\text{N}^2\text{H} \rightarrow \text{N}^1$ ). Enfin dans le conformère  $\text{Gly}_2\text{-Na}^+ \text{IV}$ , le sodium est coordonné uniquement au premier résidu par l'amine terminale et le carbonyle peptidique ( $\text{O}^1\text{N}^1$ ). La structure est donc analogue au complexe  $\text{Gly-Na}^+ \text{a}$ , et le deuxième résidu présente une liaison hydrogène de type  $\text{C}_5$  entre l'hydrogène amidique et le carbonyle acide ( $\text{N}^2\text{H} \rightarrow \text{O}^2$ ).



**Figure V-1 :** Structures optimisées au niveau HF/6-31G\* des adduits  $\text{Gly}_2\text{-Na}^+$ , les énergies correspondantes sont données au tableau V-1.

Les énergies de ces conformères à différents niveaux de calculs sont données dans le tableau V-1. Les deux conformères les plus bas en énergie sont  $\text{Gly}_2\text{-Na}^+ \text{I}$  et  $\text{II}$ . Au niveau B3LYP/GB1//B3LYP/PB1, le conformère  $\text{II}$  est le plus stable de 5.3 kJ/mol, tandis qu'en MP2/GB1//HF/PB1, c'est l'isomère  $\text{I}$ , d'un peu plus d'une kJ/mol. Le conformère  $\text{IV}$  est le plus haut d'une dizaine de kJ/mol. Il semble que la méthode B3LYP/GB1//B3LYP/PB1 sous-estime l'énergie de la structure  $\text{I}$  et donc sous-estime l'interaction de l'amine avec  $\text{Na}^+$ . Par ailleurs, l'adduit de  $\text{Na}^+$  le plus bas en énergie avec une structure zwitterionique de  $\text{Gly}_2$  est au moins une quarantaine de kJ/mol moins stable que  $\text{Gly}_2\text{-Na}^+ \text{I}$  pour toutes les méthodes.

	$\Delta_r\text{H}^\circ(298\text{K})$ HF/PB1	$\Delta_r\text{H}^\circ(298\text{K})$ MP2/GB1/ /HF/PB1	$\Delta_r\text{G}^\circ(298\text{K})$ MP2/GB1/ /HF/PB1	$\Delta_r\text{H}^\circ(298\text{K})$ B3LYP/PB1	$\Delta_r\text{H}^\circ(298\text{K})$ B3LYP/GB1/ /B3LYP/PB1	$\Delta_r\text{G}^\circ(298\text{K})$ B3LYP/GB1/ /B3LYP/PB1
$\text{Gly}_2\text{-Na}^+ \text{I}$	0	<b>0</b>	0	0	0	0
$\text{Gly}_2\text{-Na}^+ \text{II}$	-1.3	1.8	-2.7	0.2	<b>-5.3</b>	-9.8
$\text{Gly}_2\text{-Na}^+ \text{IV}$	16.2	14.3	9.6	14.5	6.3	2.0

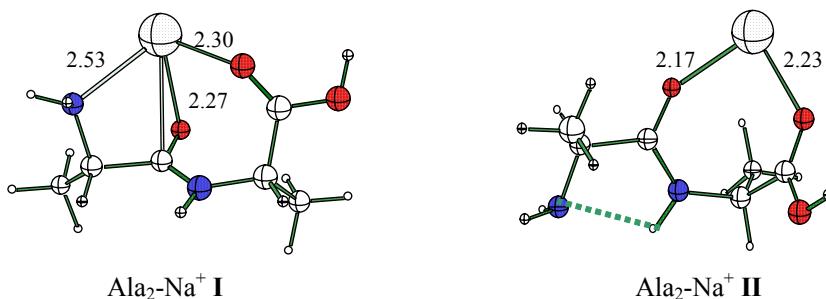
**Tableau V-1 :** Energies relatives en kJ/mol des conformères les plus stables de  $\text{Gly}_2\text{-Na}^+$  à plusieurs niveaux de calculs (GB1 : 6-311+G(2d,2p) ; PB1 : 6-31G\*).

I.A.1.b.  $\text{Ala}_2\text{-Na}^+$ 

Les trois conformères de  $\text{Ala}_2\text{-Na}^+$  les plus bas en énergie ont des structures analogues à celles de  $\text{Gly}_2\text{-Na}^+$ . Dans la plupart des cas, les conformères de  $\text{Ala}_2\text{-Na}^+$  existent sous la forme de couples de diastéréoisomères, car le sodium y est un centre stéréogénique et la configuration des carbones asymétriques doit rester S. Les structures et les valeurs d'énergies relatives sont présentées sur la figure V-2 et dans le tableau V-2. Toutes les méthodes indiquent que le conformère  $\text{Ala}_2\text{-Na}^+ \text{ I}$  est le plus bas en énergie, suivi de  $\text{Ala}_2\text{-Na}^+ \text{ II}$  à 6.8 kJ/mol au niveau MP2/GB1//HF/PB1 et 0.8 kJ/mol au niveau B3LYP/GB1//B3LYP/PB1. Quant au conformère **IV**, il est moins stable d'une dizaine de kJ/mol en *ab initio* et de 4.9 kJ/mol en B3LYP. Là encore, la méthode B3LYP/GB1//B3LYP/PB1 a tendance à défavoriser la structure **I**.

	$\Delta_r\text{H}^\circ(298\text{K})$	$\Delta_r\text{H}^\circ(298\text{K})$	$\Delta_r\text{G}^\circ(298\text{K})$
	HF/PB1	MP2/GB1// HF/PB1	MP2/GB1// HF/PB1
	0	0	0
$\text{Ala}_2\text{-Na}^+ \text{ I}$	2.8	3.2	2.4
	-2.0	2.4	-2.6
$\text{Ala}_2\text{-Na}^+ \text{ II}$	7.5	10.0	5.3
$\text{Ala}_2\text{-Na}^+ \text{ IV}$	16.9	13.6	8.7

**Tableau V-2** : Energies relatives en kJ/mol des conformères les plus stables de  $\text{Ala}_2\text{-Na}^+$  à plusieurs niveaux de calculs (GB1 : 6-311+G(2d,2p) ; PB1 : 6-31G\*).



**Figure V-2** : Structures des diastéréoisomères les plus bas en énergie de  $\text{Ala}_2\text{-Na}^+$ .

Les structures les plus basses en énergie de  $\text{Ala}_2\text{-Na}^+$  et  $\text{Gly}_2\text{-Na}^+$  sont similaires, on peut remarquer cependant que les liaisons métal-ligand sont légèrement plus courtes (0.02 Å) dans le cas de  $\text{Ala}_2\text{-Na}^+$ , ceci pouvant être dû au caractère inductif donneur du méthyle qui

exacerbe légèrement la nucléophilie des oxygènes et à la plus grande polarisabilité des méthyles par rapport à des atomes d'hydrogène. Cette tendance sera confirmée pour les complexes de peptides de plus grande taille.

### I.A.2.Gly<sub>3</sub>-Na<sup>+</sup> et Ala<sub>3</sub>-Na<sup>+</sup>

#### I.A.2.a. Gly<sub>3</sub>-Na<sup>+</sup>

Comme nous l'avons dit auparavant, il n'existe pas d'étude exhaustive des conformères de Gly<sub>3</sub>-Na<sup>+</sup>. Nous avons donc envisagé toutes les possibilités de cationisation de Gly<sub>3</sub> qui possèdent trois groupements carbonyle dont l'acide terminal, ainsi que l'amine terminale, avec lesquels peut interagir le cation sodium. La plus grande partie des conformères envisagés ont été construits par extension des structures de Gly<sub>2</sub>-Na<sup>+</sup>. Les structures des complexes les plus bas en énergie sont présentées sur la figure V-3, leurs énergies sont données dans le tableau V-3 au niveau MP2/GB1//HF/PB1 et pour les plus stables au niveau B3LYP1/GB//B3LYP/PB1.

Le conformère Gly<sub>3</sub>Na<sup>+</sup> **4sON** correspond à la maximisation de l'interaction électrostatique, puisque les trois carbonyles et l'amine ( $\text{O}^1\text{O}^2\text{O}^3\text{N}^1$ ) sont coordinés au sodium. Il est issu d'une extension du côté C-terminal du conformère Gly<sub>2</sub>-Na<sup>+</sup> **I**. Les conformères Gly<sub>3</sub>-Na<sup>+</sup> **3s2** et **3s**, construits par extension du côté N-terminal de Gly<sub>2</sub>-Na<sup>+</sup> **I**, présentent tous deux une coordination du métal avec les trois carbonyles ( $\text{O}^1\text{O}^2\text{O}^3$ ). Ces deux structures sont décrites comme les plus basses en énergie (niveau B3LYP/6-31G) par M. Bowers *et al.*<sup>2</sup> Elles sont quasi-dégénérées avec toutes les méthodes de calcul. Ces trois conformères **4sON**, **3s** et **3s2** sont les plus bas en énergie ; au niveau MP2/GB1//HF/PB1 ils sont quasi-dégénérés, alors qu'au niveau B3LYP/GB1//B3LYP/PB1, on remarque que la structure Gly<sub>3</sub>Na<sup>+</sup> **4sON** est plus haute que **3s** et **3s2** (situation analogue à celle de Gly<sub>2</sub>-Na<sup>+</sup>). Le conformère **4sON**, en maximisant le nombre de sites d'interaction peptide-Na<sup>+</sup>, conduit à la plus grande enthalpie d'interaction. Ceci implique des contraintes fortes sur la structure du peptide, défavorables entropiquement, ce qui conduit à un compromis sur l'énergie libre, qui s'avère être défavorable à **4sON**. Le conformère Gly<sub>3</sub>-Na<sup>+</sup> **3s-boucle** met en jeu une structure plus compacte du peptide où l'acide terminal interagit avec l'amine terminale par le biais d'une liaison hydrogène intramoléculaire ( $\text{O}^4\text{H}\rightarrow\text{N}^1$ ). Les trois carbonyles sont parallèles et pointent alors vers le métal. Nous parlerons, par la suite, de peptide « coiffé » par le sodium, en contraste avec les trois premières structures où le peptide s'enroule autour du cation. Cette structure est relativement haute en énergie, 18 et 24 kJ/mol par rapport à **4sON**, aux niveaux MP2/GB1//HF/PB1 et B3LYP/GB1//B3LYP/PB1, respectivement. Le transfert du proton de l'acide vers l'amine dans la structure **3s-boucle** forme le premier zwitterion **3s-boucle-Z** qui

a pu être identifié sur la surface de potentiel au niveau HF/PB1. Par contre cette structure n'a pu être optimisée au niveau B3LYP/PB1 : elle s'effondre sur la structure **3s-boucle**. La structure **3sON-b**, une trentaine de kJ/mol moins stable de **4sON**, voit l'amine N-terminale ainsi que les deux premiers carbonyles ( $\text{O}^1\text{O}^2\text{N}^1$ ) complexés au sodium, cette structure est issue de  $\text{Gly}_2\text{Na}^+$  **I** à laquelle on a rajouté un résidu du côté C-terminal, le carbonyle de la fonction acide recevant une liaison hydrogène de l'amide du dernier résidu de type  $\text{C}_5$  ( $\text{N}^3\text{H} \rightarrow \text{O}^3$ ). La structure **3sON** est similaire à **3sON-b**, seule la liaison hydrogène  $\text{C}_5$  n'est pas présente, diminuant son énergie de 2 kJ/mol. Quant aux deux structures  $\text{Gly}_3\text{Na}^+$  **2s** et  $\text{Gly}_3\text{Na}^+$  **2s-C<sub>7</sub>**, elles sont issues de  $\text{Gly}_3\text{Na}^+$  **II** étendue du côté C- et N-terminal, respectivement. Elles présentent deux carbonyles complexés au sodium, ce qui laisse plus de degrés de liberté au peptide pour favoriser l'auto-solvatation. Ainsi chacun de ces conformères a deux liaisons hydrogène : deux  $\text{C}_5$  ( $\text{N}^2\text{H} \rightarrow \text{N}^1$  et  $\text{N}^3\text{H} \rightarrow \text{O}^3$ ) et un  $\text{C}_5$  plus un  $\text{C}_7$  ( $\text{N}^2\text{H} \rightarrow \text{N}^1$  et  $\text{N}^3\text{H} \rightarrow \text{O}^1$ ), respectivement. Cependant cette auto-solvatation ne suffit pas pour compenser la perte de l'interaction électrostatique du sodium avec le troisième carbonyle. Ainsi ces structures se trouvent environ 45 kJ/mol moins stables que le conformère le plus bas en énergie. Enfin,  $\text{Gly}_3\text{Na}^+$  **1s-Z** est un complexe zwitterionique où seul le groupement carboxylate est complexé au sodium. Cette structure, qui présente aussi une liaison hydrogène de type  $\text{C}_8$  ( $\text{N}^1\text{H} \rightarrow \text{O}^2$ ), se situe une centaine de kJ/mol au dessus du minimum global.

$\text{Gly}_3\text{Na}^+$	$\Delta_f\text{H}^\circ(298\text{K})$	$\Delta_f\text{H}^\circ(298\text{K})$	$\Delta_f\text{G}^\circ(298\text{K})$	$\Delta_f\text{H}^\circ(298\text{K})$	$\Delta_f\text{H}^\circ(298\text{K})$	$\Delta_f\text{G}^\circ(298\text{K})$	$\Delta E$
	HF/PB1	MP2/GB1/ HF/PB1	MP2/GB1/ HF/PB1		B3LYP/PB1	B3LYP/GB1/ B3LYP/PB1	B3LYP/ 6-31G <sup>2</sup>
<b>4sON</b>	0	<b>0</b>	0.0	0	0	<b>0</b>	
<b>3s2</b>	-5.1	<b>0.0</b>	-2.8	-2.0	<b>-6.3</b>	-9.3	<b>0.1</b>
<b>3s</b>	-5.5	2.3	-1.4	-3.2	<b>-8.4</b>	-13.0	<b>0.0</b>
<b>3sboucle</b>	33.1	18.0	23.2	16.8	23.8	32.3	<b>15.6</b>
<b>3sON-b</b>	32.7	31.9	25.9	29.6	22.6	16.8	
<b>3sON</b>	36.2	33.7	27.6	37.2	27.5	20.6	
converge							
<b>3sboucle-Z</b>	85.7	35.0	43.8	sur 3s-boucle-neutre			
<b>2s</b>	44.2	44.2	30.9	42.2			
<b>2s-C<sub>7</sub></b>	46.8	46.1	39.1				
<b>1s-Z</b>	138.2	106.1	101.2				

Tableau V-3 : Energies relatives en kJ/mol des différents conformères de  $\text{Gly}_3\text{Na}^+$ .

La conclusion de cette exploration est qu'un groupe de conformères (**4sON**, **3s2** et **3s**) est nettement plus stable que les autres, et que dans ces structures,  $\text{Na}^+$  chélate les trois carbonyles et éventuellement l'amine. Nous verrons que cette tendance reste vraie pour  $\text{Gly}_4\text{-Na}^+$ . On peut cependant anticiper qu'elle ne pourra pas être maintenue pour les peptides de taille croissante, la coordination du sodium atteignant la saturation.

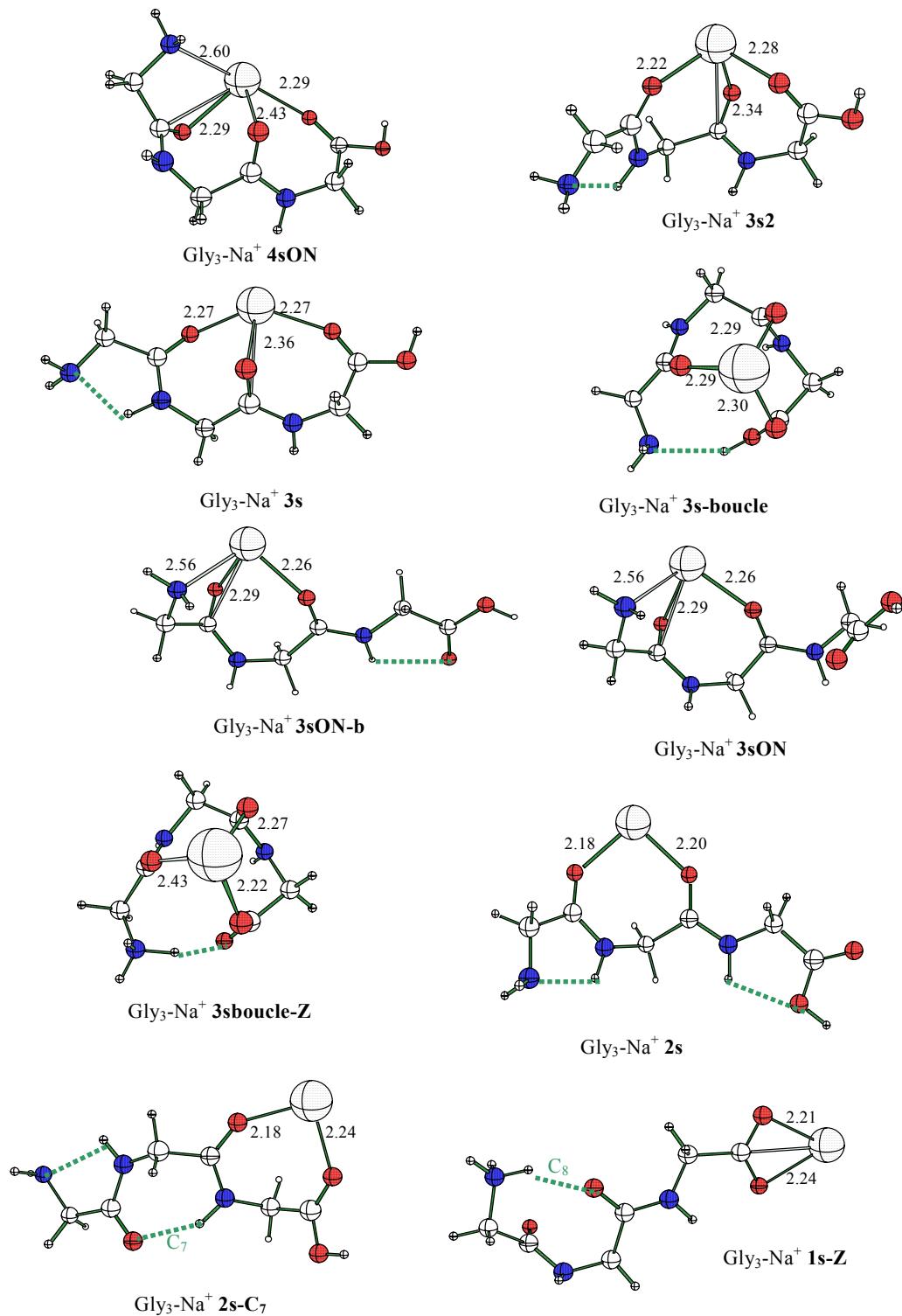


Figure V-3 : Structures des adduits  $\text{Gly}_3\text{-Na}^+$  les plus bas en énergie.

I.A.2.b.  $\text{Ala}_3\text{-Na}^+$ 

Dans le cas de l'étude conformationnelle de  $\text{Ala}_3\text{-Na}^+$ , nous avons utilisé comme points de départ les structures obtenues pour  $\text{Gly}_3\text{-Na}^+$ . Le tableau V-4 donne les énergies relatives des différents conformères (avec les notations de  $\text{Gly}_3\text{-Na}^+$ ) et celles des couples de diastéréoisomères les plus bas en énergie. Nous pouvons remarquer que l'ordre énergétique est peu perturbé par la substitution d'un hydrogène par un groupement méthyle ; de même les énergies relatives sont relativement proches. Par ailleurs, la différence d'énergie des diastéréoisomères **4sON** (0.6 kJ/mol) est faible, contrairement aux **3s2** (7.4 kJ/mol). Ceci est dû au fait que lorsque deux carbonyles consécutifs avec des dipôles parallèles sont coordinés au sodium, on retrouve un motif géométriquement proche du  $\gamma$ -turn ( $\text{C}_7$ ). Il existe alors une position optimale pour le groupement méthyle : équatoriale, qui minimise les interactions stériques. Ceci est illustré sur la figure V-4, pour les diastéréoisomères **3s2** et **3s2-diest** où les méthyles sont en position équatoriale et axiale, respectivement. Dans le cas des deux diastéréoisomères **4sON**, les carbonyles interagissant avec le cation ont des moments dipolaires non parallèles, le motif cyclique à sept chaînons est déformé ; il n'existe pas alors de position particulièrement stabilisante ou déstabilisante pour le méthyle (figure V-4).

$\text{Ala}_3\text{Na}^+$	$\Delta_r\text{H}^\circ(298\text{K})$		$\Delta_r\text{G}^\circ(298\text{K})$
	HF/PB1	MP2/GB1/ /HF/PB1	MP2/GB1/ /HF/PB1
<b>4sON</b>	0	0	0
	1.4	0.6	0.1
<b>3s2</b>	-9.7	<b>-2.6</b>	-6.6
	1.2	4.8	2.5
<b>3s</b>	0.6	9.5	4.8
	3.3	8.8	5.2
<b>3sboucle</b>	30.5	17.0	22.0
<b>3sON-b</b>	34.7	33.7	28.0
<b>3sON</b>	79.8	33.7	41.9
<b>3sboucle-Z</b>	48.5	46.8	40.4
<b>2s</b>	48.2	48.6	36.2
<b>2s-C<sub>7</sub></b>	56.6	53.0	45.9

Tableau V-4 : Energies relatives en kJ/mol des différents conformères de  $\text{Ala}_3\text{-Na}^+$ .

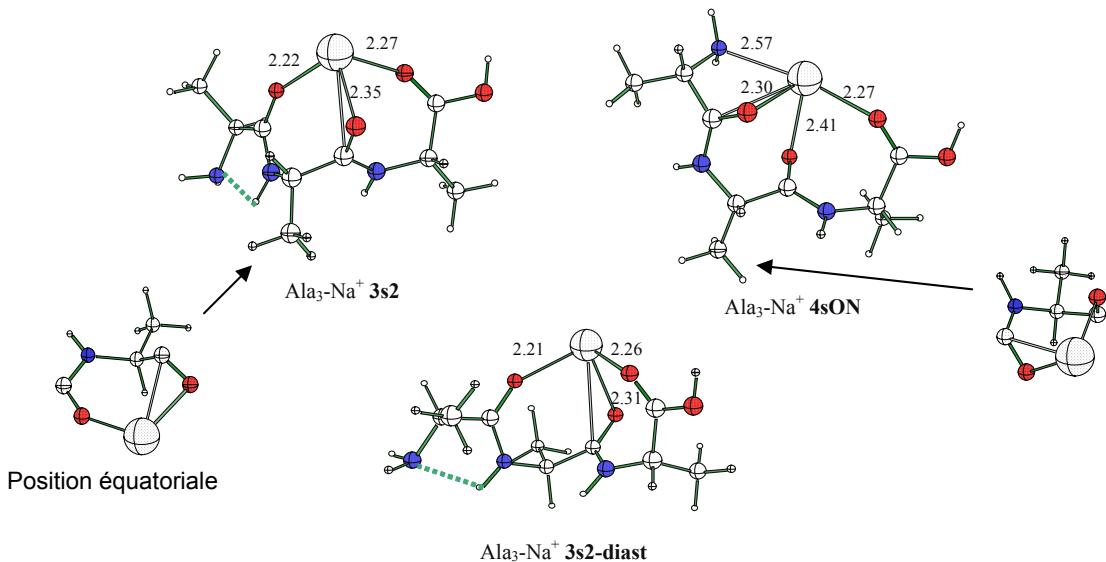


Figure V-4 : Structures des diastéréoisomères les plus bas en énergie de  $\text{Ala}_3\text{-Na}^+$ .

### I.A.3. Gly<sub>4</sub>-Na<sup>+</sup> et Ala<sub>4</sub>-Na<sup>+</sup>

#### I.A.3.a. Gly<sub>4</sub>-Na<sup>+</sup>

Les structures de plus basses énergies sont présentées sur la figure V-5, leurs énergies relatives au conformère le plus bas sont données dans le tableau V-5. Le conformère le plus bas en énergie pour toutes les méthodes de calculs est  $\text{Gly}_4\text{Na}^+ \mathbf{4s}$ , issu de  $\text{Gly}_3\text{Na}^+ \mathbf{3s}$  par extension du côté C-terminal. Le sodium est coordonné aux quatre carbonyles ( $\text{O}^1\text{O}^2\text{O}^3\text{O}^4$ ). Les deux structures suivantes sont  $\mathbf{5sON-b}$  et  $\mathbf{5sON}$  (issues de  $\text{Gly}_3\text{Na}^+ \mathbf{4sON}$  par extension du côté C-terminal). Elles présentent une coordination du sodium par les quatre carbonyles et par l'amine terminale ( $\text{O}^1\text{O}^2\text{O}^3\text{O}^4\text{N}^1$ ) et diffèrent d'une part par l'orientation de l'amine terminale, et d'autre part par le fait que le reste des squelettes carbonés sont énantiomères l'un de l'autre. Comme nous l'avons déjà remarqué, ces structures, où l'amine terminale joue le rôle de ligand, sont défavorisées en B3LYP/GB1//B3LYP/PB1 puisqu'elles se trouvent environ 14 kJ/mol plus hautes en énergie que la structure sans cette interaction, alors qu'au niveau MP2/GB1//HF/PB elles sont quasi-dégénérées.

$\text{Gly}_4\text{Na}^+$	$\Delta_r\text{H}^\circ(298\text{K})$ HF/PB1	$\Delta_r\text{H}^\circ(298\text{K})$	$\Delta_r\text{G}^\circ(298\text{K})$	$\Delta\text{E}$ B3LYP/ 6-31G <sup>2</sup>
		MP2/GB1/ /HF/PB1	MP2/GB1/ /HF/PB1	
<b>4s</b>	0.0	<b>0.0</b>	0.0	0
<b>5sON-b</b>	10.3	3.1	5.0	
<b>5sON</b>	12.1	4.1	6.0	
<b>3sboucle</b>	34.4	11.8	20.3	
<b>3sboucle-b</b>	33.6	21.7	21.7	
<b>3s-C<sub>10</sub></b>	40.9	24.8	23.5	
<b>4sboucle</b>	50.7	26.8	34.2	20.3
<b>3s-Z</b>	79.0	27.6	38.2	
<b>4sON</b>	36.7	28.1	25.7	
<b>3s-C<sub>7</sub></b>	41.9	36.3	33.8	

**Tableau V-5** : Energies relatives en kJ/mol des différents conformères de  $\text{Gly}_4\text{Na}^+$ .

Les trois structures suivantes présentent un peptide coiffé par le sodium en interaction avec trois des carbonyles. La plus stable, **3sboucle**, est coordinée par ( $\text{O}^1\text{O}^3\text{O}^4$ ) au sodium, et le peptide porte une liaison hydrogène intramoléculaire ( $\text{O}^4\text{H}\rightarrow\text{N}^1$ ) qui lie l'acide terminal à l'amine. Ce conformère est le premier qui ne présente pas une coordination des carbonyles tous consécutifs, ce qui permet de minimiser l'interaction stérique entre ces groupements : toutes les liaisons métal-peptide sont alors courtes (2.25 Å en moyenne). Nous verrons que ce type d'interaction faisant intervenir des carbonyles non consécutifs sera présent dans les conformères de plus basse énergie de plus grands peptides (cf. chapitre VI). Les deux conformères **3s-C<sub>10</sub>** et **3sboucle-b** sont issus de  $\text{Gly}_3\text{Na}^+$  **3sboucle** étendu du côté N-terminal et C-terminal, respectivement. Le composé  $\text{Gly}_4\text{Na}^+$  **3s-C<sub>10</sub>** fait apparaître une liaison hydrogène de type  $\text{C}_{10}$  ( $\text{N}^4\text{H}\rightarrow\text{O}^1$ ), alors qu'il existe une interaction faible entre un des hydrogènes de l'amine et le carbone terminal (distance 2.75 Å) pour le composé **3sboucle-b**. La structure **4sboucle** est dérivée de la structure  $\text{Gly}_3\text{Na}^+$  **3sboucle**, où le peptide a une structure beaucoup plus compacte avec une liaison hydrogène intramoléculaire ( $\text{O}^4\text{H}\rightarrow\text{N}^1$ ), liant ses deux extrémités. Les groupements carbonyles sont parallèles et le sodium vient coiffer le peptide. Cette structure est 27 kJ/mol moins stable que le conformère le plus bas ; cette forte déstabilisation est due à une forte contrainte géométrique et stérique du peptide. Le premier isomère zwitterionique **3s-Z** a une énergie relative de 28 kJ/mol ; il présente une structure du peptide coiffé ( $\text{O}^2\text{O}^3\text{O}^4$ ) et une liaison H ( $\text{N}^1\text{H}\rightarrow\text{O}^4$ ). Cette structure est iso-énergétique de **4sON**, dérivée de **5sON**, avec le dernier résidu en structure de brin  $\beta$ . La structure la plus haute (36 kJ/mol) **3s-C<sub>7</sub>** présente une conformation proche de **3s-C<sub>10</sub>**, les

deux premiers résidus étant impliqués dans une interaction de type  $\text{C}_7$  ( $\text{N}^3\text{H} \rightarrow \text{O}^1$ ). La liaison H étant moins directionnelle dans une structure  $\text{C}_7$  que dans un  $\text{C}_{10}$ , cela explique qu'elle soit moins stable par rapport à **3s-C<sub>10</sub>**.

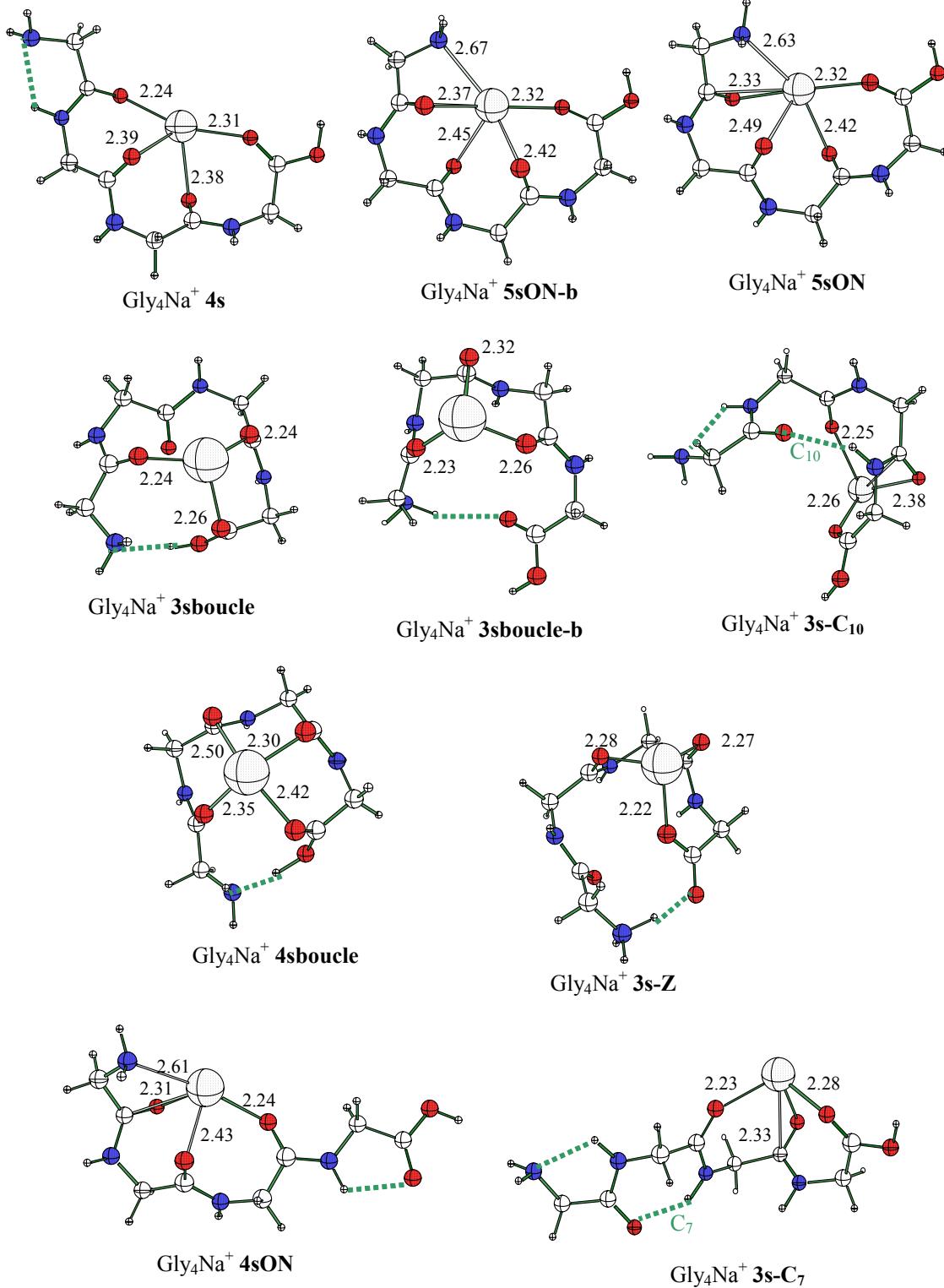


Figure V-5 : Structures des adduits  $\text{Gly}_4\text{-Na}^+$  les plus bas en énergie.

I.A.3.b.  $\text{Ala}_4\text{-Na}^+$ 

$\text{Ala}_4\text{Na}^+$	$\Delta_f\text{H}^\circ(298\text{K})$ HF/PB1	$\Delta_f\text{H}^\circ(298\text{K})$ MP2/GB1/ /HF/PB1	$\Delta_f\text{G}^\circ(298\text{K})$ MP2/GB1/ /HF/PB1
<b>4s</b>	0	0	0
<b>5sON-b</b>	7.1	-0.7	1.1
	4.2	-0.6	1.0
<b>5sON</b>	5.3	-1.2	0.8
	4.3	<b>-2.6</b>	-0.3
<b>3sboucle</b>	18.4	<b>-3.8</b>	6.0
	22.7	4.0	12.6
<b>4sboucle</b>	38.1		
<b>3sboucle-b</b>	56.8	50.4	50.8
<b>3s-Z-2</b>	73.3		
<b>4sON</b>	27.7	17.0	16.7

 Tableau V-6 : Energies relatives en kJ/mol des différents conformères de  $\text{Ala}_4\text{-Na}^+$ .

Seuls les conformères les plus stables de  $\text{Gly}_4\text{-Na}^+$  ont été étendus pour le calcul des structures de  $\text{Ala}_4\text{-Na}^+$ , nous avons notamment optimisé les structures des couples de diastéréoisomères **4s**, **5sON-b**, **5sON** et **3sboucle**. Leurs énergies relatives à la structure **4s** sont présentées dans le tableau V-6. Les structures les plus basse **3sboucle** et **5sON** ainsi que celle du diastéréoisomère **3sboucle-diast** sont données dans la figure V-6. Les structures **4s**, **5sON** et **5sON-b** sont très proches en énergie comme dans le cas de  $\text{Gly}_4\text{-Na}^+$ , par ailleurs la différence d'énergie entre les couples de diastéréoisomères sont très faibles : 1.4 kJ/mol au plus. Ceci est du, comme dans le cas de  $\text{Ala}_3\text{-Na}^+$  **4sON**, au fait que les carbonyles consécutifs ont des moments dipolaires non parallèles, le motif cyclique à sept chaînons est alors déformé ; la position du méthyle influence peu l'énergie de ces structures. Au contraire le conformère **3sboucle** présente deux carbonyles consécutifs et parallèles coordinés au sodium, la position du méthyle la plus stable est équatoriale. On retrouve le gain de 8 kJ/mol de cette structure par rapport celle où le méthyle est axial (**3sboucle-diast**) que l'on avait observé pour  $\text{Ala}_2\text{-Na}^+$  **3s2**.

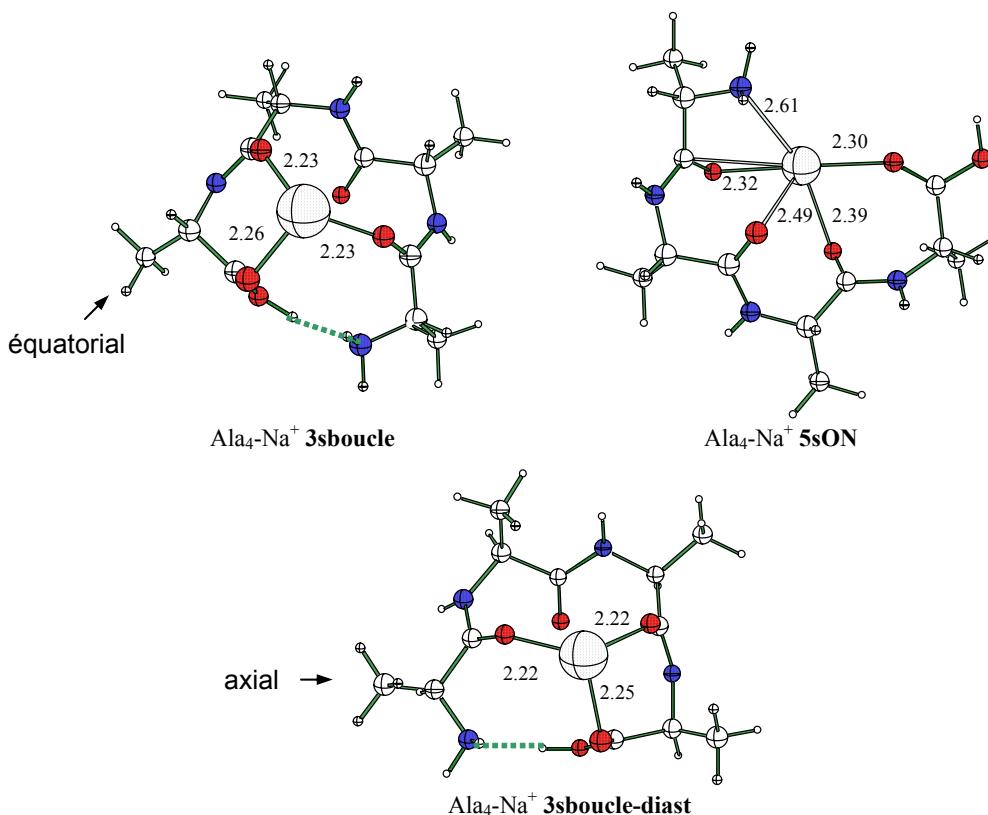


Figure V-6 : Structures des diastéréoisomères les plus bas en énergie de  $\text{Ala}_4\text{-Na}^+$ .

#### I.A.4. Conclusion

Pour les structures de  $\text{Gly}_2\text{-Na}^+$  à  $\text{Gly}_4\text{-Na}^+$ , les conformères de plus basse énergie sont ceux qui maximisent l'interaction électrostatique du métal avec les groupements carbonyles, avec ou sans l'amine terminale. On peut distinguer deux types de structure du peptide autour du métal, d'une part une structure enroulée et d'autre part une structure coiffée, cette dernière étant défavorisée entropiquement. Les structures de plus basse énergie sont identiques dans les cas des oligoalanines et des oligoglycines, excepté pour les tétra-peptides où  $\text{Ala}_4\text{-Na}^+$  présente une structure coiffée par le sodium, alors que  $\text{Gly}_4\text{Na}^+$  a une structure enroulée.

#### I.B. Spectres IRMPD

Le principe de cette spectroscopie a été développé dans le chapitre II, et le montage original utilisé a fait l'objet d'une partie du chapitre III.

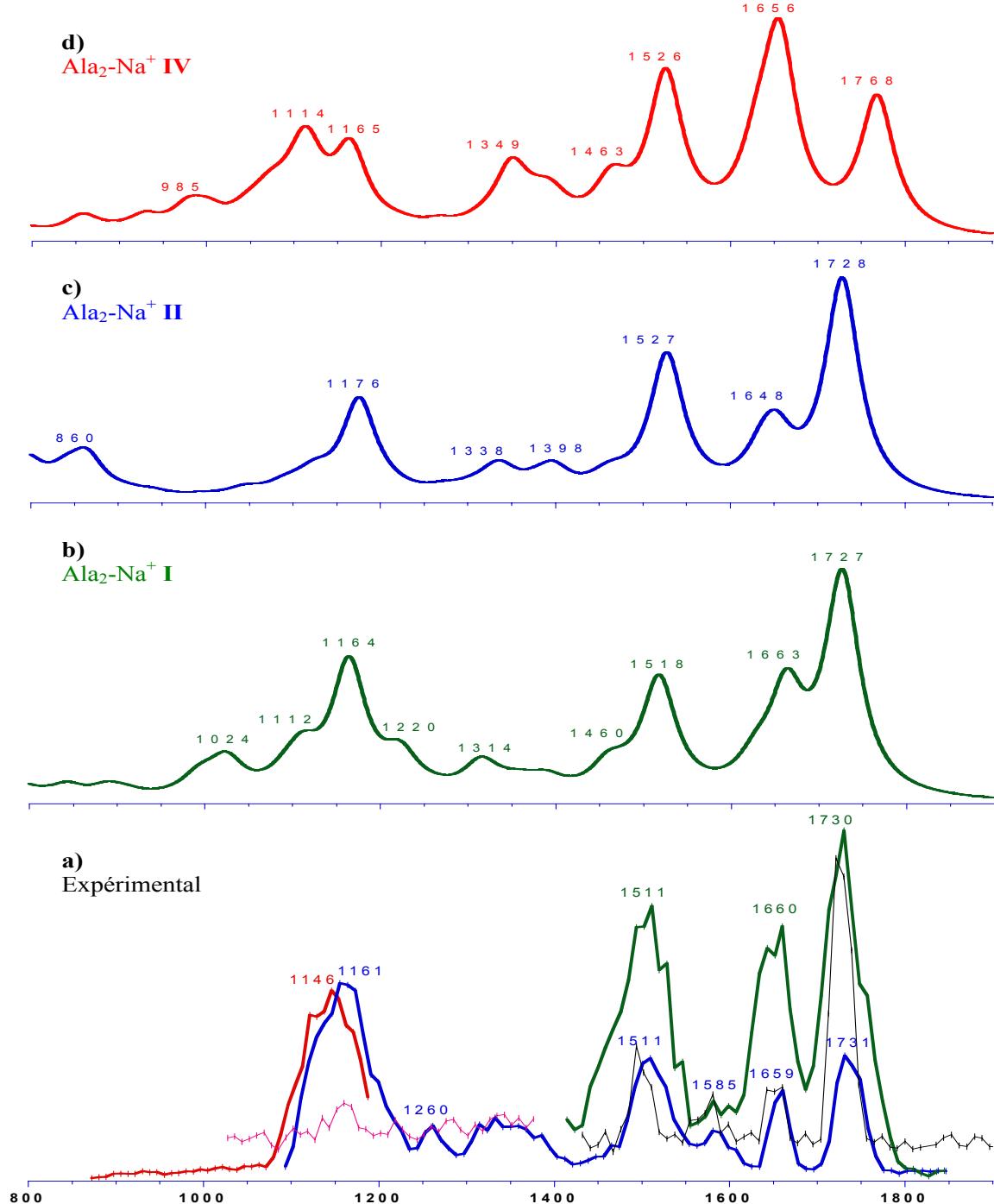
### I.B.1. Echantillons MALDI

Bien que le type d'échantillon « sandwich » utilisé,<sup>5</sup> dans le cadre de la production d'ions par la technique de MALDI, ait déjà été évoqué dans le chapitre III sur l'exemple des acides aminés, nous allons donner ici la composition des échantillons utilisés pour produire les ions  $\text{Gly}_2\text{-Na}^+$ ,  $\text{Gly}_3\text{-Na}^+$ ,  $\text{Ala}_2\text{-Na}^+$  et  $\text{Ala}_3\text{-Na}^+$ . Nous avons utilisé dans tous les cas les mêmes proportions d'analyte et de matrice. Le premier dépôt correspond à 10  $\mu\text{L}$  de solution de matrice à  $5 \cdot 10^{-2}$  mol/L, le deuxième à 3  $\mu\text{L}$  d'une solution aqueuse de NaCl à  $5 \cdot 10^{-2}$  mol/L, puis 20  $\mu\text{L}$  d'une solution aqueuse d'oligopeptide à  $5 \cdot 10^{-2}$  mol/L et enfin la dernière couche est identique à la première. Nous avons privilégié l'utilisation de l'acide  $\alpha$ -cyano-4-hydroxy-cinnamique (4-HCCA) comme matrice dans un solvant eau/acétone dans les proportions (1/3) en volume excepté dans le cas de  $\text{Gly}_3\text{-Na}^+$ , où nous avons utilisé l'acide 2,5-dihydroxy-benzoïque (DHB) dans l'eau, à cause de la coïncidence de masse entre  $\text{Gly}_3\text{-Na}^+$  et 4-HCCA-Na<sup>+</sup> (MM=212 g/mol). Dans ce cas, nous avons utilisé le laser Nd :YAG dans son mode non-focalisé. Cette technique de MALDI interne permet de former les complexes peptides-Na<sup>+</sup> en quantité suffisante pour pouvoir sélectionner l'ion puis le fragmenter par IRMPD. Cependant elle présente certaines limites : d'une part une instabilité du nombre d'ions produits à chaque impact laser, d'autre part le dépôt étant relativement mince, il est nécessaire de changer le point d'impact du laser quand la quantité d'échantillon devient trop faible. Le spectre IRMPD global est alors la superposition de spectres obtenus à différentes positions d'impact du laser sur la cible et à différentes intensités du LEL, afin d'éviter de saturer la fragmentation.

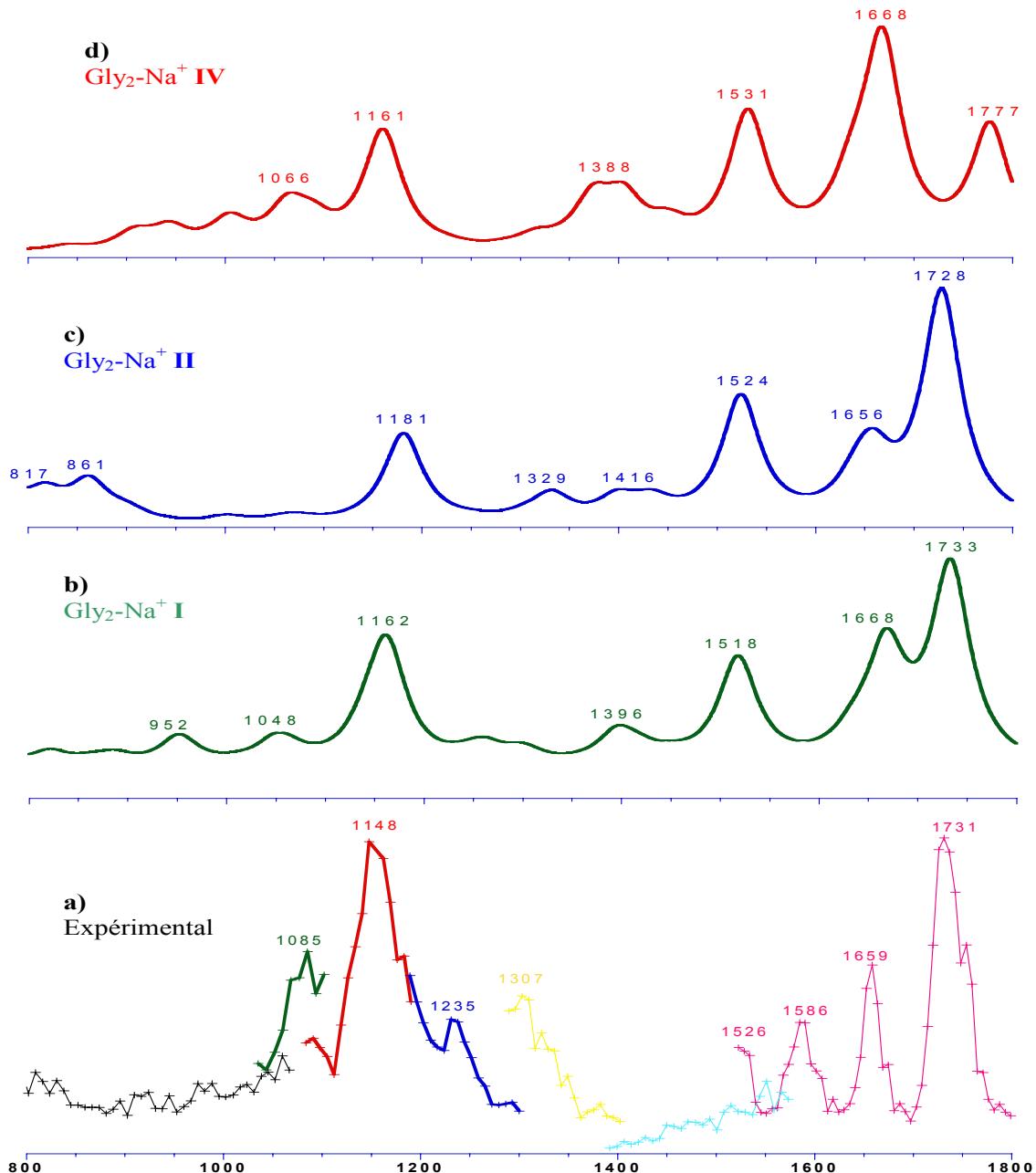
### I.B.2. Spectroscopie IRMPD de $\text{Ala}_2\text{-Na}^+$ et $\text{Gly}_2\text{-Na}^+$

Les ions  $\text{Ala}_2\text{-Na}^+$  et  $\text{Gly}_2\text{-Na}^+$  ont été sélectionnés en masse 50 à 70 ms après leur formation par MALDI. Ces ions sont ensuite piégés pendant 1000 ms dans la cellule afin de les thermaliser, puis ils sont irradiés pendant 1250 ms par le laser à électrons libres, avec des photons dont l'énergie varie entre 800 et 1800  $\text{cm}^{-1}$ . Les spectres IRMPD expérimentaux de  $\text{Ala}_2\text{-Na}^+$  et  $\text{Gly}_2\text{-Na}^+$  sont présentés sur la partie a) des figures V-7 et V-8, respectivement, tandis que les spectres d'absorption IR à un photon simulés au niveau B3LYP/6-31G\* dans l'approximation harmonique (scaling factor standard de 0.9613) pour les conformères I, II et IV sont donnés dans les parties b), c) et d) des figures V-7 et V-8. Comme nous l'avons dit précédemment, le spectre expérimental global est composé d'une superposition de spectres partiels qui sont définis d'une part par un emplacement différent sur la cible et d'autre part par la puissance du LEL. En effet, comme le montrent les simulations théoriques, la section efficace d'absorption des modes entre 1600 et 1800  $\text{cm}^{-1}$  est la plus importante dans la

gamme étudiée, ce qui peut provoquer une saturation de la fragmentation (rendement proche de 100%), les bandes expérimentales étant alors mal résolues. Une solution est d'utiliser un atténuateur de la puissance du LEL. Les parties du spectre en trait gras sont celles obtenues sans atténuateur, celles avec un trait fin, en atténuant trois fois la puissance du laser.



**Figure V-7 :** (a) Spectre IRMPD de  $\text{Ala}_2\text{-Na}^+$ , (b-d) spectres IR des conformères de plus basse énergie calculés au niveau B3LYP/6-31G\*.



**Figure V-8 :** (a) Spectre IRMPD de  $\text{Gly}_2\text{-Na}^+$ , (b-d) spectres IR des conformères de plus basse énergie calculés au niveau B3LYP/6-31G\*.

Le spectre IRMPD de  $\text{Ala}_2\text{-Na}^+$  présente quatre bandes d'absorption intenses à (1146-1161), 1511, 1660 et 1730  $\text{cm}^{-1}$ . On peut remarquer que la courbe en noir, qui correspond au laser atténué dans la gamme 1400-1800  $\text{cm}^{-1}$ , permet d'avoir des pics beaucoup plus résolus que dans le cas du LEL non atténué (courbe verte). Afin d'interpréter ce spectre nous l'avons comparé à ceux simulés pour les trois conformères de plus basse énergie I, II et IV. De plus les fréquences expérimentales de dissociation ainsi que les fréquences, intensités et une description succincte des modes de vibration théoriques d'intensité supérieure à 30  $\text{km/mol}$  sont tabulées dans la partie droite du tableau V-7. La

bande à  $1511 \text{ cm}^{-1}$  est caractéristique d'un couplage entre les modes de déformation angulaire du N-H peptidique et du C-H voisin (amide II). La bande à  $1660 \text{ cm}^{-1}$  correspond à la vibration d'elongation du C=O amidique (amide I). Ces bandes de fréquences voisines pour les trois conformères ne permettent pas d'identification certaine. Par contre, l'absorption expérimentale à  $1146-1161 \text{ cm}^{-1}$  peut être associée au mode d'elongation C-OH de tous les conformères, cependant l'absence du profil de double bande caractéristique de l'isomère **IV** (la bande à  $1114 \text{ cm}^{-1}$  de **IV** est un mode délocalisé qui existe aussi pour **I** et **II**, mais avec des intensités plus faibles), indique une présence prédominante des conformères **I** et **II**. De plus, la bande de vibration d'elongation C=O de la fonction acide à  $1730 \text{ cm}^{-1}$ , qui est bien résolue, est compatible avec les bandes des conformères **I** et **II** à  $1727$  et  $1728 \text{ cm}^{-1}$  respectivement, mais pas avec celle du conformère **IV** à  $1768 \text{ cm}^{-1}$ .

En conclusion, si on se base sur les fréquences au maximum des absorptions, on peut remarquer une excellente concordance avec les fréquences simulées pour le conformère **I**. Cependant la largeur des bandes, les déplacements possibles dus aux processus multiphotoniques et les incertitudes sur les fréquences calculées ne nous permettent pas d'affirmer de manière définitive que le conformère **II** n'est pas formé. Par contre, les bandes relatives aux vibrations d'elongation C=O et C-OH permettent d'exclure la présence du conformère **IV**.

Le spectre IRMPD de  $\text{Gly}_2\text{-Na}^+$  présenté figure V-5 a) est de bien moins bonne qualité que celui de  $\text{Ala}_2\text{-Na}^+$ , ceci étant dû notamment à la difficulté de produire les ions de manière stable. Cependant il présente trois bandes d'absorption intenses à  $1586$ ,  $1659$  et  $1731 \text{ cm}^{-1}$ , une partie de bande vers  $1526 \text{ cm}^{-1}$  et une bande intense à  $1148 \text{ cm}^{-1}$ . La zone entre  $1300$  et  $1526 \text{ cm}^{-1}$ , paraît difficilement exploitable. La portion de bande à  $1526 \text{ cm}^{-1}$  semble correspondre au couplage entre les modes de déformation angulaire du N-H peptidique et du  $\text{CH}_2$  voisin. On retrouve la bande à  $1659 \text{ cm}^{-1}$  caractéristique de la vibration d'elongation C=O amidique. Cependant comme dans le cas du spectre de  $\text{Ala}_2\text{-Na}^+$ , ces deux bandes ne sont pas caractéristiques d'un des trois conformères de plus basse énergie. Par contre la bande de vibration d'elongation de C=O acide situé à  $1731 \text{ cm}^{-1}$  n'est caractéristique que des conformères **I** et **II** et l'absence de fragmentation à  $1777 \text{ cm}^{-1}$  exclut la présence du conformère **IV**. Par ailleurs, on observe une bande à  $1586 \text{ cm}^{-1}$ . On peut remarquer qu'un résultat similaire a été obtenu pour  $\text{Ala}_2\text{-Na}^+$  (figure V-7-a)) où une bande de faible intensité a été enregistrée à  $1585 \text{ cm}^{-1}$ . L'origine de cette bande pour  $\text{Gly}_2\text{-Na}^+$  et  $\text{Ala}_2\text{-Na}^+$  n'est pas connue.

Ainsi bien que de moins bonne qualité, le spectre IRMPD de  $\text{Gly}_2\text{-Na}^+$  permet d'exclure la présence du conformère **IV**, comme pour  $\text{Ala}_2\text{-Na}^+$ . Par contre, les maxima des pics observés ne permettent pas de proposer une différentiation possible entre les conformères **I** et ou **II**.

Gly <sub>2</sub> -Na <sup>+</sup>						Ala <sub>2</sub> -Na <sup>+</sup>				
Fréquences (cm <sup>-1</sup> )		Intensité calculée (Km/mol)			modes	Intensité calculée (Km/mol)		Fréquences (cm <sup>-1</sup> )		
Exp	Calc	I	II	IV		I	II	IV	Calc	Exp
	817	55								
	861	83					37		840	
	907	33						31	858	
	944	40					79		864	
	952	56			$\omega \text{ NH}_2 + \nu \text{ C-N}$ + $\tau \text{ NH ...}$		38		978	
	1005	53					52		1024	
	1048	37					32		1075	
1085	1064	75					58		1112	
(1060-1100)	<b>1089</b>	<b>40</b>					136		1114	
	1142	41					38		1125	1146-1161
1156	1161	212					<b>222</b>		<b>1164</b>	(1107-1196)
(1126-1189)	1162	223			$\nu \text{ C-OH}$		<b>134</b>		<b>1165</b>	
	1181	203					206		1176	
1235 ? (1225-1248)										
	1261	32					58		1223	
1307?										
	1329	40			$\omega \text{ NH} + \omega \text{ CHR}$		47		1314	
	1374	87			+ $\delta \text{ R} + \delta \text{ COH}$		44		1338	
	1396	57			+ ...		102		1349	
	1398	48					31		1397	
	1405	83					49		1398	
					$2 \delta \text{ CH}_3$		41		1463	
	1518	201					<b>196</b>		<b>1518</b>	
	1524	282			$\omega \text{ NH} + \omega \text{ CHR}$		262		1526	1511
1526?					+ $\delta \text{ R} + \delta \text{ COH}$					(1485-1530)
	1531	253			+ ...		299		1527	
1586 ? (1565-1605)										1585?
	1635	34					40		1628	
	1636	80			$\delta \text{ NH}_2$		86		1630	
	1642	58					57		1636	
1659	<b>1658</b>	<b>120</b>					103		1653	1660
(1645-1673)	1668	385			$\nu \text{ C=O}$		317		1656	(1644-1670)
	1668	208					<b>163</b>		1663	
1731	<b>1728</b>	519					<b>372</b>		<b>1727</b>	1730
(1717-1760)	<b>1733</b>	371			$\nu \text{ C=O(acide)}$		<b>460</b>		<b>1728</b>	(1712-1742)
	1777	232					229		1768	

Tableau V-7 : Fréquences mesurées et calculées pour les complexes Gly<sub>2</sub>-Na<sup>+</sup> et Ala<sub>2</sub>-Na<sup>+</sup>.

### I.B.3. Spectroscopie IRMPD de $\text{Gly}_3\text{-Na}^+$ et $\text{Ala}_3\text{-Na}^+$

Les ions  $\text{Gly}_3\text{-Na}^+$  et  $\text{Ala}_3\text{-Na}^+$  ont été sélectionnés en masse, 100 ms après leur formation par MALDI. Ces ions sont ensuite piégés pendant 1000 ms dans la cellule afin de les thermaliser, puis ils sont irradiés pendant 1250 ms par le laser à électrons libres, avec des photons dont l'énergie varie entre 800 et 2000  $\text{cm}^{-1}$ . Les spectres IRMPD expérimentaux de  $\text{Gly}_3\text{-Na}^+$  et  $\text{Ala}_3\text{-Na}^+$  sont présentés sur la partie a) des figures V-9 et V-10, respectivement, tandis que les spectres d'absorption IR à un photon simulés au niveau B3LYP/6-31G\* dans l'approximation harmonique (scaling factor standard de 0.9613) pour les conformères enroulés autour du sodium **4sON**, **3s** et **3s2**, pour le conformère coiffé **3s-boucle** et pour le conformère **3sON-b** sont donnés dans les parties b), c) et d) des figures V-9 et V-10.

Comparons, tout d'abord, les spectres IR théoriques des différents conformères de  $\text{Gly}_3\text{-Na}^+$ . On peut remarquer que les trois spectres théoriques de **4sON**, **3s** et **3s2** sont relativement semblables. Les quatre bandes d'absorption les plus intenses sont dues d'une part à la vibration d'elongation C-OH dans la gamme 1161 à 1165  $\text{cm}^{-1}$ , un couplage de vibrations de déformation angulaire des groupements N-H C-OH et  $\text{CH}_2$  de 1512-1522  $\text{cm}^{-1}$ , la vibration d'elongation C=O des groupements peptidiques (1657-1689) et enfin l'elongation C=O acide (1742-1749). La différence structurale entre les trois conformères est due essentiellement à la position du  $\text{NH}_2\text{-CH}_2$  terminal, or les vibrations associées à ce groupement étant peu intenses, cela explique la forte similitude de ces trois spectres infrarouge. Il sera donc difficile au vu de la résolution accessible dans les spectres IRMPD expérimentaux, de différencier ces trois conformères. Le conformère **3s-boucle** présente quatre pics caractéristiques intenses, les deux premières bandes relativement larges correspondent à la superposition de plusieurs pics, d'une part de torsions  $\text{NH}_2$  et du squelette peptidique vers 1084  $\text{cm}^{-1}$ , et d'autre part de déformation angulaire (ou bending) des groupements NH,  $\text{CH}_2$  et C-OH vers 1298  $\text{cm}^{-1}$ . La bande à 1504  $\text{cm}^{-1}$  correspond à la vibration de bending  $\text{NH}_2$ . Quant aux vibrations d'elongation C=O, elles sont toutes les trois couplées, contrairement aux autres conformères où la vibration du carbonyle acide reste découplée de celles des carbonyles amidiques. Seule la composante symétrique conduit à une bande d'intensité supérieure à 50 km/mol, elle atteint 656 km/mol et se situe à 1721  $\text{cm}^{-1}$ . Le spectre IR calculé pour le conformère **3sON-b** présente lui aussi quatre bandes intenses, la première à 1160  $\text{cm}^{-1}$  correspond à la convolution de deux pics : la vibration de torsion des groupements  $\text{NH}_2$  et  $\text{N-C}_\alpha\text{-C}$  et la vibration de l'elongation C-OH. La bande à 1521  $\text{cm}^{-1}$  correspond à la superposition de deux déformations angulaires, notamment des NH peptidiques. Enfin les deux bandes à 1687 et 1778  $\text{cm}^{-1}$  correspondent aux vibrations d'elongation C=O amidiques et acide respectivement. Dans le cas du

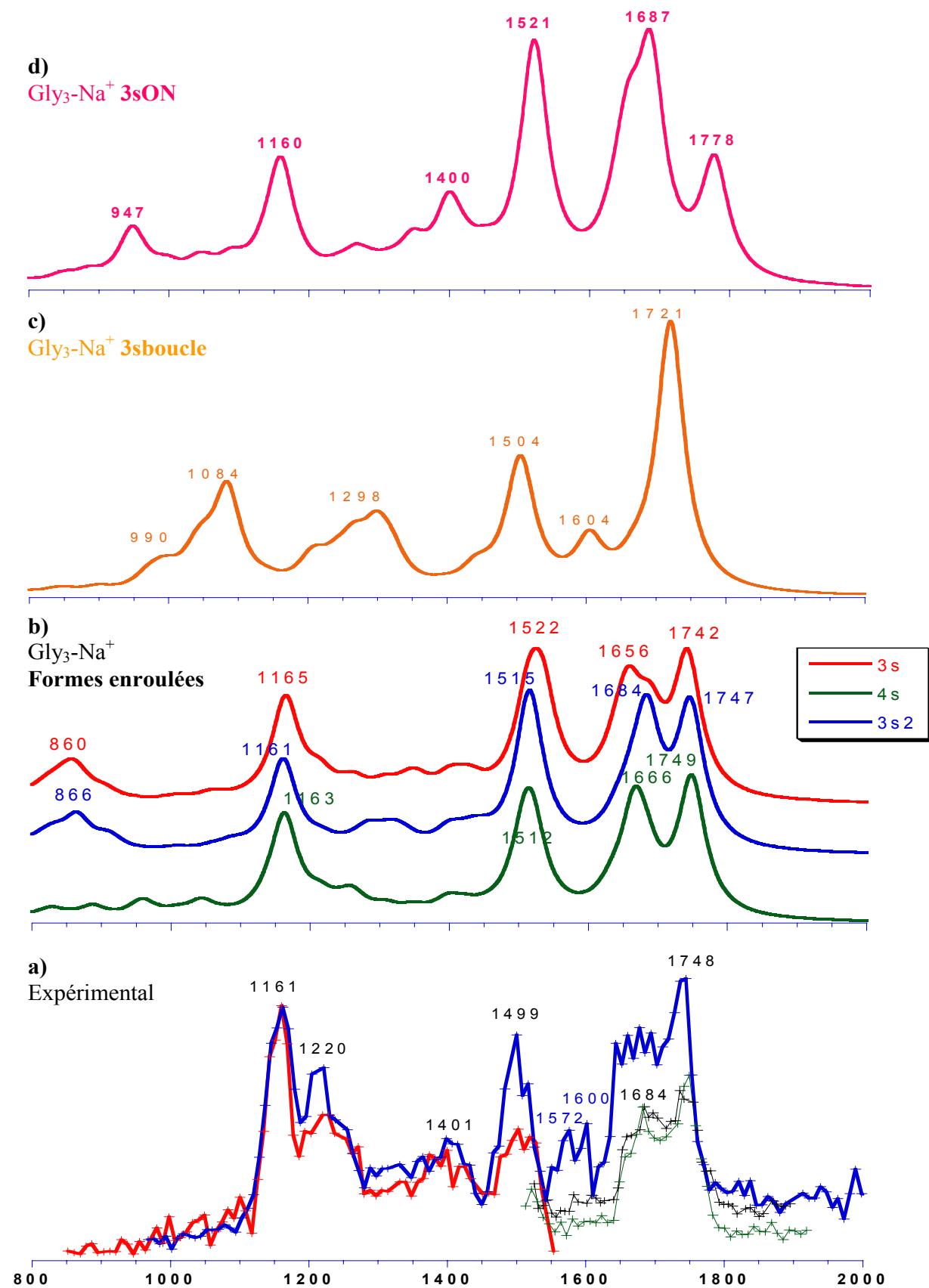
conformère **3sON-b**, le carbonyle acide n'étant pas impliqué dans une interaction avec le métal, sa fréquence de vibration d'elongation est plus élevée que dans le cas des autres complexes.

Le spectre IRMPD expérimental de  $\text{Gly}_3\text{-Na}^+$  est présenté sur la figure V-9-a. On peut constater que le rapport signal sur bruit est relativement mauvais, cependant cela est dû en partie à la congestion de bandes d'absorption. En effet, dans la zone  $800\text{-}2000\text{ cm}^{-1}$ , on dénombre trente bandes d'absorption théoriques pour chaque conformère avec des intensités variant de 1 à  $660\text{ km/mol}$  ; les bandes d'intensité supérieure à  $50\text{ km/mol}$  ainsi que leurs assignations sont présentées dans le tableau V-8. Ce spectre présente quatre bandes d'absorption intenses dont les maxima se situent à  $1161$ ,  $1220$ ,  $1499$  et la gamme de  $1648$  à  $1771\text{ cm}^{-1}$ . Nous avons pu constater une saturation de la fragmentation en utilisant le LEL non atténué dans la gamme  $1650\text{-}1750\text{ cm}^{-1}$  (courbe bleue). Cette même zone infrarouge sondée avec le LEL atténué (courbes verte et noire) permet de dédoubler cette bande en deux bandes mieux résolues centrées à  $1684$  et  $1748\text{ cm}^{-1}$ . La bande à  $1161\text{ cm}^{-1}$  peut être associée à la vibration d'elongation C-OH des conformères **4sON**, **3s2**, **3s** et **3sON-b**, avec un accord quantitativement meilleur pour **4sON**, **3s2** et **3s**. La bande relativement fine à  $1220\text{ cm}^{-1}$  ne peut être identifiée que comme la vibration de bending des groupements NH peptidiques et des  $\text{CH}_2$  avoisinants du conformère **3s** (cf. tableau V-8). La bande à  $1499\text{ cm}^{-1}$ , qui correspond à un couplage de déformation angulaire des groupements NH peptidiques, des  $\text{CH}_2$  voisins et du C-OH, est commune à tous les conformères et donc n'est pas spécifique. Comme dans le cas des dipeptides, la zone des vibrations d'elongation C=O est la plus caractéristique des conformères. En effet le profil en double bande est caractéristique des conformères **4sON**, **3s2** **3s** et **3sON-b**. Cependant dans le cas du complexe **3sON-b**, les bandes sont beaucoup plus séparées que pour les trois autres car l'acide terminal n'interagit pas avec le cation. Dans le spectre expérimental, l'absence d'absorption à  $1789\text{ cm}^{-1}$  caractéristique de **3sON** permet d'exclure la présence de ce conformère. L'étude de la gamme  $1600\text{-}1800\text{ cm}^{-1}$  nous permet donc de conclure que les structures majoritaires sont les conformères qui présentent un enroulement du peptide autour du cation. La présence du conformère **3sboucle** ne peut être confirmé ou infirmé, dans cette gamme de longueur d'onde. Cependant au vu de l'allure de double bande du spectre IRMPD, on peut déjà dire qu'il est minoritaire. Par ailleurs l'absence de bande d'absorption expérimentale autour de  $1084\text{ cm}^{-1}$  permet de confirmer l'absence de ce conformère.

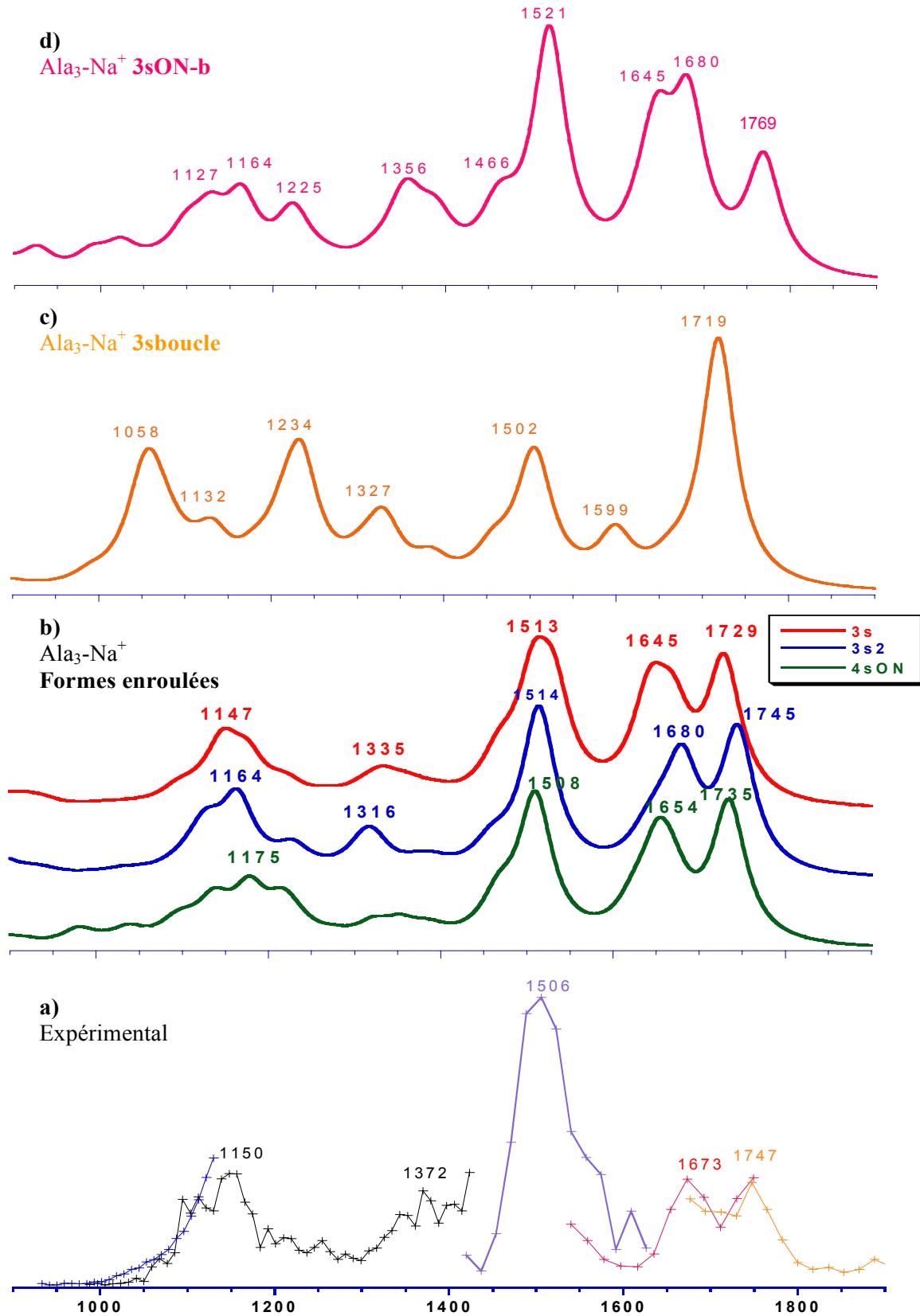
En conclusion, la spectroscopie IRMPD, si elle ne permet pas d'identifier chaque conformère (la différenciation des conformères **4sON**, **3s2** et **3s** est impossible), permet de différencier des classes de conformères : d'une part ceux qui comportent des nombres différents de carbonyles complexés au métal (par exemple **4sON**, **3s2**, **3s**, **3sboucle** vs.

**3sON-b**), d'autre part ceux dans lesquels le peptide est enroulé autour du métal et ceux où il est « coiffé » par le sodium (**4sON**, **3s2**, **3s** vs. **3sboucle**). Dans le cas de  $\text{Gly}_3\text{-Na}^+$ , la spectroscopie IRMPD nous a permis d'identifier la présence du conformère **4sON** et/ou **3s2** et/ou **3s**. Rappelons que ce sont les trois conformères calculés les plus bas en énergie.

Nous avons effectué une expérience similaire avec  $\text{Ala}_3\text{-Na}^+$ . Le spectre IRMPD présenté dans la figure V-10-a). Bien que de moins bonne qualité que celui de  $\text{Gly}_3\text{-Na}^+$  (due essentiellement à des problèmes techniques de production d'ions et de temps), ce spectre en présente les mêmes caractéristiques, avec quatre bandes d'absorption situées à 1150, 1506, 1673 et  $1747\text{ cm}^{-1}$ . L'allure des spectres théoriques est identique à celle des conformères de  $\text{Gly}_3\text{-Na}^+$ , excepté quelques légères modifications de fréquences et d'intensité des bandes autour de 1000 à  $1300\text{ cm}^{-1}$ , qui correspondent à des modes très couplés, notamment avec des déformations angulaires des groupements méthyle. La conclusion est alors identique à celle de  $\text{Gly}_3\text{-Na}^+$ : La spectroscopie IRMPD permet d'identifier des classes de conformères qui diffèrent soit par le nombre de carbonyles complexés au sodium, soit par le mode de complexation enroulé ou coiffé du peptide. Ainsi dans le cas de  $\text{Ala}_3\text{-Na}^+$ , seuls les conformères **4sON** et/ou **3s2** et/ou **3s** ont pu être identifiés. En effet le profil de double bande et la différence de fréquences dans la région des vibrations d'elongation C=O permet d'identifier les conformères **4sON**, **3s2** et **3s** et d'exclure le conformère **3sON-b**. De plus, l'absence de fragmentation autour de  $1058\text{ cm}^{-1}$ , montre que le conformère **3sboucle** est absent.



**Figure V-9 :** (a) Spectre IRMPD de  $\text{Gly}_3\text{-Na}^+$ , (b-d) spectres IR des conformères de plus basse énergie calculés au niveau B3LYP/6-31G\*.



**Figure V-10 :** (a) Spectre IRMPD de  $\text{Ala}_3\text{-Na}^+$ , (b-d) spectres IR des conformères de plus basse énergie calculés au niveau B3LYP/6-31G\*.

<b>.Gly<sub>3</sub>-Na<sup>+</sup></b>						
<b>Fréquence (cm<sup>-1</sup>)</b>		<b>Intensité calculée (Km/mol)</b>				
<b>Exp</b>	<b>Calc</b>	<b>4sON</b>	<b>3s2</b>	<b>3s</b>	<b>3s</b> <b>boucle</b>	<b>mode</b>
	860			75		
	866			50		$\tau \text{ NH}_2 + \tau$
	947				70	NCC
	958		52			
	1038			51		tt est couplé
	1084			239		
1161	<b>1160</b>				233	
(1132-	<b>1161</b>		<b>268</b>			$\nu \text{ C-OH}$
1178)	<b>1163</b>		<b>292</b>			
	<b>1165</b>			<b>295</b>		
1220	1214			57		$\omega \text{ NH} + \omega$
	1264			71		tt est couplé
	1298			107		
	1347				63	
	1400				139	
	<b>1501</b>			<b>149</b>		
	1508		178			
1499	1509			179		
	1514		265			$\omega \text{ NH} + \delta \text{ C-}$
(1468-	1514			249		$\text{OH} + \omega \text{ CH}_2$
1527)	1517		200			+ ...
	1519				154	
	1519		222			
	1522				317	
	1534			262		
1572 ?						
1600	<b>1604</b>			<b>111</b>		$\delta \text{ NH}_2$
	1641			68		
	1657			249		
	1654				228	$2 \nu^{\text{as}} \text{ C=O}$
	1663		71			
	1666		288			
	<b>1683</b>		<b>85</b>			
1684	<b>1684</b>		<b>358</b>			$2 \nu^{\text{s}} \text{ C=O}$
(1648-	1687				407	
	1689			176	407	
	1721			<b>656</b>		$3 \nu^{\text{s}} \text{ C=O}$
	1742			400		
1748	<b>1747</b>		<b>406</b>			
	<b>1749</b>		<b>397</b>			$\nu \text{ C=O(acide)}$
	1778				230	

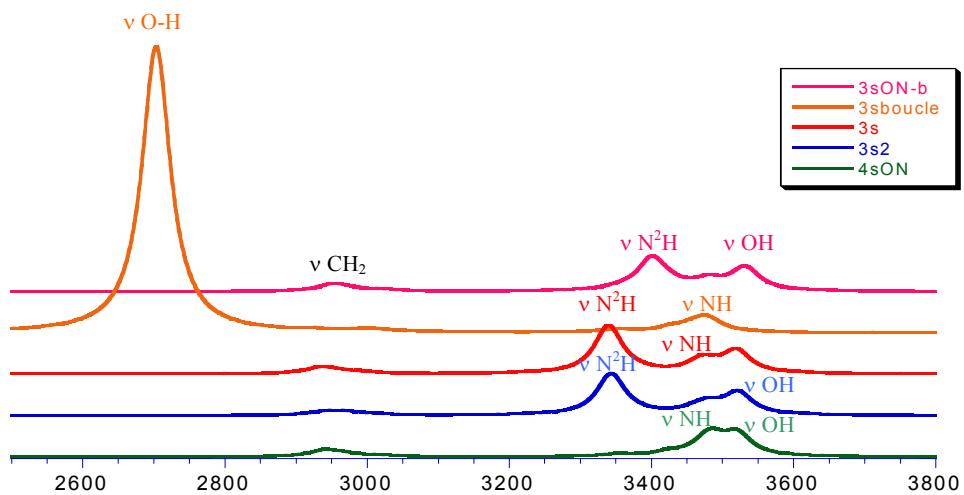
Tableau V-8 : Fréquences mesurées et calculées pour les complexes Gly<sub>3</sub>-Na<sup>+</sup> et Ala<sub>3</sub>-Na<sup>+</sup>.

Ala <sub>3</sub> -Na <sup>+</sup>						
Fréquence (cm <sup>-1</sup> )		Intensité calculée (Km/mol)				mode
Exp	Calc	4sON	3s2	3s	3s boucle	
	865			71		$\tau \text{ NH}_2 + \tau$
	871			82		NCC + $\tau$
	926				50	CH <sub>3</sub>
	1051			159		tt est couplé
	1063			114		
	1127				92	
	1128		102			$\tau \text{ NH}_2 + \tau$
1150	1132			85		NCC + $\tau$ CH <sub>3</sub>
(1090-	1137	107				
	1147		160			
1177)	1164		205		131	
	1164					$\nu \text{ C-OH}$
	1176		110			
	1178	139				
	1214			83		
	1217	102				
	1225				82	
	1228	65				
	1234			237		
	1316	62				
	1326			63		
	1334			64		
	1335		51			
	1347				55	$\omega \text{ NH} + \delta \text{ C-}$
	1356				83	OH + $\omega \text{ CH}_2$
	1464				55	+ $\omega \text{ CH}_3 \dots$
	1500		138			
	1507		298			
	1508	217				
1506	1510		161			
(1468-	1512	196				
1555)	1514	248				
	1514	222				
	1519			73		
	1521			391		
	1531	265				
	1599		105			
	1625	69				$\delta \text{ NH}_2$
	1637	53				
	1637		77			
	1645		202			
	1645			226		2 $\nu^{\text{as}} \text{ C=O}$
	1652	231				
	1654		73			
1673	1669	109				
	1670		198			2 $\nu^{\text{s}} \text{ C=O}$
(1644-	1680	286				
1778)	1682			302		
	1719			551		3 $\nu^{\text{s}} \text{ C=O}$
	1729		395			
	1735	394				
1747	1745		400			$\nu \text{ C=O(acide)}$
	1769			223		

Tableau V-8bis : Fréquences mesurées et calculées pour les complexes Ala<sub>3</sub>-Na<sup>+</sup>.

### I.B.4. Conclusion

La spectroscopie IRMPD est une technique efficace pour identifier différentes classes de conformères, d'une part celles qui possèdent des nombres différents de carbonyles complexés au sodium, d'autre part celles dont le motif de complexation est tel que le peptide est enroulé autour du sodium de celles où le cation coiffe le peptide. Cependant cette technique reste limitée pour identifier les conformères dans une même classe, comme les conformères **I** et **II** de  $\text{AA}_2\text{-Na}^+$  (AA : Acide Aminé) ou les conformères **4sON**, **3s2** et **3s** de  $\text{AA}_3\text{-Na}^+$ , dans la gamme de longueur d'onde accessible au laser à électrons libres. Ceci est notamment du au problème de la résolution des bandes IRMPD dont la largeur est d'environ  $50 \text{ cm}^{-1}$ . Cette largeur de bande est due d'une part à la largeur de bande du laser qui est déterminée comme étant 0.5% de  $\lambda$  (soit entre 5 et 10  $\text{cm}^{-1}$ , dans la gamme 1000-2000  $\text{cm}^{-1}$ ), mais surtout au fait que l'on ne travaille pas à basse température. Les raies sont alors élargies à cause du couplage ro-vibrationnel. Une solution pour améliorer la résolution serait de produire des ions froids, ce qui est en principe possible en couplant la technique de MALDI avec une détente supersonique, mais pas dans la configuration actuelle du montage expérimental MICRA. Une autre solution serait de sonder les vibrations X-H dans la gamme 2500-3800  $\text{cm}^{-1}$ . En effet les spectres théoriques des différents conformères ont des caractéristiques spectrales différentes dans cette gamme de longueur d'onde. Ceci est illustré par les spectres calculés pour les cinq conformères de  $\text{Gly}_3\text{-Na}^+$  discutés plus haut. Ils sont présentés par la figure V-11. Le conformère **3sboucle** voit la fréquence de sa bande d'élongation O-H fortement déplacée vers les petits nombres d'onde, car cet hydrogène est impliqué dans une liaison hydrogène. Par ailleurs **3s2** et **3s** et **3sON-b** possèdent une bande supplémentaire aux autres conformères vers 3330  $\text{cm}^{-1}$  qui correspond à la vibration d'élongation N-H du deuxième résidu pour **3s2** et **3s**, du troisième pour **3sON-b** : ces hydrogènes étant impliqués dans une liaison hydrogène, voient leurs fréquences de vibration diminuées de manière plus importante pour **3s2** et **3s**. Cette gamme de longueur d'onde permet alors de différencier les conformères **3s2** et **3s** des conformères **4sON**, **3sON-b** et **3sboucle**. Seuls les conformères **3s2** et **3s** ne peuvent être discriminés. Le problème actuel pour réaliser ce type d'expérience est que le LEL CLIO est accordable jusqu'à 3000  $\text{cm}^{-1}$  mais avec de faibles puissances dans cette gamme de longueur d'onde. Une alternative serait d'utiliser un laser OPO, mais sa puissance moyenne, beaucoup plus faible que celle d'un LEL ( $\approx 20 \text{ mW}$ ), rend le processus IRMPD moins efficace. Les premiers essais n'ont pas été concluants.



**Figure V-11 :** Spectres IR simulés dans la gamme 2500-3800  $\text{cm}^{-1}$  pour les cinq conformères de plus basse énergie de  $\text{Gly}_3\text{-Na}^+$  au niveau B3LYP/6-31G\*.

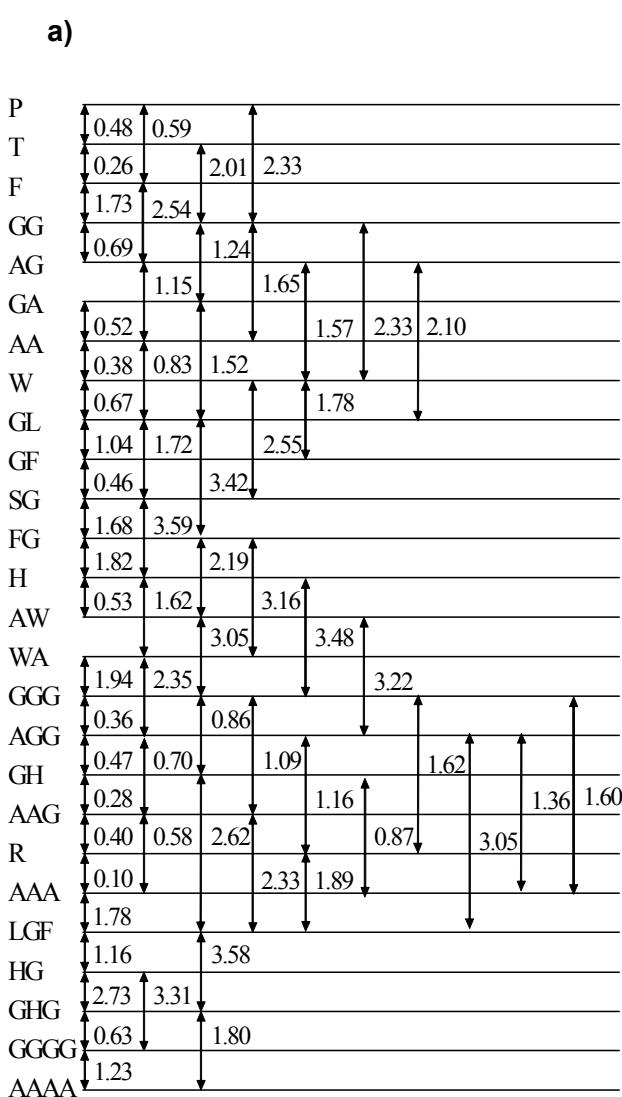
## II. Energies de liaison peptide- $\text{Na}^+$

Cette seconde partie concerne la détermination des énergies de liaison de peptides avec le cation  $\text{Na}^+$ . Ce travail a été mené en collaboration avec l'équipe de C. Wesdemiotis qui a produit les résultats expérimentaux. Nous nous sommes intéressés à la détermination des énergies de liaison théoriques. Cette partie sera décomposée en trois paragraphes qui présenteront en premier lieu les résultats expérimentaux, puis la détermination des conformères neutres les plus bas en énergie pour les peptides concernés :  $\text{Gly}_2$ ,  $\text{Ala}_2$ ,  $\text{Gly}_3$ ,  $\text{Ala}_3$ ,  $\text{Gly}_4$  et  $\text{Ala}_4$ . Le dernier paragraphe concerne la confrontation entre expérience et calculs.

### II.A. Echelle expérimentale d'affinités de peptides pour le sodium

L'équipe de C. Wesdemiotis utilise la technique de détermination d'énergie de liaison par la méthode cinétique de Cooks. Le principe de cette technique a été développé dans le chapitre II. Rappelons que dans le cadre de cette expérience, on étudie la décomposition induite par collision d'un dimère cationisé  $\text{Pep-Na}^+-\text{Ref}$  entre le peptide (Pep) et un composé de référence (Ref) d'affinité connue pour  $\text{Na}^+$ , produit par ESI. En général, les seules voies de dissociation de basse énergie sont les éliminations compétitives de Ref (formant  $\text{Pep-Na}^+$ ) et de Pep (formant  $\text{Na}^+-\text{Ref}$ ). Moyennant quelques approximations, le logarithme du rapport entre les nombres d'ions  $\text{Pep-Na}^+$  et  $\text{Na}^+-\text{Ref}$  est directement égal au logarithme du rapport des constantes de vitesse de dissociation de  $\text{Pep-Na}^+$  et  $\text{Na}^+-\text{Ref}$ . Des expériences similaires ont été faites précédemment pour déterminer une échelle d'affinité du sodium pour 17 acides aminés<sup>6</sup> et pour  $\text{Gly}_2$ ,<sup>7</sup> ce qui permet d'avoir plusieurs molécules de référence. La

série de  $\ln\left(\frac{k_{\text{Pep}-\text{Na}^+}}{k_{\text{Na}^+-\text{Ref}}}\right)$  obtenue dans le cas de plusieurs peptides est représentée dans le tableau V-9-a.



b)

Pep	$\ln(k/k_{\text{GG}})$	$\Delta H_{\text{Na}}^{a,c}$	$\Delta H_{\text{Na}}^{b,c}$
	(std. dev.)	362 K	435 K
Pro	-2.40 (0.17)	<b>196</b>	194
Thr	-2.03 (0.16)	<b>197</b>	196
Phe	-1.79 (0.22)	<b>198</b>	196
Gly <sub>2</sub>	0.00	<b>203</b>	203
AlaGly	0.69 (0.06)	<b>205</b>	205
GlyAla	1.24 (0.07)	<b>207</b>	207
Ala <sub>2</sub>	1.75 (0.31)	<b>208</b>	209
Trp	2.24 (0.23)	<b>210</b>	211
GlyLeu	2.76 (0.25)	<b>211</b>	213
GlyPhe	3.91 (0.25)	215	<b>217</b>
SerGly	4.55 (0.26)	217	<b>219</b>
PheGly	6.20 (0.35)	222	<b>225</b>
His	8.08 (0.33)	227	<b>232</b>
AlaTrp	8.51 (0.35)	229	<b>234</b>
TrpAla	9.53 (0.35)	232	<b>237</b>
Gly <sub>3</sub>	11.53 (0.36)	238	<b>245</b>
AlaGly <sub>2</sub>	11.83 (0.40)	239	<b>246</b>
GlyHis	12.35 (0.38)	240	<b>248</b>
AlaAlaGly	12.59 (0.40)	241	<b>248</b>
Arg	13.05 (0.39)	242	<b>250</b>
Ala <sub>3</sub>	13.17 (0.40)	243	<b>251</b>
LeuGlyPhe	14.93 (0.40)	248	<b>257</b>
HisGly	16.09 (0.46)	251	<b>261</b>
GlyHisGly	18.67 (0.48)	259	<b>270</b>
Gly <sub>4</sub>	19.35 (0.50)	261	<b>273</b>
Ala <sub>4</sub>	20.53 (0.49)	265	<b>277</b>

Tableau V-9 : a) Echelle de  $\ln(k_1/k_2)$  b) Energies de liaison en kJ/mol obtenues pour différent peptides à deux températures effectives. Les valeurs de  $\ln(k/k_{\text{GG}})$  sont des moyennes de celles du tableau a). Les valeurs en gras sont considérées comme les meilleures estimations expérimentales.

La conversion des valeurs de  $\ln\left(\frac{k_{\text{Pep}-\text{Na}^+}}{k_{\text{Na}^+-\text{Ref}}}\right)$  en énergies de liaison relatives demande de déterminer la température effective  $T_{\text{eff}}$ , qui n'est pas une grandeur thermodynamique puisque les ions ne sont pas à l'équilibre. En pratique deux températures effectives ont été

utilisées pour la détermination des énergies de liaison,  $T_{\text{eff}}=362$  K pour les di-peptides  $\text{Gly}_2$ ,  $\text{Ala}_2$ ,  $\text{GlyAla}$ ,  $\text{AlaGly}$  et  $\text{GlyLeu}$  et  $T_{\text{eff}}=435$  K pour les autres di-peptides, les tri-peptides et les tétra-peptides. En effet la température effective peut être plus importante pour les systèmes ayant un plus grand nombre de degrés de liberté, ce qui correspond dans ce cas à une plus forte énergie d'interaction métal-ligand. Les valeurs des enthalpies standard de dissociation sont données dans le tableau V-9-b.

## II.B. Détermination théorique d'affinités de peptides pour le sodium

### II.B.1. Rappel sur la détermination d'énergie de liaison

Afin de déterminer théoriquement l'affinité pour le sodium de peptides (Pep), il est nécessaire d'optimiser les géométries du complexe Pep- $\text{Na}^+$  et du peptide seul, afin de déterminer les propriétés énergétiques de chacun. En effet on a pour la réaction :  $\text{Pep-}\text{Na}^+ \rightarrow \text{Pep} + \text{Na}^+$

$$\Delta_r\text{U}^\circ(\text{T}) = E(\text{Pep}) + E(\text{Na}^+) - E(\text{Pep-}\text{Na}^+) + (E_{\text{th}}(\text{T}, \text{Pep}) + E_{\text{th}}(\text{T}, \text{Na}^+) - E_{\text{th}}(\text{T}, \text{Pep-}\text{Na}^+))$$

où E est l'énergie à 0 K et  $E_{\text{th}}(\text{T})$  est l'énergie thermique à la température T qui prend en compte les composantes vibrationnelle, rotationnelle et translationnelle. Dans le cas d'un cation atomique, seule l'énergie translationnelle est à prendre en compte donc

$$E_{\text{th}}(\text{T}, \text{Na}^+) = \frac{3}{2}RT.$$

L'affinité cationique correspond à l'enthalpie standard  $\Delta_r\text{H}^\circ(\text{T})$  de la réaction précédente. Rappelons que  $\Delta_r\text{H}^\circ(\text{T}) = \Delta_r\text{U}^\circ(\text{T}) - RT$  dans le cas de cette réaction.

On peut aussi accéder à la variation d'enthalpie libre où  $\Delta_r\text{G}^\circ(\text{T}) = \Delta_r\text{H}^\circ(\text{T}) - T \Delta_r\text{S}^\circ(\text{T})$ .

Des travaux précédents au laboratoire ont montré que le calcul des affinités pour le sodium au niveau MP2/6-311+G(2d,2p)//HF/6-31G\* (MP2/GB1//HF/PB1) sans correction de BSSE donne des résultats en bon accord avec les expériences<sup>8, 9</sup> nous avons donc utilisé ce niveau de calcul.

### II.B.2. Conformères de plus basse énergie de $\text{Gly}_2$ et $\text{Ala}_2$

La conformation du di-peptide glycylglycine a déjà été étudiée par les groupes de C. J. Cassady,<sup>10</sup> de G. Ohanessian<sup>3</sup> et C. W. Tsang<sup>11</sup>. Les deux structures les plus basses en énergie sont présentées sur la figure V-12, leurs énergies relatives sont données dans le tableau V-10. Ces deux structures sont quasi-dégénérées au niveau MP2/GB1//HF/PB1, cependant le facteur entropique favorise le conformère I.  $\text{Gly}_2$  I présente une structure de

type brin  $\beta$  avec deux liaisons hydrogène intramoléculaires de type  $\text{C}_5$  ( $\text{N}^2\text{H} \rightarrow \text{N}^1$  et  $\text{N}^2\text{H} \rightarrow \text{O}^2$ ), tandis que  $\text{Gly}_2$  **II** en présente une de type  $\text{C}_5$  ( $\text{N}^2\text{H} \rightarrow \text{N}^1$ ) et une de type  $\text{C}_7$  ( $\text{O}^2\text{H} \rightarrow \text{O}^1$ ) où l'hydrogène de l'acide en conformation trans peut interagir avec le groupement carbonyle peptidique. Remarquons que la structure **II** existe sous la forme des conformères droite et gauche ou  $\text{C}_7$  axial et équatorial. Dans le cas de la glycine qui ne possède pas de chaîne latérale, ces deux conformations sont dégénérées.

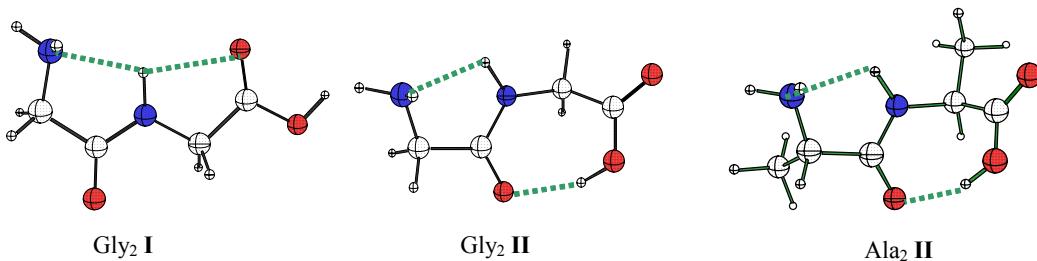


Figure V-12 : Structures de  $\text{Gly}_2$  et structure de  $\text{Ala}_2$  les plus basses en énergie.

Dans le cas de  $\text{Ala}_2$ , les structures les plus basses sont similaires à celles de  $\text{Gly}_2$ . La présence des méthyles dans  $\text{Ala}_2$  introduit l'existence des deux conformations axiale et équatoriale, non dégénérées en énergie ; la conformation  $\text{C}_7$  équatoriale est la plus stable, comme nous l'avons vu dans le chapitre I. Les conformères **I** et **II** restent très proches en énergie.

	$\Delta_r\text{H}^\circ(298\text{K})$ HF/PB1	$\Delta_r\text{H}^\circ(298\text{K})$ MP2/GB1/ /HF/PB1	$\Delta_r\text{G}^\circ(298\text{K})$ MP2/GB1/ /HF/PB1		$\Delta_r\text{H}^\circ(298\text{K})$ HF/PB1	$\Delta_r\text{H}^\circ(298\text{K})$ MP2/GB1/ /HF/PB1	$\Delta_r\text{G}^\circ(298\text{K})$ MP2/GB1/ /HF/PB1
$\text{Gly}_2$ <b>I</b>	0	0	0	$\text{Ala}_2$ <b>I</b>	0	0	0
$\text{Gly}_2$ <b>II</b>	10.6	-0.2	4.6	$\text{Ala}_2$ <b>II</b>	6.1	-3.2	0.6

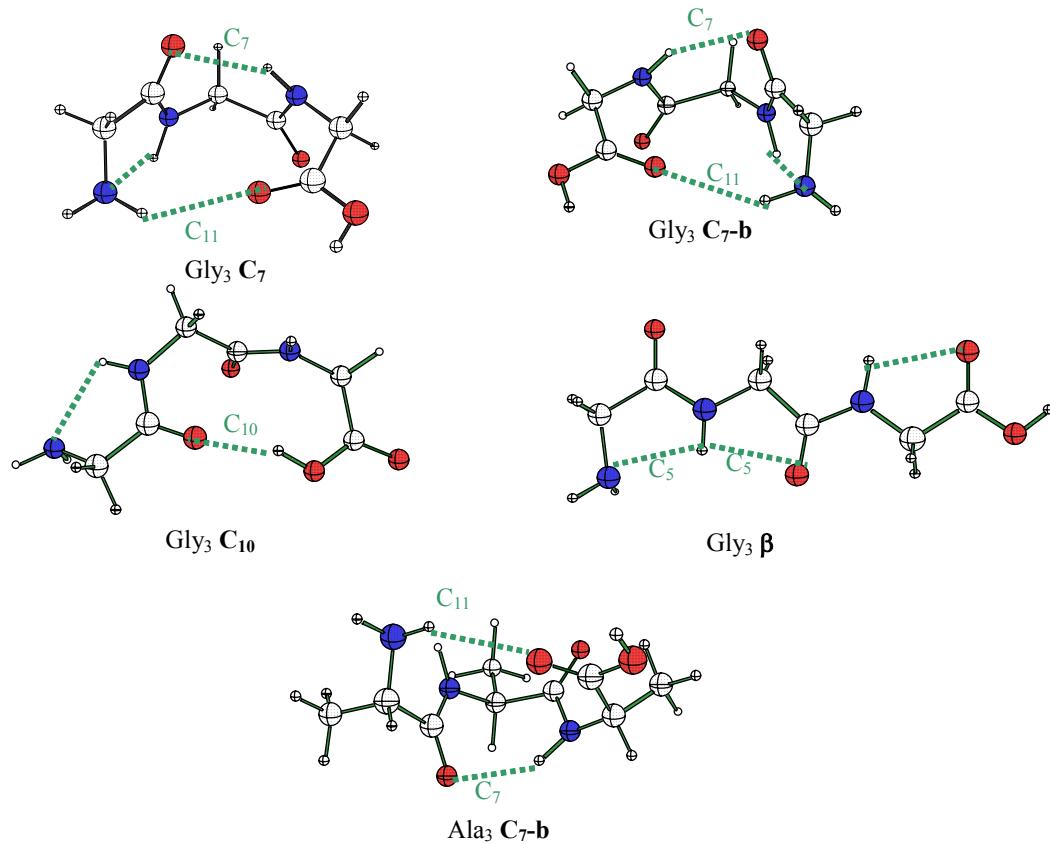
Tableau V-10 : Energies relatives (kJ/mol) des deux conformères les plus bas en énergie de  $\text{Gly}_2$  et  $\text{Ala}_2$ .

### II.B.3. Conformères de plus basse énergie de $\text{Gly}_3$ et $\text{Ala}_3$

La conformation du tri-peptide  $\text{Gly}_3$  a déjà été étudiée par le groupe de K. W. M. Siu.<sup>12</sup> Les auteurs ont identifié les trois conformères les plus bas en énergie. Leurs structures et leurs énergies relatives sont données dans la figure V-13 et le tableau V-11, respectivement. Les conformères  $\text{Gly}_3$  **C**<sub>7</sub> et **C**<sub>7-b</sub>, présentant les conformations droite et gauche, possèdent trois liaisons hydrogène intramoléculaires de type  $\text{C}_5$  ( $\text{N}^2\text{H} \rightarrow \text{N}^1$ ),  $\text{C}_7$

( $\text{N}^3\text{H} \rightarrow \text{O}^1$ ) et  $\text{C}_{11}$  ( $\text{N}^1\text{H} \rightarrow \text{O}^3$ ). Ces deux conformères sont iso-énergétiques dans le cas de  $\text{Gly}_3$  qui ne possède pas de chaîne latérale. La structure  $\text{Gly}_3 \beta$ , qui correspond à un brin  $\beta$  et met en jeu trois liaisons H de type  $\text{C}_5$  ( $\text{N}^2\text{H} \rightarrow \text{N}^1$ ), ( $\text{N}^2\text{H} \rightarrow \text{O}^2$ ) et ( $\text{N}^3\text{H} \rightarrow \text{O}^3$ ), se situe 19 kJ/mol au dessus des deux précédentes. Cependant il a été montré que la méthode MP2 a une tendance à défavoriser énergétiquement ce type de structure.<sup>13, 14</sup> Nous avons trouvé une structure intermédiaire :  $\text{Gly}_3 \text{C}_{10}$ . Cette structure possède deux liaisons hydrogène intramoléculaires, une de type  $\text{C}_{10}$  ( $\text{O}^3\text{H} \rightarrow \text{O}^1$ ) qui lie l'hydrogène de l'acide en conformation trans avec l'oxygène du premier résidu et une  $\text{C}_5$  ( $\text{N}^2\text{H} \rightarrow \text{N}^1$ ) ; elle se situe 17 kJ/mol au dessus des structures **C<sub>7</sub>** et **C<sub>7-b</sub>**.

Dans le cas de la tri-alanine, aucune étude à notre connaissance n'avait été menée précédemment sur ses conformations de plus basse énergie. Les structures équivalentes à celles de  $\text{Gly}_3$ ,  $\text{Ala}_3 \text{C}_7$  et  $\text{Ala}_3 \text{C}_{7-b}$  présentent une conformation du  $\text{C}_7$  axiale et équatoriale, respectivement, la conformation **C<sub>7-b</sub>** étant la plus basse de 5.3 kJ/mol. Quant à la structure  $\text{Ala}_3 \beta$ , elle moins stable de 10.8 kJ/mol que **C<sub>7-b</sub>**.



**Figure V-13 :** Structures de  $\text{Gly}_3$  les plus basses en énergie et structure de  $\text{Ala}_3$  la plus basse en énergie.

	$\Delta_f\text{H}^\circ(298\text{K})$ HF/PB1	$\Delta_f\text{H}^\circ(298\text{K})$ MP2/GB1// HF/PB1	$\Delta_f\text{G}^\circ(298\text{K})$ MP2/GB1// HF/PB1		$\Delta_f\text{H}^\circ(298\text{K})$ HF/PB1	$\Delta_f\text{H}^\circ(298\text{K})$ MP2/GB1// HF/PB1	$\Delta_f\text{G}^\circ(298\text{K})$ MP2/GB1// HF/PB1
Gly <sub>3</sub> <b>C<sub>7</sub>-b</b>	0	0	0	Ala <sub>3</sub> <b>C<sub>7</sub>-b</b>	0	0	0
Gly <sub>3</sub> <b>C<sub>7</sub></b>				Ala <sub>3</sub> <b>C<sub>7</sub></b>	3.9	5.3	4.7
Gly <sub>3</sub> <b>C<sub>10</sub></b>	23.5	17.0	14.5	Ala <sub>3</sub> <b>C<sub>10</sub></b>	23.8	18.7	13.7
Gly <sub>3</sub> $\beta$	6.5	18.9	9.0	Ala <sub>3</sub> $\beta$	-0.5	10.8	3.0

Tableau V-11 : Energies relatives (kJ/mol) des conformères les plus bas en énergie de Gly<sub>3</sub> et Ala<sub>3</sub>.

#### II.B.4. Conformères de plus basse énergie de Gly<sub>4</sub> et Ala<sub>4</sub>

Il n'existe pas à ce jour de littérature sur les conformations de Gly<sub>4</sub>. Seule une publication du groupe de E. R. Williams en 1999, sur l'affinité protonique des poly-glycines, donne une structure de Gly<sub>4</sub>.<sup>15</sup> Afin de déterminer la conformation la plus basse en énergie, nous avons été amenés à effectuer une recherche conformationnelle de type Monte-Carlo utilisant le champ de force AMBER, puis une ré-optimisation des structures les plus stables par calcul *ab initio* au niveau HF/PB1, complétée par un calcul d'énergie au niveau MP2/GB1//HF/PB1. Ce protocole a été développé dans le chapitre III.

Les dix structures les plus basses en énergie et la structure en brin  $\beta$  sont présentées sur la figure V-14 ; leurs énergies sont données dans le tableau V-12. On peut décomposer ces structures en deux grandes classes, l'une comportant au moins un motif C<sub>7</sub>, l'autre au moins un motif C<sub>10</sub>.

Les deux conformères de plus basse énergie de Gly<sub>4</sub> de type C<sub>7</sub>, **C<sub>7</sub>-C<sub>7</sub>-C<sub>14</sub>** et **C<sub>7</sub>-C<sub>7</sub>-C<sub>14</sub>-b**, présentent des réseaux de liaisons H très similaires, incluant deux motifs C<sub>7</sub> consécutifs (N<sup>3</sup>H→O<sup>1</sup> et N<sup>4</sup>H→O<sup>2</sup>), un motif C<sub>5</sub> (N<sup>2</sup>H→N<sup>1</sup>) et un motif C<sub>14</sub> qui fait interagir l'un des hydrogènes de l'amine terminale avec le groupement carbonyle de l'acide terminal (N<sup>1</sup>H→O<sup>4</sup>) pour **C<sub>7</sub>-C<sub>7</sub>-C<sub>14</sub>** et l'hydrogène de la fonction acide avec l'amine (O<sup>4</sup>H→N<sup>1</sup>) pour le conformère **C<sub>7</sub>-C<sub>7</sub>-C<sub>14</sub>-b**. Ce dernier est très similaire au conformère décrit comme étant le plus stable par Strittmatter et Williams,<sup>15</sup> sans autre précision. La liaison hydrogène de l'amine vers l'oxygène du carbonyle étant plus stable que celle de l'acide vers l'amine, la forme **C<sub>7</sub>-C<sub>7</sub>-C<sub>14</sub>** est favorisée d'une dizaine de kJ/mol. La structure **C<sub>7</sub>-C<sub>7</sub>-C<sub>14</sub>** est iso-énergétique d'un conformère de type C<sub>10</sub> : **C<sub>10</sub>-II-C<sub>14</sub>-C<sub>7</sub>** qui présente un motif C<sub>10</sub> de type II (N<sup>4</sup>H→O<sup>1</sup>), ces deux extrémités sont liées par un C<sub>14</sub> identique à celui de la structure **C<sub>7</sub>-C<sub>7</sub>-C<sub>14</sub>-b**.

**C<sub>14</sub>**, la fonction acide trans interagit avec l'oxygène O<sup>3</sup> (O<sup>4</sup>H→O<sup>3</sup>) par le biais d'un C<sub>7</sub> et l'amine terminale est stabilisée par un C<sub>5</sub> (N<sup>2</sup>H→N<sup>1</sup>). Le conformère **C<sub>7</sub>-C<sub>7</sub>-C<sub>14</sub>-b**, quant à lui, fait parti d'un groupe de quatre conformères très proches en énergie avec **C<sub>7</sub>-C<sub>7</sub>-C<sub>11</sub>**, **C<sub>7</sub>-C<sub>11</sub>**, et **C<sub>10</sub>-I-C<sub>7</sub>**. **C<sub>7</sub>-C<sub>7</sub>-C<sub>11</sub>** présente aussi la succession des deux motifs C<sub>7</sub>, mais le carbonyle terminal reçoit une liaison hydrogène de la première fonction amide (N<sup>2</sup>H→O<sup>4</sup>) pour former un cycle C<sub>11</sub>; quant à l'amine terminale, elle est stabilisée par interaction avec le premier carbonyle peptidique (N<sup>1</sup>H<sub>2</sub>→O<sup>1</sup>). Le conformère **C<sub>7</sub>-C<sub>11</sub>** comporte un seul motif C<sub>7</sub> faisant interagir le premier et le troisième résidu (N<sup>3</sup>H→O<sup>1</sup>) et un motif C<sub>11</sub> liant l'amine terminale avec le carbonyle du troisième résidu (N<sup>1</sup>H→O<sup>3</sup>). Il présente aussi deux interactions C<sub>5</sub> stabilisant l'amine terminale (N<sup>2</sup>H→N<sup>1</sup>) et l'acide terminal (N<sup>4</sup>H→O<sup>4</sup>). **C<sub>10</sub>-I-C<sub>7</sub>** porte une liaison C<sub>10</sub> de type I (N<sup>1</sup>H→O<sup>4</sup>), un C<sub>7</sub> (O<sup>4</sup>H→O<sup>3</sup>) identique à **C<sub>10</sub>-II-C<sub>14</sub>-C<sub>7</sub>** et un C<sub>5</sub> (N<sup>2</sup>H→N<sup>1</sup>). Les deux conformères suivants, **C<sub>10</sub>-I-C<sub>14</sub>** et **C<sub>10</sub>-II-C<sub>14</sub>** possèdent tous deux une interaction C<sub>10</sub> (N<sup>4</sup>H→O<sup>1</sup>) et un C<sub>14</sub> (N<sup>1</sup>H→O<sup>4</sup>). Ils ne diffèrent que par le type de C<sub>10</sub> I et II, respectivement. Deux conformères d'assez haute énergie, à 22 et 25 kJ/mol, **C<sub>7</sub>-C<sub>11</sub>-b** et **C<sub>7</sub>-C<sub>11</sub>-c**, présentent un seul motif C<sub>7</sub> et un motif C<sub>11</sub> (N<sup>4</sup>H→O<sup>2</sup> et N<sup>2</sup>H→O<sup>4</sup>) et (N<sup>3</sup>H→O<sup>1</sup> et N<sup>1</sup>H→O<sup>3</sup>), respectivement. Le conformère **C<sub>7</sub>-C<sub>11</sub>-c** a une structure relativement proche de celle de **C<sub>7</sub>-C<sub>11</sub>**, cependant une simple rotation du groupement CH<sub>2</sub>-NH<sub>2</sub> fait perdre le motif C<sub>5</sub> (N<sup>2</sup>H→N<sup>1</sup>), ce qui coûte 9.2 kJ/mol. Dans le cas de **C<sub>7</sub>-C<sub>11</sub>-b**, la structure est très semblable à celle de **C<sub>7</sub>-C<sub>11</sub>** excepté pour le résidu non impliqué dans le C<sub>7</sub> ou le C<sub>11</sub>. Quant à la structure en brin  $\beta$ , qui était parmi les plus stables pour Gly<sub>2</sub> et Gly<sub>3</sub>, elle se trouve à une énergie relative d'une trentaine de kJ/mol dans le cas de Gly<sub>4</sub>. En effet, les interactions de type C<sub>5</sub>, étant très éloignées de la linéarité, sont beaucoup moins stabilisantes que des interactions de type C<sub>7</sub>, C<sub>10</sub>, C<sub>11</sub> ou C<sub>14</sub>. Comme nous l'avons vu, cet effet est certainement exagéré par la méthode de calcul MP2. En effet, de nombreuses études ont montré que la méthode MP2 a tendance à défavoriser les formes linéaires comme les brins  $\beta$  au profit de structures plus compactes.<sup>13, 14</sup>

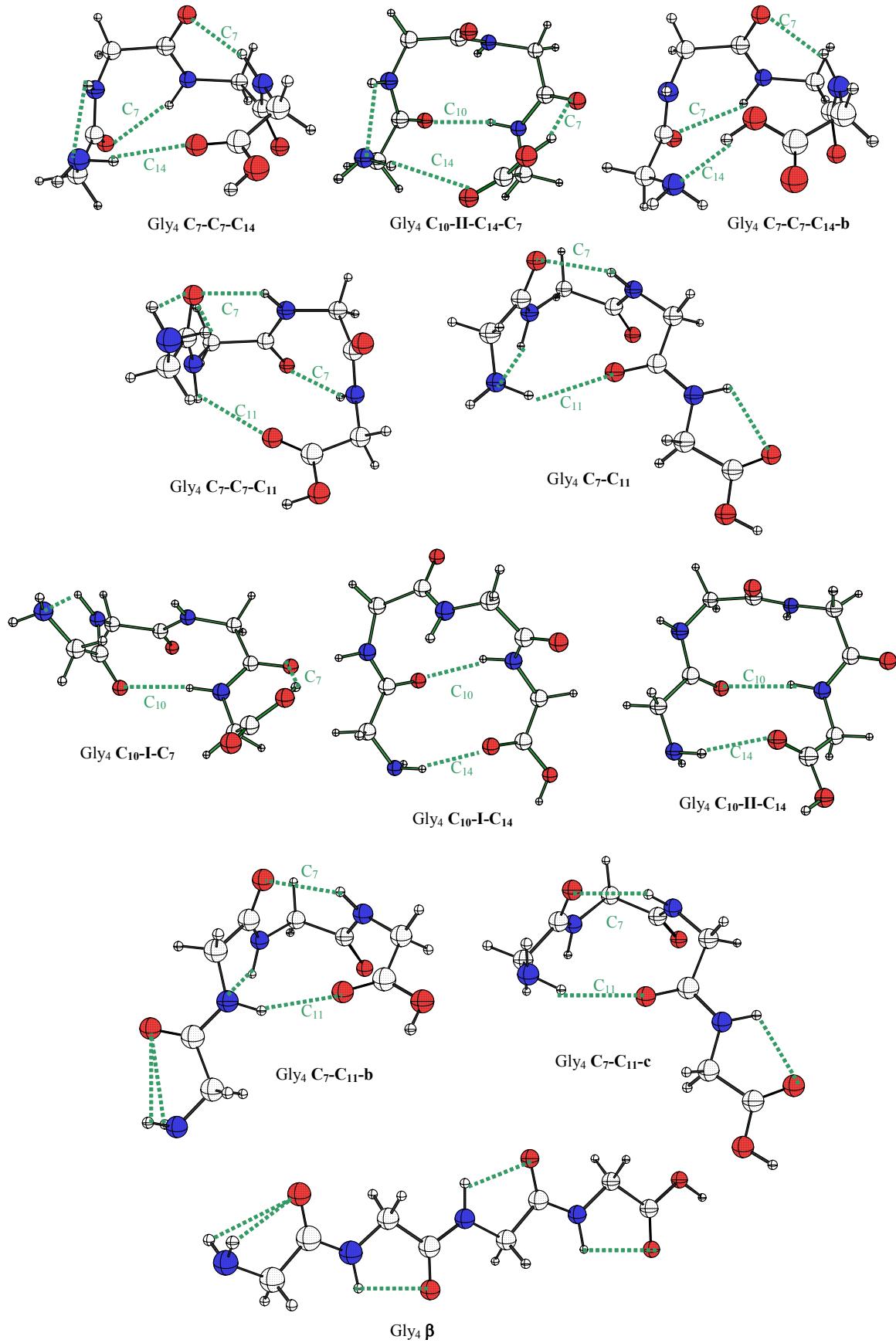


Figure V-14 : Conformères de plus basse énergie de Gly<sub>4</sub>.

Pour l'étude de  $\text{Ala}_4$ , nous avons étendu les structures obtenues dans le cas de  $\text{Gly}_4$ . Nous avons pris soin d'étudier les différents conformères axiaux et équatoriaux, puisque nous avons vu dans le cas de  $\text{Ala}_3$ , que la stabilisation du  $\text{C}_7$  équatorial était de 5 kJ/mol par rapport au motif axial. Par ailleurs on peut remarquer que dans le cas de deux motifs  $\text{C}_7$  consécutifs, il est nécessaire d'avoir un  $\text{C}_7$  axial et un  $\text{C}_7$  équatorial. Les énergies relatives des différents conformères sont données dans le tableau V-12, les énergies de chaque conformère axial et équatorial étant présentées sur deux lignes. On peut remarquer que l'ordre énergétique des différents conformères est peu perturbé par la substitution d'un hydrogène par un méthyle. Les deux conformères les plus bas sont représentés sur la figure V-15.

<b>Gly<sub>4</sub></b>	$\Delta_r\text{H}^\circ(298\text{K})$	$\Delta_r\text{H}^\circ(298\text{K})$	$\Delta_r\text{G}^\circ(298\text{K})$	<b>Ala<sub>4</sub></b>	$\Delta_r\text{H}^\circ(298\text{K})$	$\Delta_r\text{H}^\circ(298\text{K})$	$\Delta_r\text{G}^\circ(298\text{K})$
	HF/PB1	MP2/GB1/ HF/PB1	MP2/GB1/ HF/PB1		HF/PB1	MP2/GB1/ HF/PB1	MP2/GB1/ HF/PB1
<b>C<sub>7</sub>-C<sub>7</sub>-C<sub>14</sub></b>	0	<b>0</b>	0	<b>C<sub>7</sub>-C<sub>7</sub>-C<sub>14</sub></b>	0	<b>0</b>	0
					10.5	8.8	11.2
<b>C<sub>10</sub>-II-C<sub>14</sub>-C<sub>7</sub></b>	13.8	1.2	4.6	<b>C<sub>10</sub>-II</b>	13.8	<b>0.3</b>	5.1
					16.4	7.4	8.7
<b>C<sub>7</sub>-C<sub>7</sub>-C<sub>14</sub>-b</b>	16.8	10.9	10.7	<b>C<sub>7</sub>-C<sub>7</sub>-C<sub>14</sub>-b</b>	15.8	10.4	10.6
					13.9	11.5	10.3
<b>C<sub>7</sub>-C<sub>7</sub>-C<sub>11</sub></b>	9.1	11.3	7.3	<b>C<sub>7</sub>-C<sub>7</sub>-C<sub>11</sub></b>	19.0	16.3	15.4
					7.7	10.3	9.4
<b>C<sub>7</sub>-C<sub>11</sub></b>	7.2	13.1	8.5	<b>C<sub>7</sub>-C<sub>11</sub></b>	10.2	15.7	12.8
					16.9	12.6	8.1
<b>C<sub>10</sub>-I-C<sub>7</sub></b>	19.5	13.7	11.0	<b>C<sub>10</sub>-I</b>	23.1	13.9	13.5
					18.4	23.7	15.7
<b>C<sub>10</sub>-I-C<sub>14</sub></b>	13.8	17.3	13.6	<b>C<sub>10</sub>-I-C<sub>14</sub></b>			
<b>C<sub>10</sub>-II-C<sub>14</sub></b>	18.4	21.9	14.8	<b>C<sub>10</sub>-II-C<sub>14</sub></b>			
<b>C<sub>7</sub>-C<sub>11</sub>-b</b>	14.8	22.2	15.6	<b>C<sub>7</sub>-C<sub>11</sub>-b</b>	27.7	31.6	28.6
					18.4	26.2	20.2
<b>C<sub>7</sub>-C<sub>11</sub>-c</b>	16.2	25.2	17.7	<b>C<sub>7</sub>-C<sub>11</sub>-c</b>	18.7	24.8	13.0
<b><math>\beta</math>-b</b>	11.9	33.0	17.0	<b><math>\beta</math>-b</b>	5.2		

**Tableau V-12 :** Energies relatives (kJ/mol) des six conformères les plus bas en énergie de  $\text{Gly}_4$  et  $\text{Ala}_4$  et du brin  $\beta$  (les deux colonnes pour les énergies de  $\text{Ala}_4$  correspondent aux énergies de chacun des conformères droite et gauche).

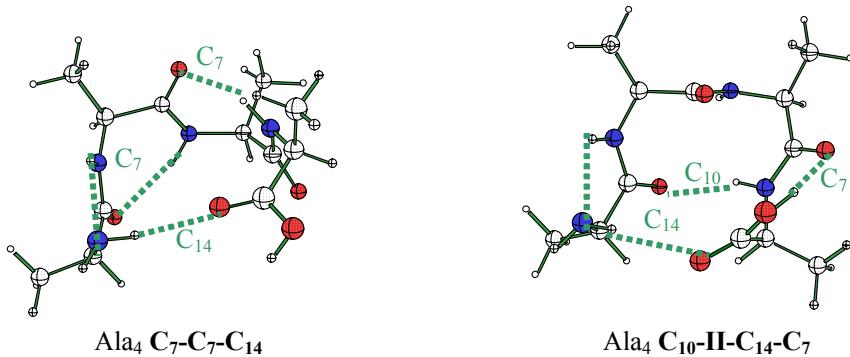


Figure V-15 : les deux conformères de plus basse énergie de  $\text{Ala}_4$ .

### II.C. Comparaison expériences-théorie

Les valeurs théoriques obtenues au niveau MP2/GB1//HF/PB1 et expérimentales déterminées par la méthode cinétique de Cooks, des affinités pour le sodium des peptides évoqués précédemment sont présentées dans le tableau V-13. Les valeurs expérimentales et théoriques sont en très bon accord, puisque les déviations varient de 1 à 8 kJ/mol, avec 4 kJ/mol en moyenne. On peut remarquer que cette déviation va toujours dans le même sens : les valeurs théoriques étant inférieures aux valeurs expérimentales. Un meilleur accord serait obtenu pour  $\text{Gly}_3$  en maintenant une température effective de 362 K jusqu'à ce peptide.

Pour les cas de GlyHis et HisGly, une étude détaillée des structures des complexes sodiés et des peptides isolés fera l'objet de la partie III de ce chapitre. L'histidine existant sous la forme de deux tautomères  $\text{N}^\delta\text{-H}$  et  $\text{N}^\varepsilon\text{-H}$ , nous avons fait l'hypothèse, pour déterminer leur énergie de liaison avec  $\text{Na}^+$ , que la dissociation s'effectue sans étape d'isomérisation. Nous avons donc considéré que la forme de l'histidine était la même dans le complexe et dans la molécule isolée. On observe un très bon accord entre les valeurs expérimentales et les valeurs théoriques uniquement dans le cas des tautomères  $\text{N}^\varepsilon\text{-H}$ , ce qui indique que seul ce tautomère a été formé en ESI. Ce résultat expérimental est cohérent avec les calculs, puisque nous avons déterminé que les complexes  $\text{Gly-His}(\text{N}^\varepsilon\text{-H})\text{-Na}^+$  et  $\text{His}(\text{N}^\varepsilon\text{-H})\text{-Gly-Na}^+$  sont plus stables d'une quarantaine de kJ/mol par rapport à leurs homologues  $\text{N}^\delta\text{-H}$ .

Molécule	$\Delta_r\text{H}_{\text{Na}}$	$\Delta_r\text{H}_{\text{Na}}$	$\Delta_r\text{H}_{\text{Na}}$ Calc	Nom du complexe Nom du peptide isolé
	362 K Exp	435 K Exp		
<b>Pro</b>	<b>196</b>	194		
<b>Thr</b>	<b>197</b>	196		
<b>Phe</b>	<b>198</b>	196		
<b>GlyGly</b>	<b>203</b>	203	199	Gly <sub>2</sub> -Na <sup>+</sup> <b>I</b> Gly <sub>2</sub> <b>II</b>
<b>AlaGly</b>	<b>205</b>	205		
<b>GlyAla</b>	<b>207</b>	207		
<b>AlaAla</b>	<b>208</b>	209	200	Ala <sub>2</sub> -Na <sup>+</sup> <b>I</b> Ala <sub>2</sub> <b>II</b>
<b>Trp</b>	<b>210</b>	211		
<b>GlyLeu</b>	<b>211</b>	213		
<b>GlyPhe</b>	215	<b>217</b>		
<b>SerGly</b>	217	<b>219</b>		
<b>PheGly</b>	222	<b>225</b>		
<b>His</b>	227	<b>232</b>		
<b>AlaTrp</b>	229	<b>234</b>		
<b>TrpAla</b>	232	<b>237</b>		
<b>GlyGlyGly</b>	238	<b>245</b>	237	Gly <sub>3</sub> -Na <sup>+</sup> <b>4sON</b> Gly <sub>3</sub> <b>C</b>
<b>AlaGlyGly</b>	239	<b>246</b>		
<b>GlyHis</b>	240	<b>248</b>	N <sup>δ</sup> -H 201	I I-1-ad
			N <sup>ε</sup> -H 245	I I-1-jb
<b>AlaAlaGly</b>	241	<b>248</b>		
<b>Arg</b>	242	<b>250</b>		
<b>AlaAlaAla</b>	243	<b>251</b>	247	Ala <sub>3</sub> Na <sup>+</sup> 3s2 Ala <sub>3</sub> C <sub>7</sub> -b
<b>LeuGlyPhe</b>	248	<b>257</b>		
<b>HisGly</b>	251	<b>261</b>	N <sup>δ</sup> -H 219	I I-1-h
			N <sup>ε</sup> -H 260	I I-1-a
<b>GlyHisGly</b>	259	<b>270</b>		
<b>GlyGlyGlyGly</b>	261	<b>273</b>	272	Gly <sub>4</sub> -Na <sup>+</sup> <b>4s</b> Gly <sub>4</sub> C <sub>7</sub> -C <sub>7</sub> -C <sub>14</sub>
<b>AlaAlaAlaAla</b>	265	<b>277</b>	273	Ala <sub>4</sub> -Na <sup>+</sup> <b>3sboucle</b> Ala <sub>4</sub> C <sub>7</sub> -C <sub>7</sub> -C <sub>14</sub>

**Tableau V-13 :** Comparaison des affinités pour le sodium (kJ/mol) obtenues expérimentalement par la méthode de Cooks, et théoriquement au niveau MP2/GB1//HF/PB1.

### III.Cas des tautomères GlyHis et HisGly sodiés et isolés

#### III.A. Résumé

Le but de cette partie est la détermination théorique de l'énergie de liaison des di-peptides GlyHis et HisGly avec le cation  $\text{Na}^+$ . En effet les résultats expérimentaux montrent un effet important de la séquence sur l'énergie de liaison : l'affinité pour le sodium de HisGly est de 14 kJ/mol plus importante que celle de GlyHis. Nous nous sommes donc intéressés à la structure des conformères les plus bas en énergie pour GlyHis et HisGly ainsi que leurs adduits sodiés. Cette étude est plus complexe que dans le cas des dipeptides  $\text{Gly}_2$  et  $\text{Ala}_2$ . En effet l'acide aminé His possède une chaîne latérale imidazole qui existe sous les deux formes tautomères  $\text{N}^\delta\text{-H}$  et  $\text{N}^\epsilon\text{-H}$ , et qui peut interagir avec  $\text{Na}^+$ . La tautométrie de His est une question de portée générale en biologie structurale.

Nous avons trouvé que dans le cas des complexes avec  $\text{Na}^+$ , les peptides se présentent sous la forme des tautomères  $\text{N}^\epsilon\text{-H}$ , avec une fonction acide cis, les formes  $\text{N}^\delta\text{-H}$  étant une quarantaine de kJ/mol moins stables. Dans le cas des peptides isolés, les deux types de structures ( $\text{N}^\epsilon\text{-H}$  et  $\text{N}^\delta\text{-H}$ ) ont des énergies semblables (différence inférieure à 3.8 kJ/mol).

#### III.B. Article

**The low energy tautomers and conformers of the dipeptides  
HisGly and GlyHis and of their sodium ion complexes in the gas phase**

**Catherine Kapota and Gilles Ohanessian**

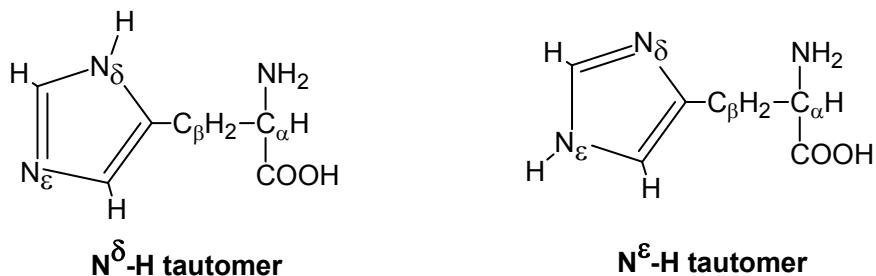
Laboratoire des Mécanismes Réactionnels, UMR CNRS 7651,  
Ecole Polytechnique, 91128 Palaiseau CEDEX, France

**Abstract**

*The low-lying conformers of the dipeptides HisGly and GlyHis, and of their sodium cation complexes, have been studied with a combination of Monte Carlo search with the Amber force field and local geometry optimization at the *ab initio* HF/6-31G(d) level, completed with MP2(full)/6-311+G(2d,2p) energetics at the HF/6-31G(d) geometries. For each dipeptide, both the  $N^\delta$ -H and  $N^\epsilon$ -H tautomers of the imidazole side chain of His were considered. For each of the four isomeric dipeptides, 20-30 conformers were fully characterized at the *ab initio* level. All low energy structures are found to involve H-bonding at the  $N_\delta$  position of imidazole, either as a N-H donor or a N acceptor, depending upon the tautomer. In three out of the four species, the most stable conformer involves a C-terminus carboxylic acid in its less favorable trans conformation, in order to maximize intramolecular H bonding. It turns out that the lowest energy tautomer of HisGly is  $N^\epsilon$ -H, while that of GlyHis is  $N^\delta$ -H. This result argues in favor of the diversity of His tautomeric states in peptides and proteins. The sodium cation complexes of both GlyHis and HisGly have been studied as well, again considering both tautomers in each case. In three out of the four species, the most stable structure involves chelation of sodium by the two carbonyl oxygens and the imidazole ring. On the contrary, the sodium complex of the  $N^\delta$ -H tautomer of HisGly favors chelation to the peptidic carbonyl oxygen, the imidazole ring and the amino terminus. In the  $N^\epsilon$ -H tautomers of both peptides, the most favorable binding site of imidazole is the  $N_\delta$  nitrogen, while in the  $N^\delta$ -H tautomers, it is the  $\pi$  cloud which provides side chain interaction. As a result, both GlyHis $\text{Na}^+$  and HisGly $\text{Na}^+$  favor the  $N^\epsilon$ -H tautomer of His, in contrast to what was found for the free peptides.*

## Introduction

Among all naturally occurring  $\alpha$ -amino acids in living systems, histidine (His) occupies a special position. The side chain of His, methylene-imidazole, bears two nitrogen atoms with variable protonation states. Their pK's are such that imidazole may easily exchange protons with its surroundings in the biologically significant range of pH, effectively supporting many chemical reactions. Because of this versatility, histidine has been found to be involved in ca. 50% of the active sites of all enzymes, in a recent analysis of the enzyme structural database.<sup>1</sup> The Lewis basicity of His also makes it a common ligand in the first metal coordination sphere of metallo-proteins. It is even considered that coordinated His may exist in its fully deprotonated form, at least as a transient species.<sup>2</sup> Therefore the protonation state of His in peptides and proteins is a matter of great interest.



**Scheme 1 :**

In the pH range 6.3-9.0 in aqueous solution at room temperature, isolated His exists as a zwitterion with a non-protonated side chain imidazole. In this and higher pH range, the imidazole is in either of two tautomeric forms,  $\text{N}^\delta\text{-H}$  and  $\text{N}^\epsilon\text{-H}$ . Scheme 1 shows the two  $\text{N}^\delta\text{-H}$  and  $\text{N}^\epsilon\text{-H}$  tautomers in the non-zwitterionic form. Protonation of either form at the imino nitrogen leads to an imidazolium ion from which proton release may occur from either of the two nitrogens, leading to each of the two tautomers. In His-containing peptides and proteins, it is therefore expected that the occurrence of the two tautomers is strongly influenced by the presence of H bond donors and acceptors in the surroundings, and by the chemical reactions which occur. Determining the tautomer populations in proteins usually relies on NMR measurements. One-dimensional measurements on various nuclei suffer from various drawbacks, although  $^{15}\text{N}$  spectra have been shown to yield unambiguous information on His tautomers in  $^{15}\text{N}$ -enriched proteins, and in isolated His at various pH and temperatures, in several solvents.<sup>3</sup> More recently, several techniques have been devised to identify the protonation state and tautomeric forms of His, including N-C J couplings<sup>1,4</sup> and 2D  $^1\text{H}/^{13}\text{C}_\delta$  correlations.<sup>5</sup> It remains the case, however, that the protonation and tautomeric states of most His residues in most crystallographic structures in protein databases are not unequivocally assigned. In such cases, hydrogen locations are assigned, either in the

structure determination and refinement procedure, or in post-structural analysis, using tools whose reliability is not known in general. A recent computational and statistical study<sup>6</sup> has indeed issued a warning signal about the significance of proton positions in His residues in protein databases, whenever specific NMR data is not available. This study, based on local energy minimization of His side chain torsions using the Amber force field, showed that proton assignments appear to be no better than random. Thus it is currently difficult to extract from structural databases whether there is a strong preference for one of the tautomers of His in proteins, even with the restriction to the interior of proteins, where the solvent is not expected to have a strong influence. One result of the computational study is that there appears to be no general preference for one tautomer over the other. In order to understand in details the factors determining the relative stabilities of His tautomers, it is therefore of interest to study model systems.

The situation is simple in isolated His, for which  $^{15}\text{N}$  NMR, as well as other techniques, have established that the  $\text{N}^\varepsilon\text{-H}$  tautomer is largely dominant.<sup>3</sup> This dominance has been discussed in terms of the intramolecular hydrogen bond which may establish between the  $\text{N}_\delta$  and one or two of the N-H bond(s) of the ammonium group (note that in the following, we use the notation  $\text{N}_\delta$  and  $\text{N}_\varepsilon$ , with  $\delta$  and  $\varepsilon$  as subscripts, for the imidazole nitrogen atom at the  $\delta$  and  $\varepsilon$  position, respectively, so that  $\text{N}_\delta\text{-H}$  and  $\text{N}_\varepsilon\text{-H}$  denote the nitrogen-hydrogen bonds at these positions. On the other hand,  $\text{N}^\delta\text{-H}$  and  $\text{N}^\varepsilon\text{-H}$ , with  $\delta$  and  $\varepsilon$  as superscripts, are used to describe the tautomers with the N-H bond at the  $\delta$  and  $\varepsilon$  position, respectively). The same type of argument was found to explain the structure of the dipeptide HisGly, for which several crystal structures have been determined.<sup>7</sup> In the crystals of both the chloride<sup>7a</sup> and dichloride<sup>7b</sup> salts of HisGly (with one and two chloride ions, and singly and doubly protonated HisGly, respectively), there appears to be a strong hydrogen bond between the ammonium group and the imine  $\text{N}_\delta$  of imidazole. The crystal structure of the hemihydrate of HisGly has also been obtained,<sup>7c</sup> in which the His side chain is not protonated ; the structural parameters are consistent with the existence of a zwitterion, involving a N-terminus ammonium and a C-terminus carboxylate. Rather surprisingly, it was found that the crystal structure contains two peptide molecules per asymmetric crystal unit, in the two different tautomeric forms. One is the expected  $\text{N}^\varepsilon\text{-H}$  tautomer while the second is the  $\text{N}^\delta\text{-H}$  tautomer, involving H-bonding between  $\text{N}_\delta\text{-H}$  and the C terminal carboxylate. Thus it appears that HisGly is a simple model in which the coexistence and possible competition between the two tautomers of His already exist.

For this reason, we have explored the conformational landscape of HisGly. In order to delineate in details the interactions which influence the relative stabilities of the tautomers, we have also studied its dipeptide isomer GlyHis, for which there is no structural information

available, to the best of our knowledge. Clearly such simple models lack the side chain - side chain interactions in which His imidazoles are often involved in proteins. Yet local interactions within small dipeptides are found to have a significant influence on the relative stabilities of the tautomers. We have also studied the sodium cation complexes of HisGly and GlyHis, since in a recent study of the  $\text{Na}^+$  complexes of a series of small peptides, we found that there exists a significant sequence effect on the binding affinities of HisGly and GlyHis to  $\text{Na}^+$ .<sup>8</sup> In the present context, it is interesting to see if cation attachment may change the relative energies of the tautomers.

The detailed understanding of the conformation landscape of small biomolecules is a challenging task, both because there is generally a large to very large number of low-lying energy minima, and because energy barriers connecting these minima are often small, leading to several conformers being populated at room temperature. Various experimental methods are capable of yielding some structural data on such molecules, yet it remains very difficult to identify the lowest energy conformers with certainty, and to obtain an overview of the potential energy surface. Quantum chemistry is an efficient alternative to experimental methods to explore the conformations of flexible molecules of relatively small size, as it provides relative energies of conformers with good accuracy, and enables a thorough exploration of the potential energy surface. There has been ample work on the conformations of Gly in the 1990's,<sup>9</sup> and other amino acids have been also studied, including alanine,<sup>10</sup> serine, cysteine,<sup>11</sup> valine,<sup>12</sup> proline,<sup>13</sup> glutamic acid<sup>14</sup> and arginine.<sup>15</sup> Several oligopeptides have also been studied.<sup>16</sup> Herein we use ab initio computations to establish the structures of the low energy conformations of the dipeptides GlyHis and HisGly in each of their possible tautomeric forms. Analogous work has also been carried out for the sodium cation complexes of the four isomers.

### Computational methods.

The number of weakly hindered rotors (seven covalent single bonds between non-hydrogen atoms) is large enough in GlyHis and HisGly that constructing structures based on chemical intuition solely is not appropriate. Thus we resorted to a non local exploration of the potential energy surface, as a preliminary step prior to local geometry optimization with the ab initio methods mentioned below. Monte Carlo sampling was carried out, in which random values were generated for the torsion angles around all single bonds between non-hydrogen atoms except the peptide bond. In order to keep this procedure computationally tractable, the Amber 94 force field was used for energy calculations, with RESP atomic charges.<sup>17</sup> Random

sampling was systematically followed by local geometry optimization at the same level. However the Amber calculations lead to an energy ordering of conformers which deviates significantly from that obtained with accurate ab initio calculations. In order to locate all of the low lying structures, a large sample was selected from the Amber results and subjected to geometry re-optimization at the ab initio level, typically between 20 and 30 in each case (each tautomeric form of GlyHis and HisGly). While it is likely that the several most stable conformers have been identified for each species, the present work does not aim at providing an exhaustive description of these very complex potential energy surfaces.

Ab initio calculations were carried out at levels which represent a compromise between accuracy and tractability for consideration of a significant number of structures. Geometries were optimized at the HF/6-31G(d) level, vibrational analyses were carried out at the same level to determine zero-point vibrational energies, thermal corrections to total energies, and entropies. Final energetics were determined at the MP2(full)/6-311+G(2d,2p) level using the HF/6-31G(d) geometries, a level of computation which has been previously shown to yield accurate energetics.<sup>18</sup> All results mentioned below are computed at this MP2(full)/6-311+G(2d,2p)//HF/6-31G(d) level except otherwise noted. Relative energies computed at the HF and MP2 levels were usually found to be in satisfactory agreement with each other (differing by less than 10 kJ.mol<sup>-1</sup>, and often by less than 5 kJ.mol<sup>-1</sup>). It turned out, however, that in some cases the differences were as large as 10-25 kJ.mol<sup>-1</sup>. A careful inspection showed that such large differences occur when the conformers being compared have the C-terminal carboxylic acid in different conformations (cis vs trans). Test calculations on acetic acid itself indicate that the cis-trans relative energy is 7 kJ.mol<sup>-1</sup> lower at the MP2(full)/6-311+G(2d,2p)//HF/6-31G(d) level, compared to the HF/6-31G(d) result. Moreover, low energy structures always bear a trans carboxylic acid that is H-bond donating, and this H bond is not very accurately described at the HF level. Test calculations at the MP2/6-311+G(2d,2p)//MP2/6-31G(d) level confirm this result, leading to relative energies in reasonable agreement with the MP2(full)/6-311+G(2d,2p)//HF/6-31G(d) values. Therefore the latter level was judged sufficiently accurate for the purpose of the present work. Only for the matter of establishing which tautomer is the most stable for HisGly and GlyHis, were MP2/6-311+G(2d,2p)//MP2/6-31G(d) calculations carried out.

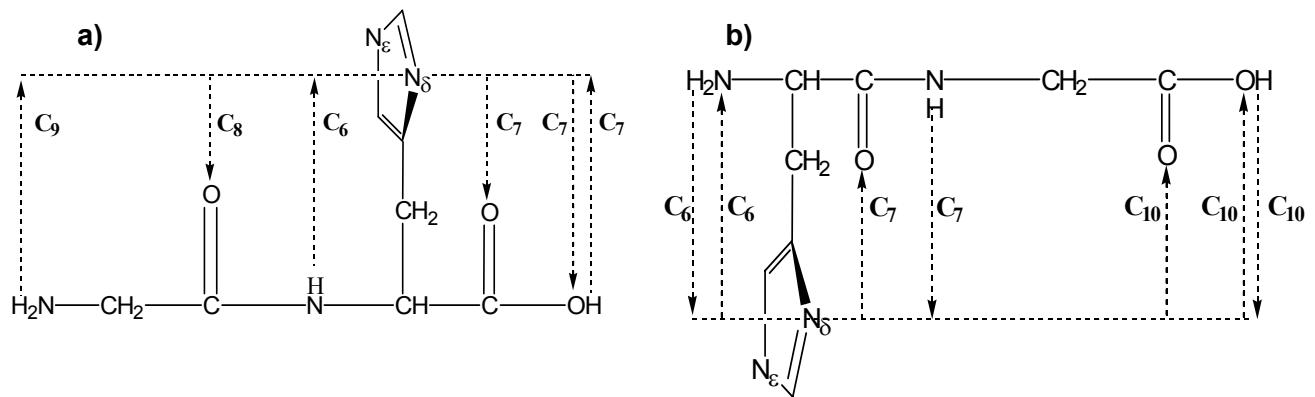
Monte Carlo calculations were carried out with HyperChem 6.0,<sup>19</sup> while the Gaussian03 package was used for ab initio calculations.<sup>20</sup>

**Results and discussion.**

*Conformer descriptors*

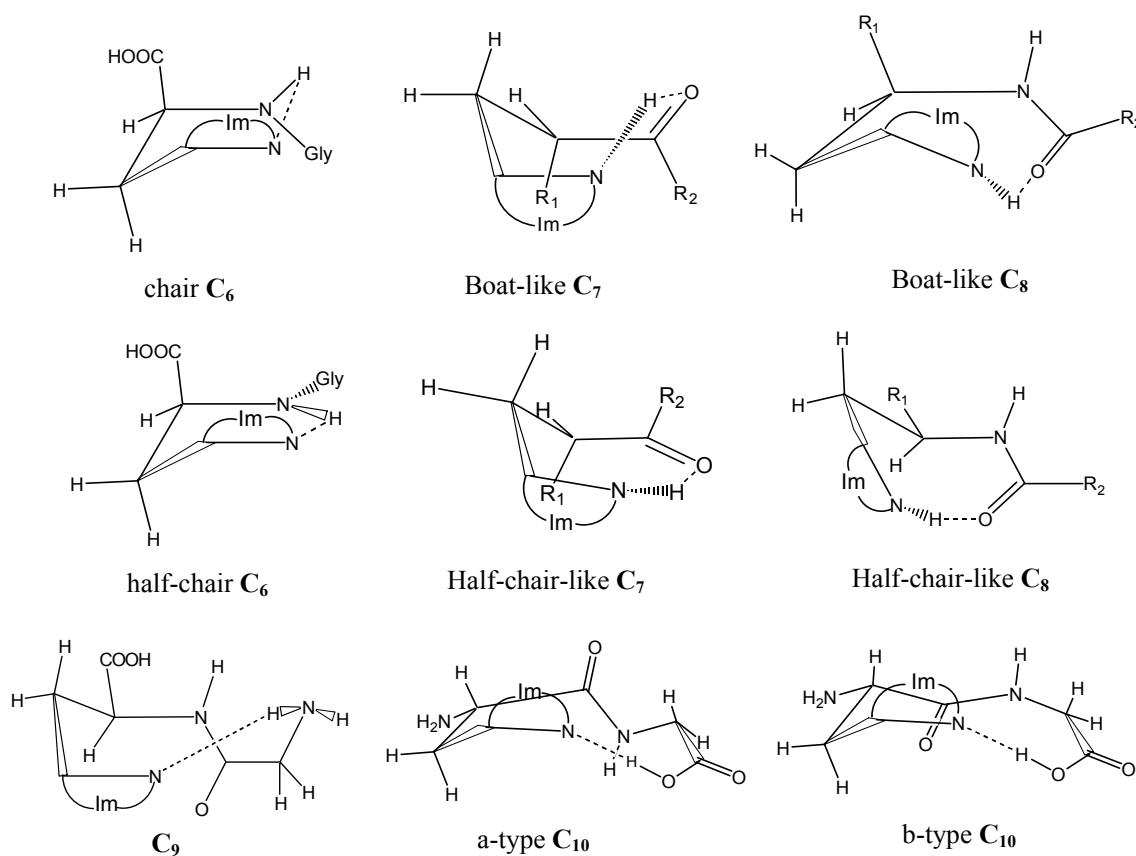
Several different ways are conceivable for the description of conformers, for each of the tautomers of HisGly and GlyHis. This is due to the presence of seven weakly hindered rotors in each species : the  $\text{C}_\alpha$ - $\text{C}(=\text{O})$  and  $\text{C}_\alpha$ - $\text{N}$  of the main chain of each residue, the  $\text{C}_\alpha$ - $\text{C}_\beta$  and  $\text{C}_\beta$ - $\text{C}_{\text{ring}}$  of the His side chain, and the  $\text{C}$ - $\text{OH}$  bond of the C-terminus acid. Describing each of the conformers by the value of the torsion angle around each of these bonds would be rather tedious, although it would carry all of the information. We have chosen to define families using a hierarchy of structural criteria (descriptors), of which several are non local.

The first descriptor is the existence of a hydrogen bond between an atom of the main chain and the  $\text{N}_\delta$  or  $\text{N}_\delta$ - $\text{H}$  of the imidazole ring. All of the low energy conformers determined in this work bear such a hydrogen bond ; for instance for the  $\text{N}^\epsilon$ - $\text{H}$  tautomer of GlyHis, the lowest lying conformer found without a hydrogen bond at  $\text{N}_\delta$  is higher in energy than the most stable by 18  $\text{kJ}\cdot\text{mol}^{-1}$ . We denote the cyclic motif formed by such H bonds as “ $\text{C}_n$ ”, meaning that the H bond generates a  $n$ -membered cycle. The various types of  $\text{C}_n$  structures,  $\text{C}_6$  –  $\text{C}_{10}$ , are depicted schematically in Scheme 2. The arrows in Scheme 2 are oriented from the hydrogen bond donor to the acceptor. For instance in the  $\text{N}^\delta$ - $\text{H}$  tautomers of HisGly and GlyHis, the  $\text{N}_\delta$ - $\text{H}$  bond can be a H-donor towards the carbonyl oxygen of His, generating a  $\text{C}_7$  motif.  $\text{C}_7$  motifs also exist for the  $\text{N}^\epsilon$ - $\text{H}$  tautomers, in which the  $\text{N}_\delta$  atom is a H-bond acceptor from an  $\text{N}$ - $\text{H}$  bond of Gly in HisGly or the C-terminus OH bond in GlyHis. On the other hand, in none of the conformers is the  $\text{N}_\epsilon$  or  $\text{N}_\epsilon$ - $\text{H}$  oriented in such a way as to engage in a H bond of any  $\text{C}_n$  type, because the dipeptide chain is too short.



**Scheme 2 :** all types of  $\text{C}_n$  for a) GlyHis, b) HisGly

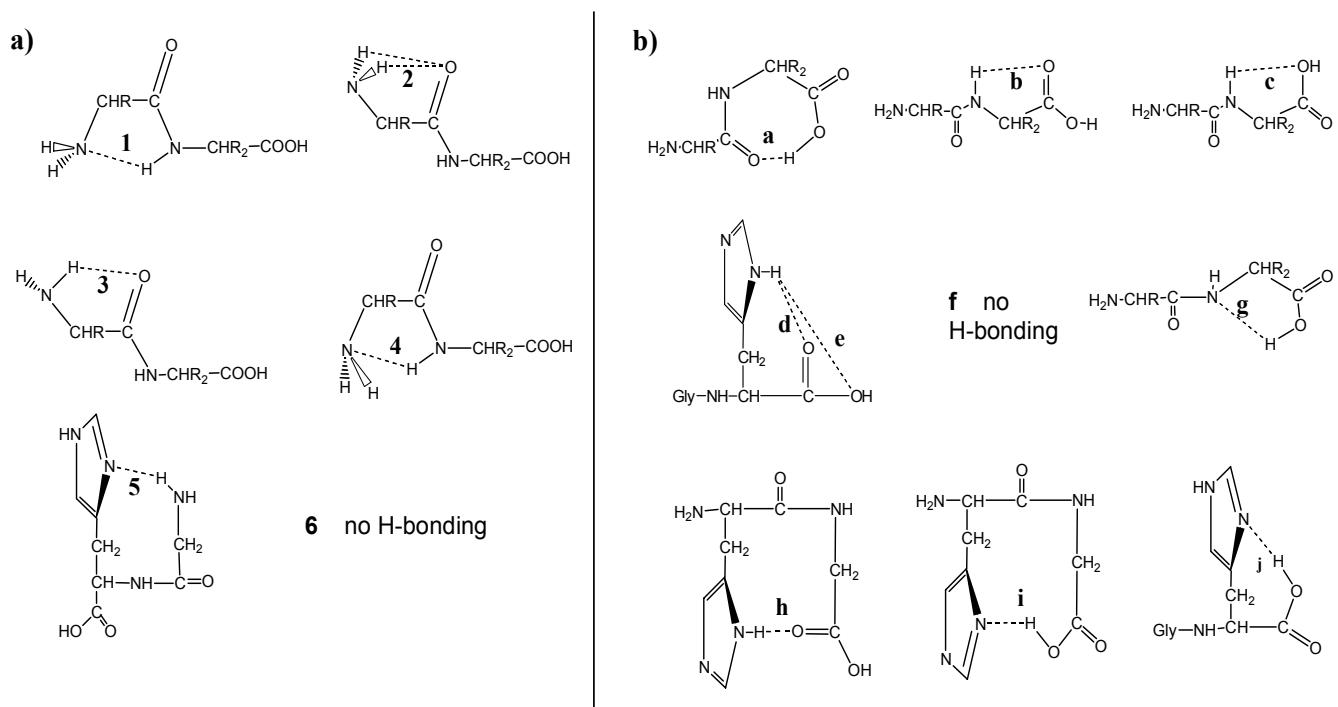
The second conformation descriptor is the conformation of the  $\text{C}_n$  ring. The types of ring conformations encountered in low energy structures are presented in more details in Figure 1.  $\text{C}_6$  rings connect the  $\text{N}_\delta$  position to the main chain nitrogen of His, which is the peptidic nitrogen in GlyHis and the N terminus in HisGly. In both  $\text{N}^\epsilon\text{-H}$  tautomers, there are two possible conformations, chair and half-chair (see Figure 1). Note that the half-chair may be inverted, leading to a different energy since steric repulsions with the rest of the molecule are different. In the  $\text{N}^\delta\text{-H}$  tautomer of HisGly, a chair conformation is formed by H bonding from the  $\text{N}_\delta\text{-H}$  to the amino terminus. For GlyHis, the H bond would point to the peptidic nitrogen, but the rigidity of the peptide linkage precludes formation of the  $\text{C}_6$  ring in this case. Other  $\text{C}_n$  rings may be formed in several isomers, but the  $\text{C}_7$  is the only one which occurs in all. The various  $\text{C}_7$  possibilities have been introduced above. We find two conformations for the  $\text{C}_7$ , half-chair-like and boat-like. Hydrogen bonding is favored when the  $\text{N}\text{-H}\cdots\text{O}=\text{C}$  segment is nearly planar in both types of conformation, since it allows the  $\text{N}\text{-H}$  bond to point approximately towards an oxygen lone pair, and the two bond dipoles to be oriented favourably for electrostatic stabilization. This restricts significantly the flexibility of the  $\text{C}_7$  motif. As shown in Figure 1, formation of a  $\text{C}_8$  ring is possible only for the  $\text{N}^\delta\text{-H}$  tautomer of GlyHis, connecting the  $\text{N}_\delta\text{-H}$  bond to the carbonyl oxygen of Gly. We found two conformations for this  $\text{C}_8$  which may be loosely defined as half-chair-like and boat-like. As for  $\text{C}_8$ , a  $\text{C}_9$  ring is compatible with one isomer only; in this case it is the  $\text{N}^\epsilon\text{-H}$  tautomer of GlyHis, in which the H bond connects  $\text{N}_\delta$  to one of the N-H bonds of the amino terminus. The  $\text{C}_9$  ring is associated in some cases with a  $\text{C}_6$  ring, when the main chain is oriented in a way which also permits interaction of  $\text{N}_\delta$  with the peptidic N-H bond. Finally,  $\text{C}_{10}$  rings exist in both tautomers of HisGly. In the  $\text{N}^\delta\text{-H}$  tautomer it is due to a H bond between the  $\text{N}_\delta\text{-H}$  bond and the oxygens at the C terminus, while in the  $\text{N}^\epsilon\text{-H}$  tautomer, the H bond is between the C terminus O-H bond of the trans conformation of carboxylic acid and  $\text{N}_\delta$ . In both cases, we find two conformers, which are deduced from each other by a  $180^\circ$  rotation around the  $\text{C}_\alpha\text{-N}$  bond of Gly (see Figure 1). This leads to opposite orientations of the  $\text{O}=\text{C}\text{-N}\text{-H}$  peptidic group with respect to the mean plane of the  $\text{C}_{10}$  cycle.



**Figure 1 :** Qualitative description of all types of  $\text{C}_n$  rings and their main conformations. The imidazole ring is schematized with “Im” for clarity of  $\text{C}_n$  conformation, explicitly showing the  $\text{C}_\gamma$  and  $\text{N}_\delta$  positions only.

The third descriptor is the relative orientations of the peptidic  $\text{O}=\text{C}-\text{N}-\text{H}$  plane with respect to that of the C terminal carboxylic acid. There are two rough relative orientations : either coplanar, defining a fragment of a  $\beta$  sheet, or perpendicular. Finally, the fourth descriptor is the relative position of imidazole and of the peptidic plane. “Open” conformers correspond to extended structures, in which the Gly residue and imidazole are distant, i.e. they cannot interact via H bonding. “Closed” conformers are such that such H bonds can be established, between the  $\text{N}_\delta$  position of His and an atom of either the peptidic bond or of the C terminus. Altogether, each existing combination of these four descriptors defines a “family”. We have noted these families with capital roman numbers, in the stability order. For instance for the  $\text{N}^\delta\text{-H}$  tautomer of GlyHis, the lowest energy conformer belongs to family “I”, which corresponds to the existence of a  $\text{C}_7$  ring, in a boat-like conformation, in an “closed” relative orientation of Gly and imidazole, and a perpendicular relative orientation of the peptidic and carboxylic planes. Within each family, there remain conformational differences. These are specified by the orientation of the terminal  $\text{NH}_2$  and  $\text{COOH}$  groups. The various possibilities encountered in low energy structures are gathered in Figure 2, together with the numbering

used hereafter. For instance in conformations “1” the  $\text{NH}_2$  terminus is a H bond acceptor towards the peptidic N-H, while in “2” it is a H bond donor towards the peptidic oxygen, etc. In conformations “a” the carboxylic C terminus is in its trans conformation and it is a H bond donor towards the peptidic oxygen, while in “b” it is cis, and a H bond acceptor from the peptidic N-H. There is some redundancy between these notations and the specification of the  $\text{C}_n$  rings, since, e.g. conformations 5 are  $\text{C}_9$  rings (see Figure 2).



**Figure 2 :** Qualitative description of all orientations **a)** of the  $\text{NH}_2$  terminus (left part) and **b)** of the COOH terminus (right part).

For all isomers, the most stable conformer of each family is depicted in Figures 3 and 5-7. In order to illustrate the conformational flexibility offered by the  $\text{NH}_2$  and COOH termini, the six most stable conformations of the most stable family of the  $\text{N}^\delta\text{-H}$  tautomer of GlyHis are shown in Figure 4.

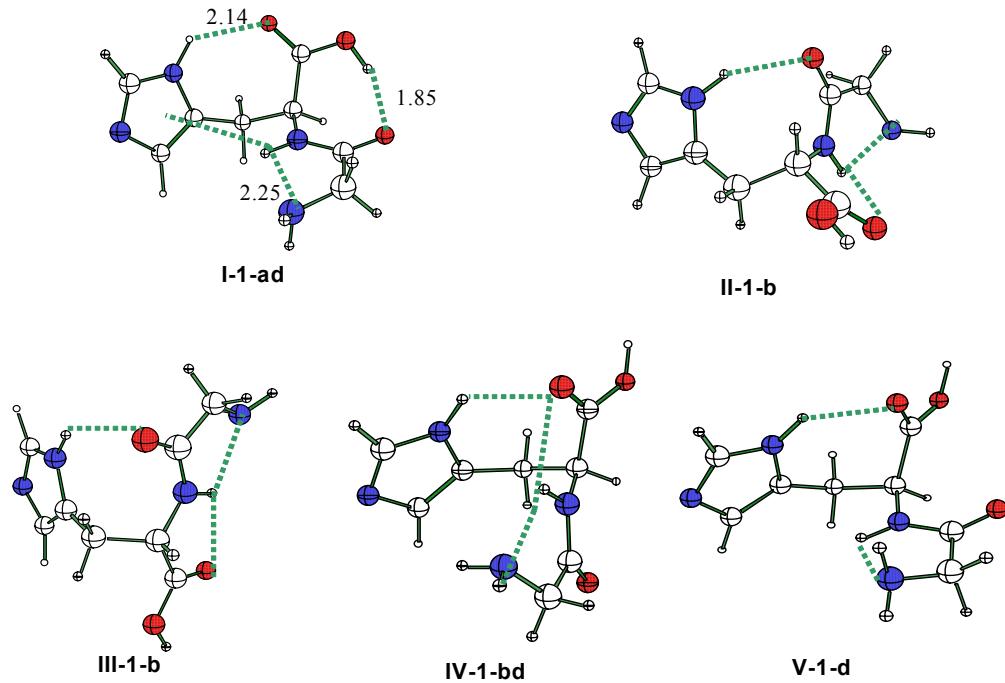
#### Low energy conformers of the $\text{N}^\delta\text{-H}$ tautomer of GlyHis

The relative energies of all conformers, grouped in five families according to the descriptors defined above, are gathered in the left part of Table I. The structures of the most stable conformers for each of the first five families are shown in Figure 3, while a series of conformers of family I are shown in Figure 4. The best structure overall is I-1-ad, which bears a  $\text{C}_7$  ring with a H bond from the  $\text{N}^\delta\text{-H}$  to the His carbonyl oxygen. In order to maximize H bonding, the carboxyl group is trans, enabling H bond donation from its O-H bond to the

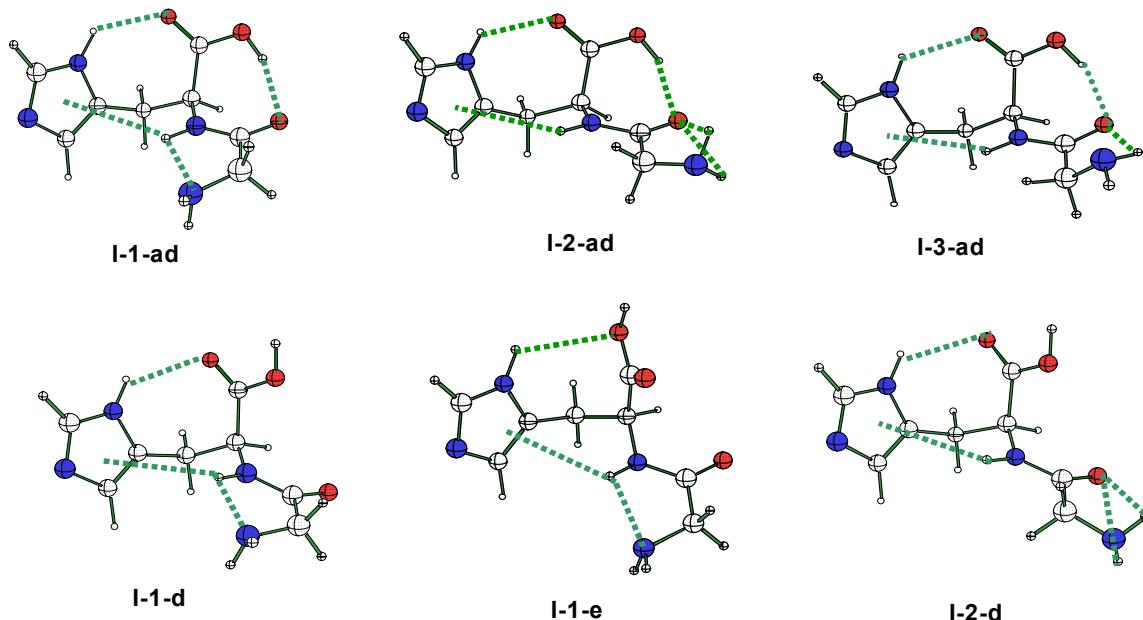
peptidic oxygen (which happens to be a  $\text{C}_7$  ring of another type, not used herein for structure specification). The so-defined orientation of the peptide linkage is such that its N-H bond may interact with the  $\pi$  cloud of imidazole. At the same time, the  $\text{NH}_2$  terminus is oriented in such as way as to be a H bond acceptor from the peptidic N-H. This structure is particularly stable since all heteroatoms are engaged in H bonding, except for the  $\text{N}_e$  (for which it is structurally impossible). The second most stable conformer is I-2-ad (see Figure 4), 10  $\text{kJ}\cdot\text{mol}^{-1}$  less stable than I-1-ad. The only difference between I-1-ad and I-2-ad is the orientation of the  $\text{CH}_2\text{NH}_2$  at the N-terminus. In I-1-ad, the amino terminus is a H bond acceptor from the peptidic N-H, while in I-2-ad it is a double H bond donor to the peptidic oxygen. In all cases studied herein, we find the same energy ordering between these two orientations of the N terminus. The next conformer of the same family (the fourth most stable overall) is I-3-ad, which differs from I-2-ad by the orientation of the amino terminus, which interacts with the peptidic oxygen via a single N-H bond rather than both. The fact that this single N-H bond is better oriented towards the peptidic oxygen than either of the N-H bonds in I-2-ad does not compensate completely for the loss of one interaction, leading to a destabilization of 3  $\text{kJ}\cdot\text{mol}^{-1}$  (i.e. 13  $\text{kJ}\cdot\text{mol}^{-1}$  less stable than I-1-ad). Other conformers of the same family have the carboxyl group in the cis conformation. Although it is intrinsically more stable than the trans (by 22  $\text{kJ}\cdot\text{mol}^{-1}$  in acetic acid), it does not allow simultaneous H bonding to the oxygen carbonyl on the one hand, and from the O-H bond on the other. The most stable of such conformers is I-1-d, which differs from I-1-ad only by the orientation of the OH group. It lies 17  $\text{kJ}\cdot\text{mol}^{-1}$  higher in energy. Other conformers add to this another less favourable interaction relative to I-1-ad, such as the orientation of the amino terminus, or the hydroxyl oxygen instead of the more basic carbonyl oxygen as a H bond acceptor (see Figure 4), and are more than 20  $\text{kJ}\cdot\text{mol}^{-1}$  less stable than I-1-ad.

We now turn to the lowest energy conformers of other families. The most stable structure in family II, II-1-b, shown in Figure 3, bears a  $\text{C}_8$  ring. Here it is the peptide carbonyl oxygen, rather than that of the carboxyl terminus, which interacts with the  $\text{N}_\delta\text{-H}$  bond of imidazole. The peptide linkage is oriented in such a way as to allow H bond donation from the peptidic N-H to both the amino terminus and the carbonyl oxygen of the C terminus. It is 12  $\text{kJ}\cdot\text{mol}^{-1}$  less stable than I-1-ad. The most stable structure in family III, III-1-b, has a  $\text{C}_8$  ring analogous to that in family II, however in a boat-like conformation. The peptide linkage has the same orientation as II-1-b, and thus forms the same H bonds involving the amidic hydrogen. The most stable structure in family IV, IV-1-bd, has a  $\text{C}_7$  ring involving the C-terminus carbonyl oxygen as in family I, however the ring has a half-chair-like conformation, while it is boat-like in family I. The two families also differ by the relative orientations of the peptidic and C-terminus planes (third descriptor above), perpendicular for family I and parallel for family IV.

This parallel orientation is common with family II, and it allows again the peptidic N-H bond to interact with both termini. This leaves the peptidic carbonyl oxygen without a H bond, so that this structure is less stable than I-1-ad by  $14 \text{ kJ.mol}^{-1}$ . Family V has a  $\text{C}_7$  ring of the same type as that of family I, but they differ by their conformation : half-chair-like in I, and boat-like in V. With a cis carboxyl group leaving the OH bond without H bonding, the most stable structure of family V, V-1-d, lies  $16 \text{ kJ.mol}^{-1}$  above I-1-a. Some additional, generally less stable, conformers of this species may be found in Table 1.



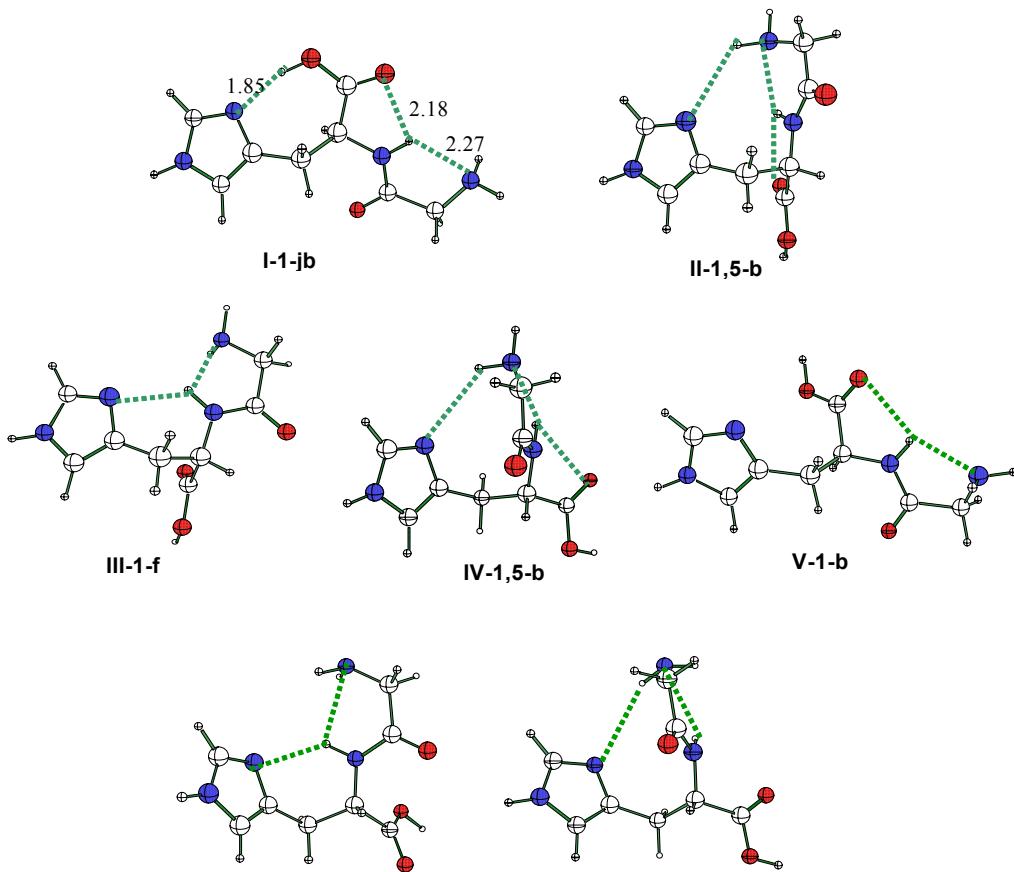
**Figure 3 :** Structure of the lowest energy conformer for each of the five families of the  $\text{N}^\delta\text{-H}$  tautomer of GlyHis.



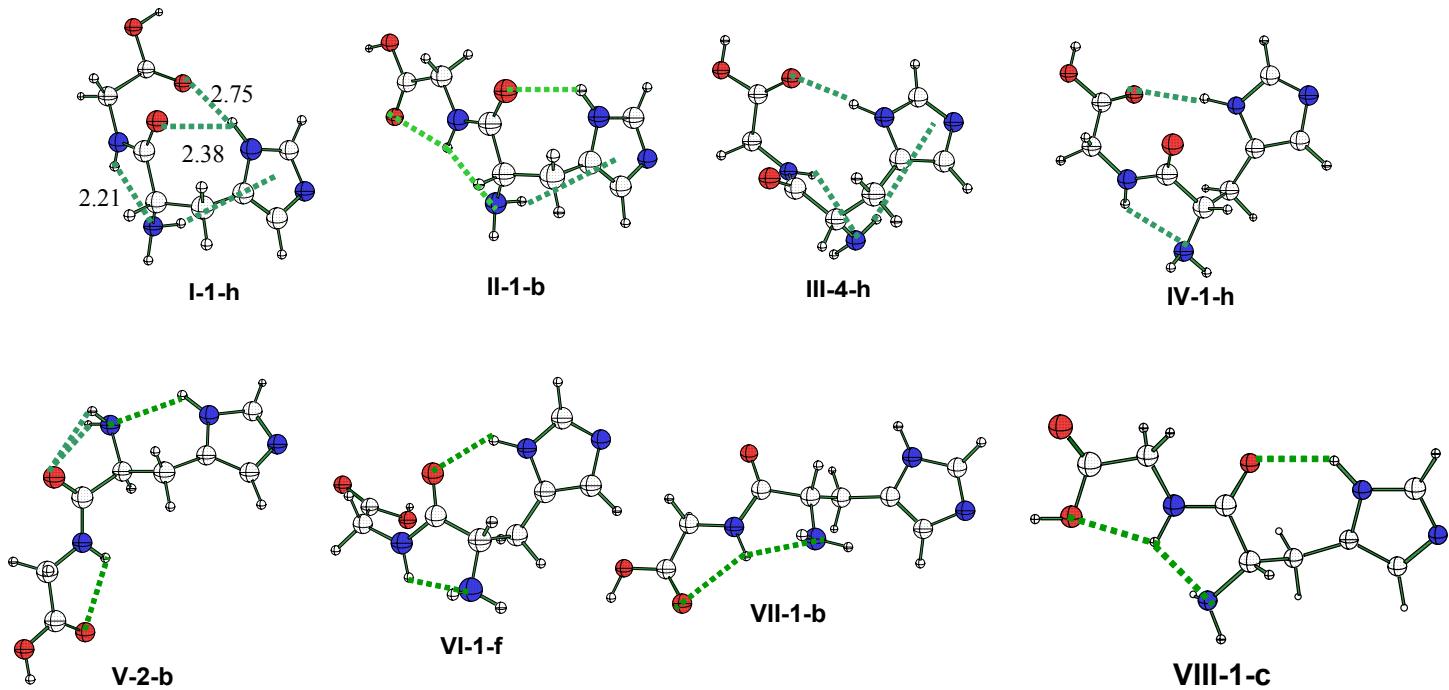
**Figure 4 :** Structures of six conformers belonging to family I of the  $\text{N}^\delta\text{-H}$  tautomer of GlyHis.

*Low energy conformers of the  $\text{N}^\varepsilon\text{-H}$  tautomer of GlyHis*

The relative energies of all conformers are gathered in the right part of Table I. The most stable conformers for each of the first seven families are shown in Figure 5. As for the most stable conformer of the  $\text{N}^\delta\text{-H}$  tautomer, the most stable structure (I-1-jb) involves a boat-like  $\text{C}_7$  ring, which is closed in this case by a  $\text{O}\text{-H---N}$  bond from the C terminus to the imidazole  $\text{N}_\delta$ . For this to occur, the carboxyl group is in its trans conformation, another feature that is common to both tautomers. The main chain has parallel peptide and carboxyl groups, and has glycine and imidazole in an open arrangement. This allows the peptidic  $\text{N}\text{-H}$  to interact with both the C-terminus carbonyl oxygen and with the amino terminus. On the other hand, the open structure precludes any interaction of the peptide carbonyl, which may explain why this conformer is less stable than the best  $\text{N}^\delta\text{-H}$  conformer (see Table 1; at the MP2/6-311+G(2d,2p)//MP2/6-31G\* level, this difference is  $6.5 \text{ kJ}\cdot\text{mol}^{-1}$ ). The second most stable conformer (I-2-jb, not shown) resembles the first, the only difference being that the amino terminus is a H bond donor towards the peptidic carbonyl, rather than an acceptor from the peptidic  $\text{N}\text{-H}$ . Thus the lower portion of the conformational spaces of the  $\text{N}^\delta\text{-H}$  and  $\text{N}^\varepsilon\text{-H}$  tautomers are fairly similar. The third most stable structure is also the first of family II (II-1,5-b, see Figure 5). It has a  $\text{C}_9$  ring connecting the amino terminus to the imidazole  $\text{N}_\delta$ , and in addition, the relative orientations of the peptide linkage and of the carboxyl group, and of the main chain and the imidazole, enable the formation of a  $\text{C}_6$  ring between  $\text{N}_\delta$  and the peptidic  $\text{N}\text{-H}$ . This structure is  $7 \text{ kJ}\cdot\text{mol}^{-1}$  higher in energy than I-1,5-jb. A structure analogous to that of II-1,5-b, with a  $\text{C}_9$  ring and planar main chain, is found with IV-1,5-b (the seventh most stable conformer in Table II). However the  $\text{C}_9$  conformation is different, precluding the formation of a  $\text{C}_6$  ring. This conformer is only  $3 \text{ kJ}\cdot\text{mol}^{-1}$  less stable than II-1,5-b. It is likely that the loss of a H bond is partly compensated by a smaller strain, permitting better relative orientations of the amino terminus and the peptidic  $\text{N}\text{-H}$  for H bonding. A slightly more stable structure is the best conformer of family III, III-1-f. Its favourable features are a  $\text{C}_6$  ring and a H bond from the peptidic  $\text{N}\text{-H}$  to the amino terminus. However none of its carboxyl oxygens can engage into H bonds, which is why this structure is  $8 \text{ kJ}\cdot\text{mol}^{-1}$  higher in energy than I-1-jb. Another structure shown in Figure 5 is V-1-b, which has a stable, planar main chain skeleton, but which lacks a H bond to the imidazole  $\text{N}_\delta$ . This leads to a high energy,  $18 \text{ kJ}\cdot\text{mol}^{-1}$  higher above I-1-jb. Other structures in Figure 5, the most stable of families VI and VII, are of even higher energies. There are also conformers of intermediate energies, in the  $12\text{--}20 \text{ kJ}\cdot\text{mol}^{-1}$  range above I-1-jb, which are not described in details here, but for which the specification of descriptors in Table 2 should be reasonably explicit.



**Figure 5 :** Structure of the lowest energy conformer for each of the seven families of the  $\text{N}^\epsilon\text{-H}$  tautomer of GlyHis.



**Figure 6 :** Structure of the lowest energy conformer for each of the eight families of the  $\text{N}^\delta\text{-H}$  tautomer of HisGly.

$\text{N}^\delta\text{-H}$	$\Delta E$	$\text{N}^\varepsilon\text{-H}$	$\Delta E$
<b>I-1-ad</b>	<b>0.0</b>	<b>I-1-jb</b>	3.8
I-2-ad	10.2	I-2-jb	7.7
<b>II-1-b</b>	<b>12.4</b>	<b>II-1,5-b</b>	10.9
I-3-ad	13.3	<b>III-1-f</b>	12.0
<b>III-1-b</b>	<b>13.8</b>	III-2-f	12.5
<b>IV-1-bd</b>	<b>14.3</b>	III-1-f	13.5
<b>V-1-d</b>	<b>15.8</b>	<b>IV-1,5-b</b>	13.9
I-1-d	16.6	II-1-c	14.9
II-1-c	18.6	III-1-g	15.6
II-2-b	22.6	III-4-f	15.9
I-1-e	22.7	IV-1,5-c	17.1
III-4-b	22.6	II-4-b	17.9
III-2-b	22.8	II-5-b	18.1
V-2-d	23.2	III-4-g	20.0
I-2-d	24.4	V-1-b	21.6
II-2-c	29.5	VI-4-f	22.6
		VII-4-b	23.1

**Table 1** : Description and relative energies of all conformers of the  $\text{N}^\delta\text{-H}$  (left) and  $\text{N}^\varepsilon\text{-H}$  (right) tautomers of GlyHis. All energetics are in  $\text{kJ}\cdot\text{mol}^{-1}$ , computed at the MP2(full)/6-311+G(2d,2p)//HF/6-31G(d) level, including zero-point vibrational energies.

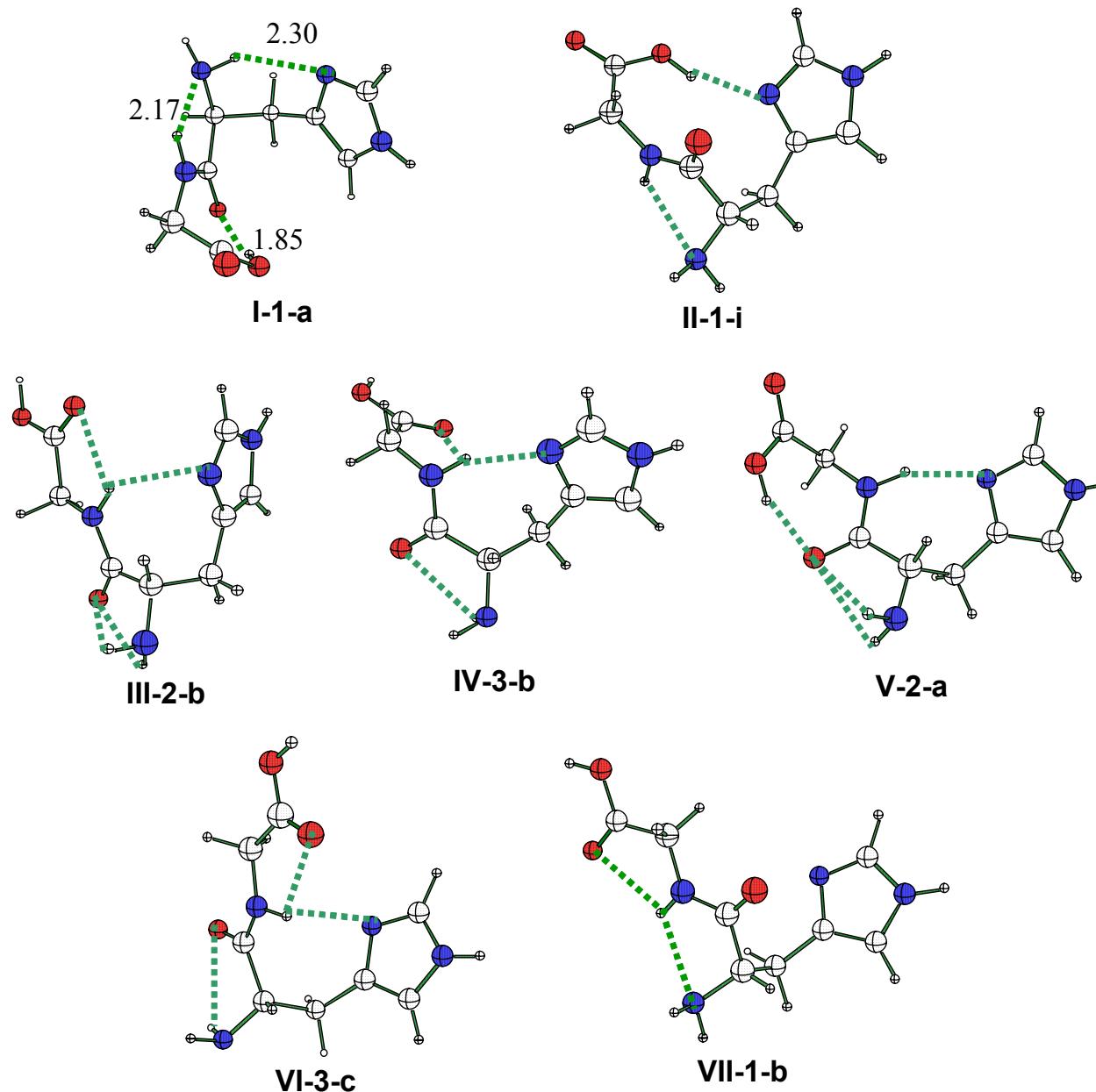
#### *Low energy conformers of the $\text{N}^\delta\text{-H}$ tautomer of HisGly*

The relative energies of all conformers found are gathered in the left part of Table II. The structures of the most stable conformers for each of the eight families identified are shown in Figure 6. The best conformer, I-1-h, has a boat-like  $\text{C}_7$  ring which involves the peptidic carbonyl, as compared to the carboxyl carbonyl in GlyHis. In addition, the perpendicular orientations of the peptidic linkage and the carboxyl terminus enable the latter to interact with the  $\text{N}_\delta\text{-H}$  of imidazole, forming a  $\text{C}_{10}$  ring. The favourable orientation of the amino terminus towards the peptidic N-H, already described previously, also permits some interaction of one of its N-H bonds with the  $\pi$  cloud of imidazole. As seen in Table 2, this is the most stable conformer, however it is slightly less stable than the best conformer of the  $\text{N}^\varepsilon\text{-H}$  tautomer. The most stable conformer of family II, II-1-b (see Figure 6), differs from I-1-h by the orientation of the Gly backbone, with the terminal carbonyl interacting with the peptidic N-H rather than with  $\text{N}_\delta\text{-H}$ . This is a weaker interaction, essentially a dipole-dipole interaction of the  $\text{C}=\text{O}$  and N-H bonds with antiparallel dipoles, rather than a H bond. As a consequence, it is less stable than I-1-h by 5  $\text{kJ}\cdot\text{mol}^{-1}$ . The next two conformers (not shown) are higher congeners of family I, differing from I-1-h only by the orientation of the COOH group. They

are less stable than I-1-h by  $6\text{-}8 \text{ kJ}\cdot\text{mol}^{-1}$ . The next most stable structure is the first of family III (III-4-h, see Figure 6). It has a  $\text{C}_{10}$  ring as does I-1-h, but with a different conformation (of type a for III-4-h), in which the  $\text{C}_7$  ring no longer exists. Although the amino terminus – N-H peptidic bond interaction is maintained, together with that of one amino N-H with the  $\pi$  cloud of imidazole, the lack of the  $\text{C}_7$  ring leads to a destabilization of  $12 \text{ kJ}\cdot\text{mol}^{-1}$ . Structure IV-1-h, the next higher conformer, presents a b-type  $\text{C}_{10}$  ring, no  $\text{C}_7$  ring, and has no amino N-H – imidazole stabilization. It is about  $1 \text{ kJ}\cdot\text{mol}^{-1}$  less stable than III-4-h. Several conformers belonging to the III, IV and I families follow with increasing energies, and the first members of families V-VIII (shown in Figure 6) are all at least  $20 \text{ kJ}\cdot\text{mol}^{-1}$  less stable than I-1-h.

#### *Low energy conformers of the $\text{N}^{\epsilon}\text{-H}$ tautomer of HisGly*

The relative energies of all conformers found are gathered in the right part of Table II. The structures of the most stable conformers for each of the seven families identified are shown in Figure 7. The best structure, I-1-a, involves a  $\text{C}_6$  ring in a half-chair conformation (as do the next two most stable, I-1-f and I-1-b). The amino terminus is both a H bond donor, to  $\text{N}_{\delta}$ , and an acceptor, from the peptidic N-H. As discussed above, the latter interaction is common in low energy structures. The carboxyl terminus adopts a trans conformation to act as a H bond donor to the peptidic carbonyl. Its other typical, cis conformation, is adopted in I-1-f, which is  $7 \text{ kJ}\cdot\text{mol}^{-1}$  less stable. Thus here again, the most stable structure involves a trans carboxylic acid. The first member of family II, II-1-i, is the fourth most stable overall. Its ring is a  $\text{C}_{10}$ , and it is now the carboxyl terminus in its trans conformation which behaves as a H bond donor towards the  $\text{N}_{\delta}$ . Again the amino terminus can accept H bonding from the peptidic N-H, however both carbonyl oxygen are left without significant interaction, leading to an energy  $10 \text{ kJ}\cdot\text{mol}^{-1}$  higher than that of I-1-a. While the  $\text{C}_7$  ring was clearly the most favorable for the three previous isomers, in this case the only possibility to form a  $\text{C}_7$  is to bind the peptidic N-H bond to the  $\text{N}_{\delta}$ , which is incompatible with H bond donation of the peptidic N-H to the amino terminus. Not only is the  $\text{C}_7$  not the best ring in this case, but the most stable conformer bearing a  $\text{C}_7$ , III-2-b, is found to lie  $15 \text{ kJ}\cdot\text{mol}^{-1}$  higher than I-1-a. The first member of family IV, IV-3-b, has a conformation very close to that of III-2-b. They differ only by the  $\text{C}_7$  conformation (half-chair-like in IV-3-b). Another  $\text{C}_7$  occurs in family V, V-2-a, with the same  $\text{C}_7$  conformation as in IV-3. The difference consists in the perpendicular orientation of the acid function with respect to the peptidic plane (third descriptor). This orientation allows the carboxyl group in trans conformation to make a H bond to the peptidic carbonyl, instead of being cis and bind electrostatically with the peptidic N-H (as in IV-3-b). As a balance of these changes, the energies of these three conformers III-2-b, IV-3-b and V-2-a are very close and differ only by  $3 \text{ kJ}\cdot\text{mol}^{-1}$ .



**Figure 7 :** Structure of the lowest energy conformer for each of the seven families of the  $\text{N}^\varepsilon\text{-H}$  tautomer of HisGly.

$\text{N}^\delta\text{-H}$	$\Delta E$	$\text{N}^\epsilon\text{-H}$	$\Delta E$
<b>I-1-h</b>	0.8	<b>I-1-a</b>	<b>0.0</b>
<b>II-1-b</b>	5.9	I-1-f	6.9
I-1-b	7.1	I-1-b	10.0
I-1-f	8.6	<b>II-1-i</b>	10.3
II-1-a	11.4	II-3-i	13.5
<b>III-4-h</b>	12.3	II-1-b	14.7
<b>IV-1-h</b>	13.1	<b>III-3-b</b>	15.0
IV-2-h	13.2	III-3-g	18.1
IV-3-h	15.5	<b>IV-2-b</b>	18.2
III-2-h	15.6	<b>V-2-a</b>	18.8
I-2-g	17.1	I-4-b	19.0
<b>V-2-b</b>	20.4	II-1-b	19.9
<b>VI-1-f</b>	20.9	IV-2-f	23.2
II-6-b	21.1	<b>VI-3-c</b>	22.6
<b>VII-1-b</b>	21.9	V-2-b	23.4
II-6-c	23.1	II-1-f	23.8
<b>VIII-1-c</b>	23.8	IV-2-f	25.6
VI-6-b	24.5	I-4-c	28.6
I-1-g	24.8	<b>VII-1-b</b>	28.7
VI-1-c	26.5		

**Table 2** : Description and relative energies of all conformers of the  $\text{N}^\delta\text{-H}$  (left) and  $\text{N}^\epsilon\text{-H}$  (right) tautomers of HisGly. All energetics are in  $\text{kJ}\cdot\text{mol}^{-1}$ , computed at the MP2(full)/6-311+G(2d,2p)//HF/6-31G(d) level, including zero-point vibrational energies.

### Zwitterions

Amino acids are known to have zwitterionic structures in aqueous solution in the biologically significant range of pH (ca. 6-9). In the gas phase, the positive and negative charges cannot be as efficiently stabilized as they are in solution, and the zwitterions are expected to be much less stable. In fact, all available evidences point to structures without formal charges in the gas phase. Thus we expect that for dipeptides such as HisGly and GlyHis, zwitterions are not the most stable isomers either. Yet it was deemed necessary to check this issue computationally. The number of stable conformers for the zwitterions of the tautomers of HisGly and GlyHis is expected to be significant smaller than for the “neutral” structures described above. Preliminary calculations indeed showed that many structures lead to collapse to non-zwitterions. As expected, the most favorable structures involve direct interaction between one of the carboxylate oxygens and one of the ammonium N-H bonds. For such motifs to be stable against proton transfer from the ammonium to the carboxylate, at least one of the charges must be stabilized by a second hydrogen bond, for instance  $\text{O}^-$  by

the imidazole  $\text{N}_{\delta}\text{-H}$ , or  $\text{N}^+\text{-H}$  by the imidazole  $\text{N}_{\delta}$ . In all cases, the zwitterions were found to be much less stable than many conformers without formal charges. The smallest difference found between the best zwitterion and the best “neutral” conformer is for the  $\text{N}^{\epsilon}\text{-H}$  tautomer of HisGly, where it is  $89 \text{ kJ}\cdot\text{mol}^{-1}$ . This zwitterion happens to have a structure that is fairly similar to that reported for the crystal structure of the hemihydrate of HisGly.<sup>7c</sup> Since all zwitterions are so high in energy, they are not described further here for the sake of brevity.

#### *General trends*

All of the low energy structures determined in this work (in a range of ca.  $20 \text{ kJ}\cdot\text{mol}^{-1}$ ) bear a  $\text{C}_n$  ring. Moreover, this ring is a  $\text{C}_7$  in the most stable conformers of three out of the four isomers. In these three cases, the ring conformation is boat-like. Clearly whenever the skeleton is compatible with such a motif, it is highly favorable. A feature that is common to all four lowest energy conformers, is the N-terminus conformation, in which the amino group is a H bond acceptor from the peptidic N-H bond. Finally, in three out of the four cases, maximizing the strength of hydrogen bonding leads to a trans carboxyl group being present in the most stable conformer, even though it is intrinsically much less stable than the cis conformer. As expected, though, the specificities of the various isomers are such that not all of the descriptors can take similar values for all most stable conformers. The relative orientations of the peptide linkage and the carboxyl group, the relative orientations of the Gly skeleton and the imidazole ring, and the cis or trans conformation of the carboxyl group, lead to structures that are significantly different, in order to optimize the network of hydrogen bonds.

Comparison of the energies of the most stable conformers, for the  $\text{N}^{\delta}\text{-H}$  and  $\text{N}^{\epsilon}\text{-H}$  tautomers, indicate that the most favorable tautomer of GlyHis is the  $\text{N}^{\delta}\text{-H}$ , while that for HisGly is the  $\text{N}^{\epsilon}\text{-H}$ . The energy differences are  $4$  and  $1 \text{ kJ}\cdot\text{mol}^{-1}$  for GlyHis and HisGly, respectively, at the MP2/6-311+G(2d,2p)//HF/6-31G\* level, including ZPE correction. Such values are too small to draw a safe conclusion regarding the lowest energy tautomers. Geometries were therefore re-optimized at the MP2/6-31G\* and final energetics were recomputed at the same level as for the HF geometries. The energy differences between the tautomers are found to be slightly increased, to  $5 \text{ kJ}\cdot\text{mol}^{-1}$  for both GlyHis and HisGly. These results leave little doubt that the most stable tautomer is  $\text{N}^{\delta}\text{-H}$  for GlyHis and  $\text{N}^{\epsilon}\text{-H}$  for HisGly.

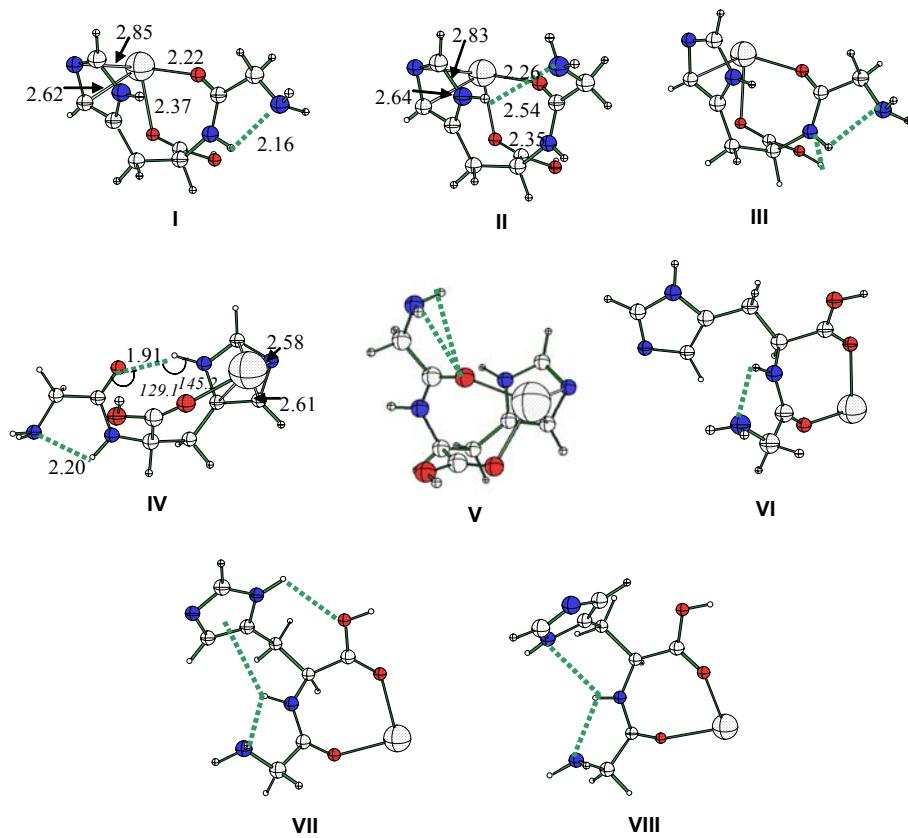
#### *Sodium complexes*

The low energy conformation spaces for the sodium complexes of the four isomers are expected to be significantly restricted as compared to those of the free peptides, since previous studies<sup>21</sup> have shown that multidentate binding of  $\text{Na}^+$  is a strong stability factor in

its complexes with amino acids and oligopeptides. For GlyGly, it was shown that the most stable structures involve ion interaction with both carbonyl oxygens, differing by the interaction, or lack thereof, with the amino terminus.<sup>22</sup> In GlyHis and HisGly, our results above suggest that for the  $\text{N}^\varepsilon\text{-H}$  tautomers, the imidazole  $\text{N}^\delta$  nitrogen may be an additional chelation site. It may also be anticipated that interaction of sodium with part or all of the imidazole  $\pi$  electrons will introduce an alternative binding capability, open to both types of tautomers. Because the number of low-energy structures is much less for sodium complexes than for free peptides, we do not introduce descriptors, and mostly describe the structures in terms of the sodium chelation sites and intramolecular H binding they involve.

*Low energy structures of the  $\text{N}^\delta\text{-H}$  tautomer of GlyHisNa<sup>+</sup>*

The principles summarized above are illustrated by the relative energies of the most stable structures (see Table 3 and Figure 8) : the most favorable chelation of sodium is to both carbonyl oxygens and the imidazole ring. Since strong binding to  $\text{N}^\varepsilon$  is sterically impossible, the ion interacts with the  $\pi$  cloud of imidazole. In both I and II,  $\text{Na}^+$  is bound to both carbonyl oxygens, and sits above the plane of imidazole, with closer interaction with one of the  $\text{C}=\text{N}$  bonds. The two structures differ by the orientation of the  $\text{NH}_2$  terminus : it interacts with the peptidic N-H in I, while it is a H-bond acceptor from the imidazole  $\text{N}^\delta\text{-H}$  in II. The distances from sodium to its chelation partners are almost the same in both cases. I is more stable than II by 12 kJ.mol<sup>-1</sup>. A third structure of the same type, III, differs from I by the terminal acid which is cis in I and trans in III, with the O-H bond pointing towards the peptidic nitrogen. Since the latter is not a particularly strong H bond acceptor, III is less stable than I by 20 kJ.mol<sup>-1</sup>. A fourth structure, V, is also bound to sodium via the same sites as I, II and III. It now has the  $\text{NH}_2$  terminus interacting with the peptidic C=O. As seen above for the conformations of the free peptides, this is less favorable for  $\text{NH}_2$  than interacting with the peptidic N-H. V is less stable than I by 28 kJ.mol<sup>-1</sup>, and is essentially degenerate with structure IV, in which the peptidic C=O is H-bound to the  $\text{N}^\delta\text{-H}$  of imidazole rather than to sodium as in the previous cases. The remaining, less stable structures of this isomers (V, VI and VII) all have sodium bound to the two carbonyl oxygens only. In all cases, the peptidic N-H bond interacts with both the  $\text{NH}_2$  terminus and imidazole in varying orientations. These three structures have similar energies, 30-40 kJ.mol<sup>-1</sup> higher than that of I.

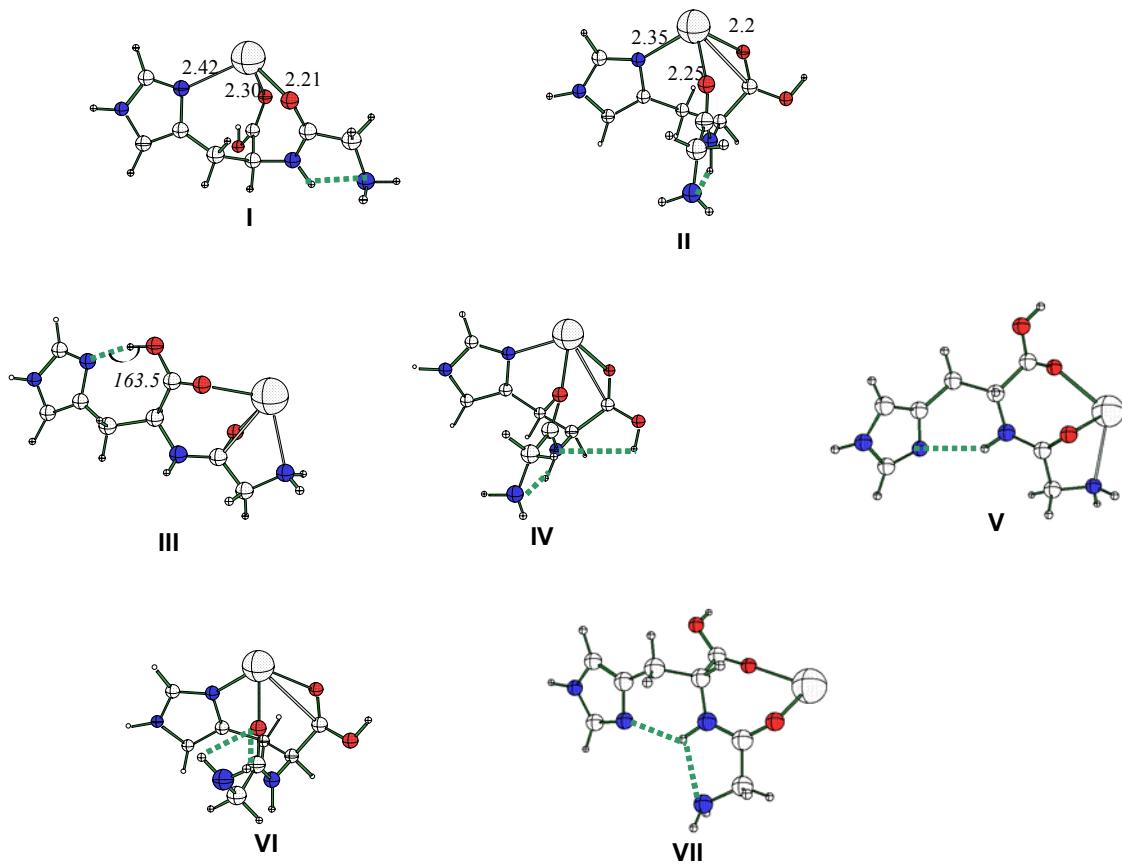


**Figure 8 :** Low energy structures of the sodium complex of the  $\text{N}^\delta\text{-H}$  tautomer of GlyHis.

*Low energy structures of the  $\text{N}^\delta\text{-H}$  tautomer of GlyHis $\text{Na}^+$*

Here as for the  $\text{N}^\delta\text{-H}$  tautomer above, the most stable structures involve chelation to both carbonyl oxygen (see Table 3 and Figure 9). However the most favorable binding mode of  $\text{Na}^+$  to imidazole is no longer to the  $\pi$  cloud N, but rather to the  $\text{N}_\delta$  nitrogen. This tridentate binding is found in I, II, IV and VI. In the lowest three of these structures, the  $\text{NH}_2$  terminus is in its most favorable conformation, interacting with the peptidic N-H. I and II differ by the side chain conformation of His. In I it is such that imidazole is closer to the C terminus, while it is closer to the main chain and peptidic carbonyl in II. I and II are isoenergetic. II and IV share a similar main chain conformation, however in IV the C terminus acid is in its less favorable trans conformation. This enables interaction of the acid O-H with the peptidic nitrogen, albeit not strongly enough to compensate for the acid destabilization. A better compromise is obtained in III, where  $\text{Na}^+$  interacts with the carbonyl oxygens and the N terminus nitrogen. Here the acid is trans as in IV, with the O-H now pointing towards the  $\text{N}_\delta$  of imidazole. This leads to favorable H binding with a O-H---N bond angle of  $164^\circ$ . Yet this structure is less stable than I by  $14 \text{ kJ}\cdot\text{mol}^{-1}$ . Another type of intramolecular H bond is found in V and VII, in which it is the peptidic N-H which binds to the  $\text{N}_\delta$ . These two structures have very similar backbones, with both carbonyls bound to  $\text{Na}^+$ . They mostly differ by the conformation of the

N terminus, and it is the interaction of  $\text{NH}_2$  with  $\text{Na}^+$  (in V) rather than with the peptidic N-H (in VII) which is the most favourable. Finally, VI differs from II by the  $\text{NH}_2$  conformation, and as seen several times before, H donation to the peptidic C=O is less favourable than H acceptation from the peptidic N-H, by 25  $\text{kJ}\cdot\text{mol}^{-1}$  in this case.



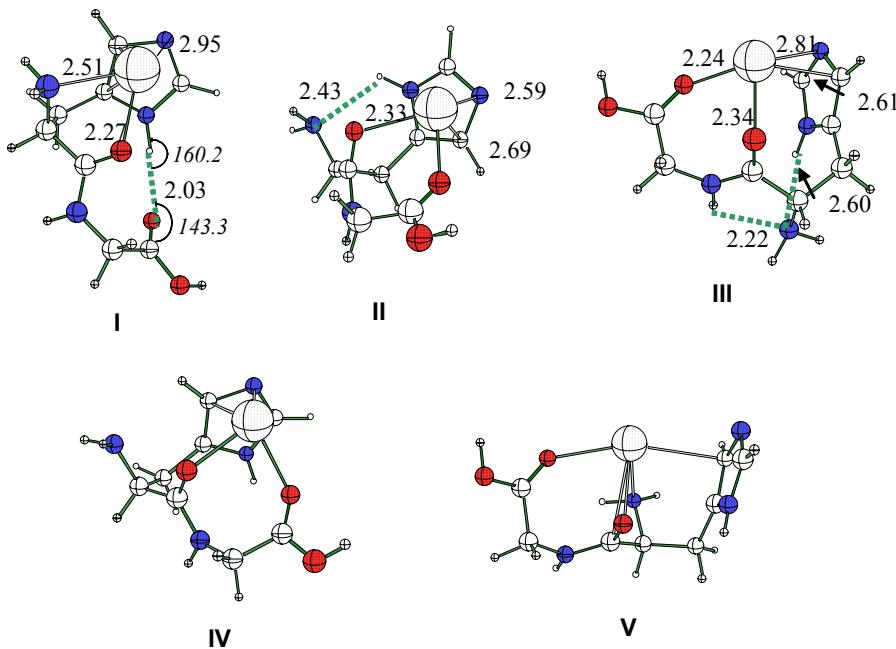
**Figure 9 :** Low energy structures of the sodium complex of the  $\text{N}^\delta\text{-H}$  tautomer of GlyHis.

$\text{N}^\delta\text{-H}$	$\Delta E$	$\text{N}^\varepsilon\text{-H}$	$\Delta E$
I	38.1	I	<b>0.0</b>
II	49.9	II	-0.1
III	58.6	III	13.7
IV	67.2	IV	16.5
V	66.4	V	20.7
VI	68.6	VI	24.7
VII	69.5	VII	36.5
VIII	78.0		

**Table 3 :** Relative energies of the lowest energy conformers of the sodium complexes of the  $\text{N}^\delta\text{-H}$  (left) and  $\text{N}^\varepsilon\text{-H}$  (right) tautomers of GlyHis. All energetics are in  $\text{kJ}\cdot\text{mol}^{-1}$ , computed at the MP2(full)/6-311+G(2d,2p)//HF/6-31G(d) level, including zero-point vibrational energies.

*Low energy structures of the  $\text{N}^\delta\text{-H}$  tautomer of  $\text{HisGlyNa}^+$*

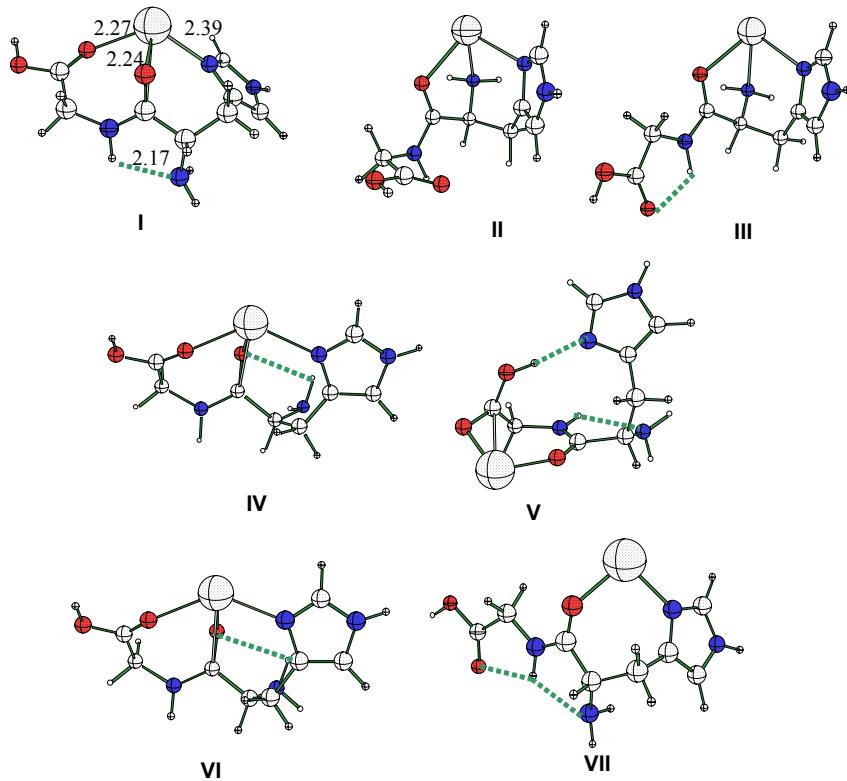
The results are summarized in Table 4 and Figure 10. As for the  $\text{N}^\delta\text{-H}$  tautomer of  $\text{GlyHisNa}^+$ , interaction of the ion with the  $\text{N}_\epsilon$  of imidazole is not possible, therefore  $\text{Na}^+$  interacts with the  $\pi$  cloud. However in this case, a significant difference may be seen : the most stable structure does not involve  $\text{Na}^+$  chelation to both carbonyls, but rather to the peptidic carbonyl and the  $\text{NH}_2$  terminus, and sits on top of the ring. This arrangement allows for efficient H bond donation from the  $\text{N}_\delta\text{-H}$  to the C terminal carbonyl oxygen, with a relatively short H---O distance of 2.03 Å. More importantly, the  $\text{N}\text{-H}^\cdots\text{O}$  and  $\text{H}^\cdots\text{O}=\text{C}$  angles are 160 and 143° respectively. The reason why the  $\text{N}^\delta\text{-H}$  tautomer of  $\text{HisGlyNa}^+$  is unique is that this H bond involves a  $\text{C}_{10}$  cycle (see Scheme 2) with enough flexibility to allow for nearly optimum  $\text{N}\text{-H}^\cdots\text{O}=\text{C}$  orientation. In contrast, in the  $\text{GlyHisNa}^+$   $\text{N}^\delta\text{-H}$  tautomer, there is only a  $\text{C}_7$  or  $\text{C}_8$  possible between the  $\text{N}_\delta\text{-H}$  and the terminal or peptidic  $\text{C}=\text{O}$ , respectively (see, e.g., isomers IV of  $\text{GlyHisNa}^+$  in Figure 8 with  $\text{N}\text{-H}^\cdots\text{O}$  and  $\text{H}^\cdots\text{O}=\text{C}$  angles of 145 and 129°). In the  $\text{N}^\delta\text{-H}$  tautomers of  $\text{GlyHisNa}^+$  and  $\text{HisGlyNa}^+$ , the  $\text{N}_\epsilon\text{-H}$  bond cannot form any significant interaction ; the  $\text{N}_\delta\text{-H}$  bond can indeed engage into H bonds (see, e.g., isomers III of  $\text{GlyHisNa}^+$  in Figure 9), but this can only occur at the expense of imidazole no longer being a sodium ligand. Therefore it is only in  $\text{N}^\delta\text{-H HisGlyNa}^+$  that such imidazole N-H bonding can compensate for the loss of a carbonyl ligand to  $\text{Na}^+$ . All other structures have  $\text{Na}^+$  bound to the two carbonyl oxygens and to the  $\pi$  cloud of imidazole. They differ by two main characteristics: (i) the conformation of the side chain of His is such that imidazole is closer either to the C terminus (as in II and IV) or to the peptidic chain (as in III and V), and (ii) the orientation of the amino terminus. The latter is a H bond acceptor from the imidazole  $\text{N}_\delta\text{-H}$  in II and III, from the  $\text{C}_\epsilon\text{-H}$  in IV and provides a fourth chelation site to  $\text{Na}^+$  in V. These four structures span an energy range of less than 20  $\text{kJ}\cdot\text{mol}^{-1}$ , while II is 6  $\text{kJ}\cdot\text{mol}^{-1}$  less stable than I.



**Figure 10 :** Low energy structures of the sodium complex of the  $\text{N}^\delta\text{-H}$  tautomer of HisGly.

*Low energy structures of the  $\text{N}^\delta\text{-H}$  tautomer of HisGly $\text{Na}^+$*

The results are summarized in Table 4 and Figure 11. As with the  $\text{N}^\varepsilon\text{-H}$  tautomer of GlyHis $\text{Na}^+$ , strong interaction of the ion with the  $\text{N}_\delta$  of imidazole is possible, so that the most favorable chelation of sodium is to both carbonyl oxygens and the  $\text{N}_\delta$  in I. Yet the next two structures in stability order, II and III, have  $\text{Na}^+$  bound to  $\text{N}_\delta$ , the peptidic oxygen and the amino terminus, differing only by the conformation of the C terminus. Both are very close in energy, 22 and 24  $\text{kJ}\cdot\text{mol}^{-1}$  higher than I. Structures IV and VI have the same tridentate chelation as I, however now the conformation of His side chain is such that imidazole is closer to the peptidic chain, rather than to the C terminus as in I. As a consequence, the amino terminus is a H bond donor to the peptidic oxygen in IV and VI, rather than being an acceptor from the peptidic N-H as in I. These differences lead to IV and VI being less stable than I by 26 and 32  $\text{kJ}\cdot\text{mol}^{-1}$ , respectively. Finally,  $\text{Na}^+$  is only bound to two sites in V and VII. V is more stable than VII by 8  $\text{kJ}\cdot\text{mol}^{-1}$ , because its two ligands are the two carbonyl oxygens. As seen for other cases previously, it has a trans acid which allows for hydrogen bonding, to the  $\text{N}_\delta$  in this case. Yet with only two chelation sites to  $\text{Na}^+$ , it is less stable than structure I by 31  $\text{kJ}\cdot\text{mol}^{-1}$ .



**Figure 11 :** Low energy structures of the sodium complex of the  $\text{N}^\delta\text{-H}$  tautomer of HisGly.

$\text{N}^\delta\text{-H}$	$\Delta E$	$\text{N}^\epsilon\text{-H}$	$\Delta E$
I	42.7	I	0.0
II	48.5	II	22.4
III	55.8	III	24.4
IV	62.7	IV	26.0
V	67.0	V	30.9
		VI	32.1
		VII	38.7

**Table 4 :** Relative energies of the lowest energy conformers of the sodium complexes of the  $\text{N}^\delta\text{-H}$  (left) and  $\text{N}^\epsilon\text{-H}$  (right) tautomers of HisGly. All energetics are in  $\text{kJ}\cdot\text{mol}^{-1}$ , computed at the MP2(full)/6-311+G(2d,2p)//HF/6-31G(d) level, including zero-point vibrational energies.

## Conclusion

A comprehensive study of the low energy structures of both tautomers of GlyHis and HisGly, and their sodium complexes has been described. Compared to what has been previously described in the literature for GlyGly, the side chain of His provides additional binding capability. While the  $\text{N}^\epsilon$  position is too remote for interaction, the  $\text{N}^\delta$  can engage in H bonding

either as an acceptor or as a donor, depending upon the tautomer. This capability has a strong influence on all low energy structures of the four peptides. Maximizing the H-bonding also leads to the C-terminal carboxylic acid being in its trans conformation in three out of the four cases. Overall, the most stable tautomer of GlyHis is the  $\text{N}^\delta\text{-H}$ , while that of HisGly is the  $\text{N}^\varepsilon\text{-H}$ .

In the sodium complexes, the  $\text{N}_\delta$  position of the imidazole side chain of His has again a strong influence on stability. In the  $\text{N}^\varepsilon\text{-H}$  tautomers,  $\text{N}^\delta$  is an efficient sodium chelator or H bond acceptor, while in the  $\text{N}^\delta\text{-H}$ , it may act as an acceptor, with  $\text{Na}^+$  interacting with the  $\pi$  cloud of imidazole. The stronger ion-molecule interactions in the  $\text{N}^\varepsilon\text{-H}$  tautomers lead them to be the most stable in  $\text{GlyHisNa}^+$  and  $\text{HisGlyNa}^+$ . These results illustrate the role of environment (here, sequences effects and ion chelation) on the relative energies of His tautomers. It is highly probable that environment effects in proteins are even stronger. Tautomeric forms of His in X-ray diffraction structures should therefore be assigned with caution.

### **Acknowledgement**

This work was supported by a generous grant of computer time at the CNRS Institut de Développement et de Recherche en Informatique Scientifique (IDRIS, project No. 040503).

### **References**

- 1 Cited in Shimba, N.; Serber, Z.; Ledwidge, R.; Miller, S. M.; Craik, C. S.; Dötsch, V. *Biochemistry*, 2003, 42, 9227.
- 2 See, e.g. Y.-I. Lin, C. Lim, *J. Am. Chem. Soc.* 2004, 126, 2602
- 3 Farr-Jones, S. ; Wong, W. Y. L.; Gutheil, W. G.; Bachovchin, W. W. *J. Am. Chem. Soc.* 1993, 115, 6813 and references therein.
- 4 Shimba, N.; Takahashi, H.; Sakakura, M.; Fujii, I.; Shimada, I. *J. Am. Chem. Soc.* 1998, 120, 10988.
- 5 Sudmeier, J. L.; Bradshaw, E. M.; Coffman Haddad, K. E.; Day, R. M.; Thalhauser, C. J.; Bullock, P. A.; Bachovchin, W. W. *J. Am. Chem. Soc.* 2003, 125, 8430.
- 6 Signorini, G. F.; Chelli, R.; Procacci, P. ; Schettino, V. *J. Phys. Chem. B* 2004, 108, 12252.
- 7 (a) Steiner, T. *Acta Cryst. Sect. C* 1997, 53, 730. (b) Steiner, T. *Acta Cryst. Sect. C* 1997, 53, 255. (c) Steiner, T.; Koellner, G. *Chem. Comm.* 1997, 1207.
- 8 Wang, P.; Kapota, C.; Wesdemiotis, C.; Ohanessian, G. submitted for publication.
- 9 See, e.g. Nguyen, D. T.; Scheiner, A. C.; Andzelm, J. W.; Sirois, S.; Salahub, D. R.; Hagler, A. T. *J. Comput. Chem.* 1997, 18, 1609 and references therein.
- 10 Csaszar, A. G. *J. Phys. Chem.* 1996, 100, 3541.
- 11 Gronert, S.; O'Hair, R. A. *J. Am. Chem. Soc.* 1995, 117, 2071.
- 12 Stepanian, S. G.; Reva, I. D.; Radchenko, E. D.; Adamowicz, L. *J. Phys. Chem. A* 1999, 103, 4404.
- 13 Stepanian, S. G.; Reva, I. D.; Radchenko, E. D.; Adamowicz, L. *J. Phys. Chem. A* 2001, 105, 10664.
- 14 Sun, W.; Kinsel, G. R.; Marynick, D. S. *J. Phys. Chem. A* 1999, 103, 4113.
- 15 (a) Gdanitz, R. J.; Cardoen, W.; Windus, T. L.; Simons, J. *J. Phys. Chem. A* 2004, 108, 515. (b) Rak, J.; Skurski, P.; Simons, J.; Gutowski, M. *J. Am. Chem. Soc.* 2001, 123, 11695. (c) Price, W. D.; Jockush, R. A.; Williams, E. R. *J. Am. Chem. Soc.* 1997, 119, 11988.
- 16 See, e.g. (a) Brooks, C. L. ; Case, D. A. *Chem. Rev.* 1993, 93, 2487. (b) Beachy, M. D.; Chapman, D.; Murphy, R. B.; Halgren, T. A. ; Friesner, R. A. *J. Am. Chem. Phys.* 1997, 119, 5908. (c) Gresh, N.; Kafafi, S. A.; Truchon, J.-F.; Salahub, D. R. *J. Comput. Chem.*

- 2004, 25, 823. (d) Masman, M. F.; Zamora, M. A.; Rodriguez, A. M.; Fidanza, N. G.; Peruchena, N. M.; Enriz, R. D.; Csizmadia, I. G. *Eur. Phys. J. D* 2002, 20, 531. (e) Elstner, M.; Jalkanen, K. J.; Knapp-Mohammady, M.; Frauenheim, T.; Suhai, S. *Chem. Phys.* 2001, 263, 203. (f) Bombasaro, J. A. ; Zamora, M. A. ; Baldoni, H. A. ; Enriz, R. D. *J. Phys. Chem. A* 2005, 109, 874.
- 17 Cornell, W. D.; Cieplak, P.; Bayly, C. I. ; Gould, I. R. ; Merz, K. M. ; Ferguson, D. M. ; Spellmeyer, D. C. ; Fox, T.; Caldwell, J. W. ; Kollman P. A. *J. Am. Chem. Soc.* 1995, 117, 5179
- 18 (a) Hoyau, S.; Norrman, K.; McMahon, T. B.; Ohanessian, G. *J. Am. Chem. Soc.* 1999, 121, 8864. (b) T. B. McMahon, G. Ohanessian, *Chem. Eur. J.* 2000, 6, 2931.
- 19 HyperChem 6.0 : Hypercube, Inc., 1115 NW 4<sup>th</sup> Street, Gainesville, FL 32601 (USA)
- 20 Gaussian03, revision B.05, Frisch, M. J. et al., Gaussian, Inc., Pittsburgh, PA, 2003.
- 21 See, e.g. (a) F. M. Siu, N. L. Ma, C. W. Tsang, *Chem. Eur. J.* 2004, 10, 1966. (b) H. A. Cox, R. R. Julian, S. W. Lee, J. L. Beauchamp *J. Am. Chem. Soc.* 2004, 126, 6485. (c) C. Kapota, J. Lemaire, P. Maitre, G. Ohanessian, *J. Am. Chem. Soc.* 2004, 126, 1836. (d) M. M. Kish, G. Ohanessian, C. Wesdemiotis, *Int. J. Mass Spectrom.* 2003, 227, 509. (e) T. Marino, N. Russo, M. Toscano, *J. Phys. Chem. B* 2003, 107, 2588. (f) T. Wyttenbach, J. E. Bushnell, M. T. Bowers, *J. Am. Chem. Soc.* 1998, 120, 5098.
- 22 (a) M. M. Kish, C. Wesdemiotis, G. Ohanessian *J. Phys. Chem. B* 2004, 108, 3086. (b) C. H. S. Wong, N. L. Ma, C. W. Tsang, *Chem. Eur. J.* 2002, 8, 4909. (c) B. A. Cerdá, S. Hoyau, G. Ohanessian, C. Wesdemiotis, *J. Am. Chem. Soc.* 1998, 120, 2437.
- 23 S. R. Carr, C. J. Cassady, *J. Am. Soc. Mass Spectrom.* 1996, 7, 1203.

## Bibliographie

- [1] Cox, H. A.; Julian, R. R.; Lee, S. W.; Beauchamp, J. L., Gas-phase H/D exchange of sodiated glycine oligomers with ND<sub>3</sub>: Exchange kinetics do not reflect parent ion structures. *J. Am. Chem. Soc.* **2004**, 126, (20), 6485-6490.
- [2] Wyttenbach, T.; Bushnell, J. E.; Bowers, M. T., Salt bridge structures in the absence of solvent? The case for the oligoglycines. *J. Am. Chem. Soc.* **1998**, 120, (20), 5098-5103, [http://www.chem.ucsb.edu/~bowersgroup/publications/supporting\\_info/JACS\\_1998\\_120\\_509\\_8.shtml](http://www.chem.ucsb.edu/~bowersgroup/publications/supporting_info/JACS_1998_120_509_8.shtml).
- [3] Cerdà, B. A.; Hoyau, S.; Ohanessian, G.; Wesdemiotis, C., Na<sup>+</sup> binding to cyclic and linear dipeptides. Bond energies, entropies of Na<sup>+</sup> complexation, and attachment sites from the dissociation of Na<sup>+</sup>-bound heterodimers and ab initio calculations. *J. Am. Chem. Soc.* **1998**, 120, (10), 2437-2448.
- [4] Benzakour, M.; McHarfi, M.; Cartier, A.; Daoudi, A., Interactions of peptides with metallic cations. I. Complexes glycylglycine-M<sup>+</sup> (M=Li, Na). *Theochem-J. Mol. Struct.* **2004**, 710, (1-3), 169-174.
- [5] Kussmann, M.; Nordhoff, E.; RahbekNielsen, H.; Haebel, S.; RosselLarsen, M.; Jakobsen, L.; Gobom, J.; Mirgorodskaya, E.; KrollKristensen, A.; Palm, L.; Roepstorff, P., Matrix-assisted laser desorption/ionization mass spectrometry sample preparation techniques designed for various peptide and protein analytes. *J. Mass Spectrom.* **1997**, 32, (6), 593-601.
- [6] Kish, M. M.; Ohanessian, G.; Wesdemiotis, C., The Na<sup>+</sup> affinities of alpha-amino acids: side-chain substituent effects. *Int. J. Mass Spectrom.* **2003**, 227, (3), 509-524.
- [7] Kish, M. M.; Wesdemiotis, C.; Ohanessian, G., The sodium ion affinity of glycylglycine. *J. Phys. Chem. B* **2004**, 108, (9), 3086-3091.
- [8] Hoyau, S.; Norrman, K.; McMahon, T. B.; Ohanessian, G., A quantitative basis for a scale of Na<sup>+</sup> affinities of organic and small biological molecules in the gas phase. *J. Am. Chem. Soc.* **1999**, 121, (38), 8864-8875.
- [9] McMahon, T. B.; Ohanessian, G., An experimental and ab initio study of the nature of the binding in gas-phase complexes of sodium ions. *Chem.-Eur. J.* **2000**, 6, (16), 2931-2941.
- [10] Cassady, C. J.; Carr, S. R.; Zhang, K.; Chungphillips, A., Experimental and Ab-Initio Studies on Protonations of Alanine and Small Peptides of Alanine and Glycine. *J. Org. Chem.* **1995**, 60, (6), 1704-1712.

- [11] Wong, C. H. S.; Ma, N. L.; Tsang, C. W., A theoretical study of potassium cation binding to glycylglycine (GG) and alanylalanine (AA) dipeptides. *Chem.-Eur. J.* **2002**, 8, (21), 4909-4918.
- [12] Rodriquez, C. F.; Cunje, A.; Shoeib, T.; Chu, I. K.; Hopkinson, A. C.; Siu, K. W. M., Proton migration and tautomerism in protonated triglycine. *J. Am. Chem. Soc.* **2001**, 123, (13), 3006-3012.
- [13] Jalkanen, K. J.; Elstner, M.; Suhai, S., Amino acids and small peptides as building blocks for proteins: comparative theoretical and spectroscopic studies. *Theochem-J. Mol. Struct.* **2004**, 675, (1-3), 61-77.
- [14] Elstner, M.; Jalkanen, K. J.; Knapp-Mohammady, M.; Frauenheim, T.; Suhai, S., Energetics and structure of glycine and alanine based model peptides: Approximate SCC-DFTB, AM1 and PM3 methods in comparison with DFT, HF and MP2 calculations. *Chem. Phys.* **2001**, 263, (2-3), 203-219.
- [15] Strittmatter, E. F.; Williams, E. R., Computational approach to the proton affinities of Gly<sub>n</sub> (n = 1-10). *Int. J. Mass Spectrom.* **1999**, 187, 935-948.

## **Chapitre VI :**

**Rôle structurant de  $\text{Na}^+$  sur les  
homo-peptides  $\text{Gly}_n\text{-Na}^+$ ,  $5 \leq n \leq 10$**

## Introduction

Lorsque la taille du peptide augmente, de nouvelles structures secondaires sont accessibles comme les structures hélicoïdales de type hélices  $\alpha$ ,  $3_{10}$  ou  $\pi$ . Il existe de nombreuses études sur la propension de peptides à former des hélices, notamment sur la transition entre structures globulaires et structures hélicoïdales, que ce soit en phase liquide ou en phase gazeuse. La plupart des modèles sont basés sur une description statistique de cette transition. Nous nous sommes intéressés aux différentes conformations des complexes Gly<sub>n</sub>-Na<sup>+</sup> pour  $5 \leq n \leq 10$ , afin d'établir de manière beaucoup plus précise leur paysage conformationnel, structuralement et énergétiquement. L'un des buts de ce travail était de déterminer le nombre de résidus  $n_H^{Gly}$  tel que la structure  $Gly_{n_H^{Gly}} - Na^+$  soit le minimum global sur la surface de potentiel. Nous présenterons tout d'abord quelques résultats bibliographiques sur la propension de peptides à former des hélices, puis les résultats sur les complexes Gly<sub>n</sub>-Na<sup>+</sup>.

## I. Propension à former des hélices $\alpha$

### I.A. Propension à former des hélices $\alpha$

#### I.A.1. Selon la nature des acides aminés

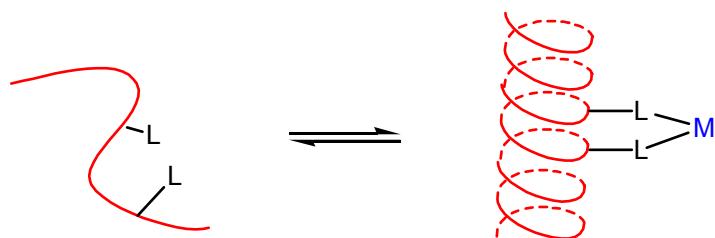
Comme nous l'avons vu dans le chapitre I, les petits peptides de moins de vingt résidus, en solution, ont très rarement une structure hélicoïdale. La stabilité relative des hélices par rapport à des structures globulaires augmente avec le nombre de résidus du peptide, mais varie aussi selon la nature des acides aminés qui le composent. L'équipe de W. DeGrado a estimé l'énergie libre de stabilisation moyenne de la forme hélicoïdale lorsque l'on introduit chacun des vingt acides aminés à une position précise dans un peptide synthétique de 29 résidus.<sup>1</sup> Ces auteurs ont ainsi pu obtenir une échelle de propension à former des hélices selon chaque acide aminé. Il en ressort que l'alanine est celui qui favorise le plus les structures en hélice  $\alpha$ , à l'inverse de la glycine. Cette tendance est également rencontrée pour les fragments hélicoïdaux de nombreuses protéines, qui ont donné lieu à des échelles de propension selon divers critères.<sup>2, 3</sup> Rappelons que la glycine et l'alanine ne diffèrent que par un méthyle en chaîne latérale. Les homo-peptides d'alanine et de glycine sont alors des systèmes modèles limites l'un favorisant la structure d'hélice  $\alpha$ , l'autre la défavorisant. Il existe plusieurs études théoriques de comparaison de stabilisation des structures hélicoïdales de  $\text{Gly}_n$  et  $\text{Ala}_n$  en phase gazeuse. L'équipe de G. E. Scuseria a notamment étudié, par des calculs de DFT en phase gazeuse, les structures en hélices  $\alpha$ ,  $3_{10}$ ,  $\text{C}_7$  et en brins  $\beta$  ( $\text{C}_5$ ) des poly-glycines et des poly-alanines infinies en utilisant des conditions aux limites périodiques.<sup>4</sup> Ils ont ainsi montré que, dans les deux cas, c'est la structure en hélice  $\alpha$  qui est la plus stable. La propension à former des hélices  $\alpha$  des poly-glycines et poly-alanines est identique du point de vue enthalpique, par contre elle est défavorisée dans le cas de la poly-glycine du point de vue entropique. L'équipe de A. Perczel a étudié l'énergie relative de  $\text{Ala}_n$ ,  $6 \leq n \leq 10$  par calcul *ab initio* et DFT.<sup>5</sup> Ces auteurs ont montré que la structure de  $\text{Ala}_n$  en hélice  $3_{10}$  est plus stable que l'hélice  $\alpha$  en phase gazeuse. A l'inverse en phase aqueuse, la structure en hélice  $\alpha$  a été calculée comme la plus stable. Il existe donc plusieurs études sur les énergies relatives des différentes hélices mais pas encore de consensus sur leur interprétation. Cependant, aucune étude théorique par calculs quantiques sur les énergies relatives des conformères les plus bas en énergie, notamment sur les formes globulaires, n'a été menée à ce jour.

## I.A.2. Augmenter la propension à former des hélices $\alpha$

Comme nous l'avons vu, la structure de petits peptides est souvent globulaire. Cependant les formes hélicoïdales peuvent être stabilisées notamment par deux méthodes, d'une part une contrainte structurale de la séquence, d'autre part en stabilisant l'hélice par interaction électrostatique.

### I.A.2.a. Par contrainte structurale

Une possibilité pour « forcer » un peptide à adopter une structure hélicoïdale est d'insérer un motif local dont la structure est contrainte à être hélicoïdale. Ceci est possible en modifiant la séquence du peptide. De nombreuses équipes ont développé des segments de séquence qui peuvent induire une stabilisation des hélices. On peut ainsi souligner le travail de Kemp *et al.* qui ont introduit le segment **He1** : analogue à Ac-Pro-Pro du côté N-terminal. La proline est souvent rencontrée comme premier résidu du côté N-terminal, dans les conformations hélicoïdales, car  $\phi(\text{Pro}) \sim 60^\circ$  est proche du dièdre de l'hélice  $\alpha$  idéale de  $57^\circ$ .<sup>6-8</sup> Il existe aussi des segments qui s'ajoutent du côté C-terminal.<sup>9</sup> Les équipes de P. Hopkins<sup>10</sup>, D. Fairlie<sup>11</sup> et C. Choi<sup>12</sup>, ont choisi de contraindre le motif hélicoïdal à l'aide d'interactions avec des métaux divalents (cf. figure VI-1). Ceci n'est possible que si le peptide possède des résidus dont les chaînes latérales peuvent servir de ligand, et ces résidus doivent être judicieusement espacés pour favoriser la forme en hélice. Ainsi C. Choi a montré que les chaînes latérales d'une histidine ou d'une cystéine en position  $i$  et  $i+4$  pouvaient favoriser l'hélice  $\alpha$ , grâce à l'interaction avec des cations  $\text{Cd}^{2+}$ ,  $\text{Ni}^{2+}$  ou  $\text{Cu}^{2+}$ , en solution.



**Figure VI-1 :** Stabilisation de la forme hélicoïdale par interaction du peptide avec un cation métallique M.

### I.A.2.b. Par interaction d'une charge avec le macro-dipôle

Une technique qui ne nécessite pas de modification chimique de la séquence peptidique est de stabiliser la forme hélicoïdale par interaction électrostatique charge-dipôle. En effet comme nous l'avons vu dans le chapitre I, les hélices sont polarisées. Elles possèdent un macro-dipôle qui correspond à une charge négative du côté C-terminal et une charge positive du côté N-terminal. Ce macro-dipôle est le plus important dans le cas de l'hélice  $\alpha$ , comparé à ceux des hélices  $3_{10}$  et  $\pi$ , car toutes les liaisons C=O et N-H

peptidiques et les liaisons hydrogènes qu'elles forment sont parallèles. Pour stabiliser une hélice, il faut donc « placer » une charge positive du côté C-terminal, ou une charge négative du côté N-terminal. Dans le cas des oligo-glycines et oligo-alanines, le site le plus basique est l'amine terminale. La protonation s'effectue donc du côté N-terminal, ce qui déstabilise la forme hélicoïdale au profit d'une conformation globulaire. Ceci a été montré par des expériences de mobilité ionique de M. Jarrold (qui sont décrites dans la section suivante). Pour favoriser la structure hélicoïdale des oligo-glycines et oligo-alanines, en phase gazeuse, deux solutions ont été envisagées. Soit ajouter un résidu très basique du côté C-terminal tel que la lysine pour imposer une protonation C-terminale, soit faire interagir le peptide avec un cation métallique judicieusement choisi. Les alcalins étant des cations durs, vont interagir préférentiellement avec les sites les plus polaires que sont les fonctions carbonyle, et notamment la fonction acide terminale. Dans les deux cas, le nouveau peptide protoné ou sodié aura une structure en hélice stabilisée.

## **I.B. Etudes de peptides dérivés de $\text{Ala}_n$ par l'équipe de M. Jarrold**

L'équipe de Jarrold s'est particulièrement intéressée à la structure d'oligomères de glycine et d'alanine en fonction du nombre de résidus, par étude de mobilité ionique. Cette technique est très efficace pour différencier des structures globulaires des structures hélicoïdales pour des peptides de plus de dix résidus. Pour un nombre inférieur à dix résidus, les sections efficaces de collision des formes globulaires et hélicoïdales sont trop proches pour pouvoir conclure. Ces auteurs ont montré que  $\text{Ala}_n\text{H}^+$  ( $n=10-20$ ) ont une structure globulaire, alors que  $\text{Ala}_n\text{LysH}^+$  et  $\text{Ala}_n\text{-Na}^+$ <sup>13, 14</sup> ont une structure hélicoïdale à partir de  $n=10$  et  $n=12$ , respectivement. Quant aux structures de  $\text{Gly}_n\text{LysH}^+$ , elles présentent toutes des structures globulaires.<sup>15</sup> Ceci a permis de montrer que la glycine a une plus faible propension à former des hélices que l'alanine en phase gazeuse, tendance identique à celle des structures en solution.

M. Jarrold *et al.* ont aussi déterminé expérimentalement les énergies de première hydratation des peptides  $\text{Ac-Lys-Ala}_n\text{+H}^+$  et  $\text{Ac-Ala}_n\text{-Lys+H}^+$ , dans des conditions d'équilibre thermodynamique. Ils ont montré que les peptides globulaires  $\text{Ac-Lys-Ala}_n\text{+H}^+$  ( $n=15$  et  $20$ ) forment une interaction avec  $\text{H}_2\text{O}$  beaucoup plus forte que les peptides hélicoïdaux  $\text{Ac-Ala}_n\text{-Lys+H}^+$  ( $n=15$  et  $20$ ).<sup>16</sup> L'énergie d'attachement d'une molécule d'eau peut alors être utilisée comme une sonde de la conformation globulaire ou hélicoïdale du peptide. Ceci a permis de déduire la structure des peptides  $\text{Ac-Lys-Ala}_n\text{+H}^+$  et  $\text{Ac-Ala}_n\text{-Lys+H}^+$ ,  $4 \leq n \leq 10$  pour lesquels il était impossible de conclure avec les expériences de mobilité ionique. Les énergies

d'hydratation de  $\text{Ac-Lys-Ala}_n\text{+H}^+$  croissent avec  $n$  pour  $4 \leq n \leq 10$  alors que pour  $\text{Ac-Ala}_n\text{-Lys+H}^+$ , on observe une augmentation de l'énergie de liaison pour  $4 \leq n \leq 7$ , puis une diminution à partir de  $n=8$ . Ceci semble indiquer que les ions  $\text{Ac-Lys-Ala}_n\text{+H}^+$   $4 \leq n \leq 10$  et  $\text{Ac-Ala}_n\text{-Lys+H}^+$ ,  $4 \leq n \leq 7$  ont une structure globulaire ; par contre les ions  $\text{Ac-Ala}_n\text{-Lys+H}^+$ ,  $n \geq 8$  présenteraient une structure hélicoïdale.

### I.C. Notre étude

Nous nous sommes intéressés aux différentes conformations accessibles pour les complexes  $\text{Gly}_n\text{-Na}^+$ ,  $5 \leq n \leq 10$  qu'elles soient globulaires ou hélicoïdales, par des méthodes de chimie quantique. Rappelons que lorsque le nombre de résidus croît, le nombre de conformations accessibles au peptide devient rapidement très grand ; nous verrons notamment que les formes globulaires sont très diverses. Nous avons couplé trois stratégies afin de déterminer les conformations les plus basses en énergie. D'une part, nous avons construit des structures par extension des systèmes  $\text{Gly}_4\text{-Na}^+$  en maximisant les interactions entre le métal et les sites de coordination des carbonyles. D'autre part, nous avons effectué plusieurs recherches conformationnelles de type Monte-Carlo basées sur des calculs de champ de forces de type AMBER, afin d'explorer la surface d'énergie potentielle (cf. chapitre III). Plusieurs recherches indépendantes ont été réalisées pour chaque taille de peptide. Toutes les structures AMBER engendrées ont été réoptimisées au niveau ri-BLYP/SV(P), puis le calcul d'énergie finale a été fait au niveau ri-BLYP/TZVPP//ri-BLYP/SV(P) (ri-BLYP/GB2//ri-BLYP/PB2), avec correction thermique. Enfin nous avons fait une recherche croisée : chacune des cinq structures les plus basses en énergie de  $\text{Gly}_n\text{-Na}^+$  déterminées au niveau ri-BLYP/GB2//ri-BLYP/PB2 a été étendue aux rangs  $n+1$  et  $n-1$  en lui ajoutant ou enlevant un résidu. Il était en effet improbable que les recherches Monte Carlo aient permis d'explorer l'espace des conformations de manière exhaustive. Nous avons donc exploité le fait que ces recherches ont été réalisées indépendamment pour toutes les dimensions de peptide, pour « faire bénéficier » chaque taille des résultats obtenus sur les tailles voisines. Ce complément s'est avéré essentiel.

## II. Les différentes structures en hélices envisagées

### II.A. Géométries

Nous avons envisagé plusieurs types d'hélices  $\alpha$  et d'hélices  $3_{10}$ , selon la nature de la coordination des carbonyles du côté C-terminal. Les trois structures d'hélices  $\alpha$  et les deux hélices  $3_{10}$  sont représentées sur la figure VI-2, pour les complexes  $\text{Gly}_9\text{-Na}^+$ . Leurs énergies

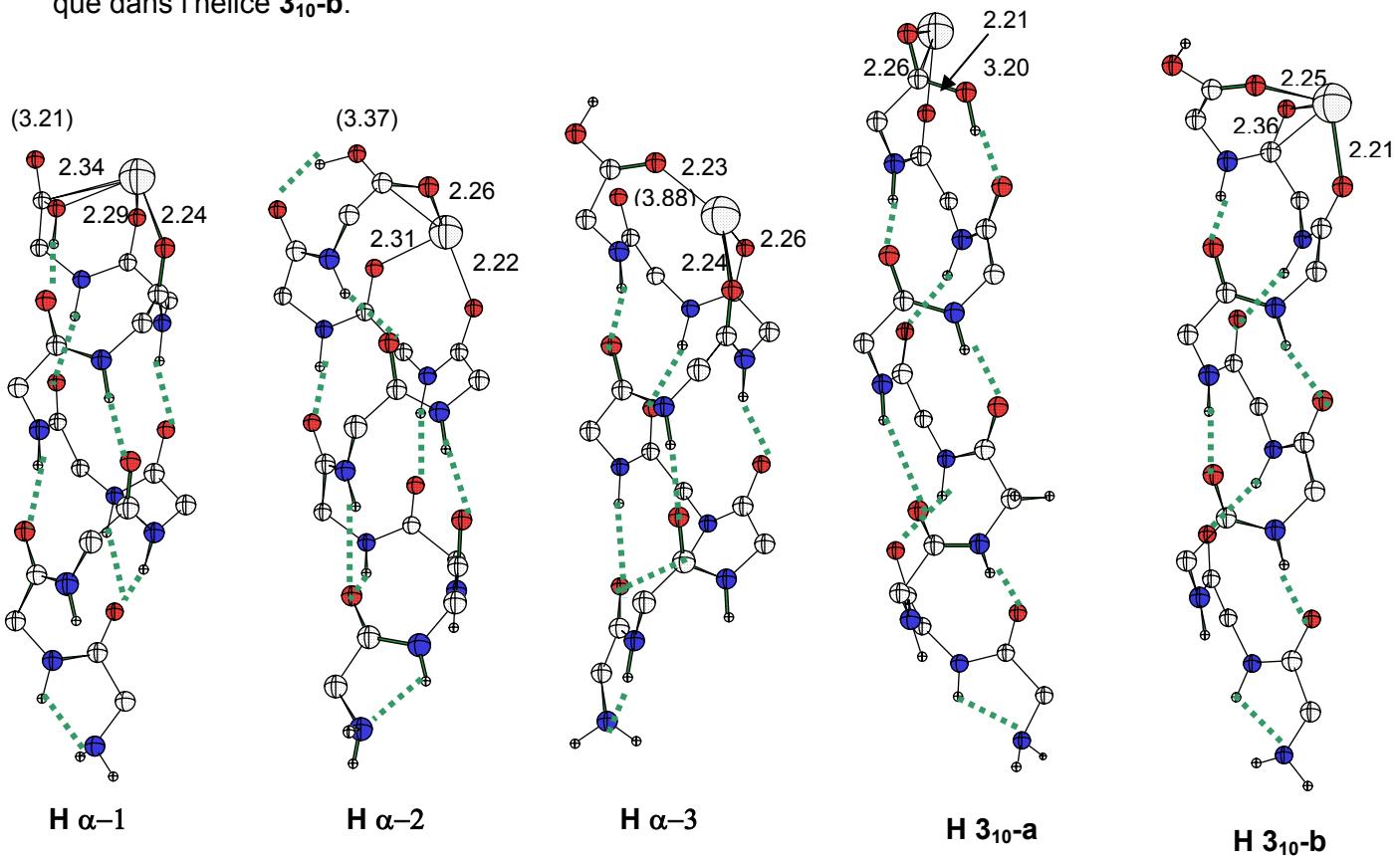
relatives sont reportées sur la figure VI-3. Les trois structures  $\text{H } \alpha$  présentent une structure en hélice  $\alpha$  pour les neuf résidus avec cinq liaisons hydrogène de type  $\text{C}_{13}$  ( $\text{N}^i\text{H} \rightarrow \text{O}^{i-4}$ ), ainsi que deux liaisons hydrogène supplémentaires du côté N-terminal de type  $\text{C}_5$  ( $\text{N}^2\text{H} \rightarrow \text{N}^1$ ) et  $\text{C}_{10}$  ( $\text{N}^4\text{H} \rightarrow \text{O}^1$ ) dues à un effet de bord. Ces trois conformations diffèrent par la complexation avec  $\text{Na}^+$ . **H  $\alpha$ -1** est complexé par  $\text{O}^7\text{O}^8\text{O}^9$ , la fonction acide trans formant une liaison hydrogène  $\text{C}_{13}$  avec le sixième résidu ( $\text{O}^9\text{H} \rightarrow \text{O}^6$ ). Dans ce cas, l'hélice  $\alpha$  se propage donc jusqu'au résidu C-terminal. **H  $\alpha$ -2** est complexé par  $\text{O}^6\text{O}^7\text{O}^9$ , et la fonction acide trans forme une liaison hydrogène de type  $\text{C}_7$  avec le huitième résidu ( $\text{O}^9\text{H} \rightarrow \text{O}^8$ ). **H  $\alpha$ -3** est complexé par  $\text{O}^6\text{O}^7\text{O}^9$ , et la fonction acide cis ne forme aucune liaison hydrogène. Nous avons calculé que ces trois structures ont des énergies relatives sensiblement constantes pour toutes les oligomères de glycine. La structure **H  $\alpha$ -3** est la plus stable, puis la structure **H  $\alpha$ -2** ( $\sim 20$  kJ/mol au dessus de **H  $\alpha$ -3**) et enfin **H  $\alpha$ -1** ( $\sim 30$  kJ/mol au dessus de **H  $\alpha$ -3**). La stabilisation due à la présence d'une liaison hydrogène supplémentaire, dans le cas de **H  $\alpha$ -1** et **H  $\alpha$ -2**, ne compense pas l'énergie défavorable de la conformation trans de l'acide. Nous avons envisagé deux structures les plus stables pour les hélices  $3_{10}$  qui présentent une structure en hélice  $3_{10}$  pour les neuf résidus avec six liaisons hydrogène de type  $\text{C}_{10}$  ( $\text{N}^i\text{H} \rightarrow \text{O}^{i-3}$ ), ainsi qu'une liaison hydrogène supplémentaire du côté N-terminal de type  $\text{C}_5$  ( $\text{N}^2\text{H} \rightarrow \text{N}^1$ ). **H  $3_{10}$ -a** est complexé au sodium par  $\text{O}^8\text{O}^9\text{O}^9\text{H}$ , la fonction acide trans formant une liaison hydrogène  $\text{C}_{10}$  avec le septième résidu ( $\text{O}^9\text{H} \rightarrow \text{O}^7$ ). **H  $3_{10}$ -b** est complexé par  $\text{O}^7\text{O}^8\text{O}^9$ , la fonction acide cis ne forme aucune liaison hydrogène. Comme dans le cas des hélices  $\alpha$ , la conformation trans de l'acide terminal donne une structure plus haute en énergie, bien qu'elle permette de former une liaison hydrogène supplémentaire. La différence d'énergie entre **H  $3_{10}$ -a** et **H  $3_{10}$ -b** est constante quand on augmente le nombre de résidus, et est de 45 kJ/mol en faveur de la conformation **H  $3_{10}$ -b**.

Tous les essais d'optimisation d'hélices  $\pi$  se sont révélés infructueux et ont conduit à des structures d'hélices  $\alpha$ , ou à une structure qui présente 3 résidus du côté C-terminal en hélice  $\pi$  et les six autres résidus en hélice  $\alpha$ . Cette structure (hélice alpha-pi- $\text{Na}^+$ ) est moins stable de 17 kJ/mol que l'hélice **H  $\alpha$ -3**.

## II.B. Energies

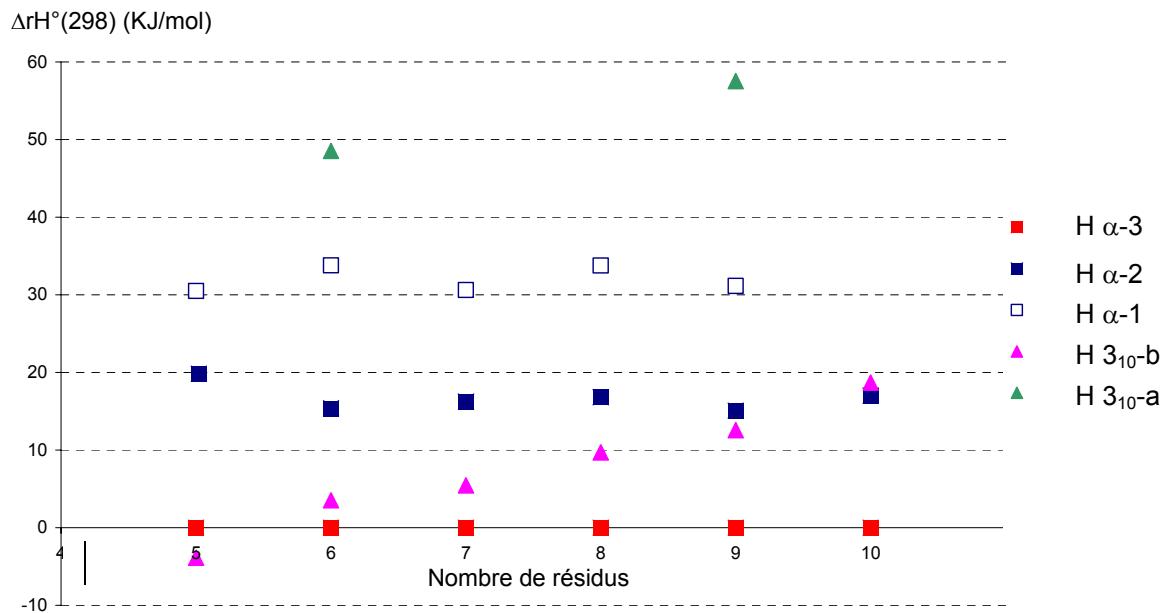
Les énergies relatives de toutes les hélices, regroupées dans la figure VI-3, sont relativement constantes entre hélices  $\alpha$ , mais montrent que les hélices  $3_{10}$  sont défavorisées par rapport aux hélices  $\alpha$  lorsque la taille du peptide augmente. On peut remarquer que pour  $n=5$ , l'hélice **3<sub>10</sub>-b** est plus stable que l'hélice  **$\alpha$ -3**, situation inversée à partir de  $n=6$ . Ceci peut s'expliquer d'une part car les hélices  $3_{10}$  forment une liaison hydrogène supplémentaire

par rapport aux hélices  $\alpha$ , pour un même nombre de résidus. Cependant les liaisons hydrogène et les carbonyles sont mieux orientés pour former un macro-dipôle plus grand dans le cas des hélices  $\alpha$ . La liaison hydrogène supplémentaire des hélices  $3_{10}$  devient progressivement négligeable par rapport à l'interaction électrostatique, quand le nombre de résidus augmente. Afin de déterminer la stabilisation due à l'interaction métal-peptide, nous avons comparé l'énergie du complexe à celle du sodium et du peptide isolé dans la géométrie du complexe, au niveau ri-BLYP/GB2. Les énergies d'interaction métal-ligand en fonction du nombre de résidus pour les conformères les plus stables de chaque type, **H  $\alpha$ -3** et **H  $3_{10}$ -b** sont présentées figure VI-4. On peut voir que, pour les deux types d'hélice, l'énergie d'interaction augmente avec le nombre de résidus, bien que la coordination du sodium en soit indépendante. Ceci illustre l'effet de l'interaction charge-macro-dipôle à longue portée. De plus on peut remarquer que pour cinq résidus, les deux structures en hélice :  **$\alpha$ -3** et  **$3_{10}$ -b** ont des énergies de stabilisation similaires (348.7 kJ/mol et 340.1 kJ/mol, respectivement). Par contre l'incrément d'énergie de stabilisation est plus important dans le cas de l'hélice  **$\alpha$ -3**, ce qui illustre le fait que le macro-dipôle est plus grand dans l'hélice  **$\alpha$ -3** que dans l'hélice  **$3_{10}$ -b**.

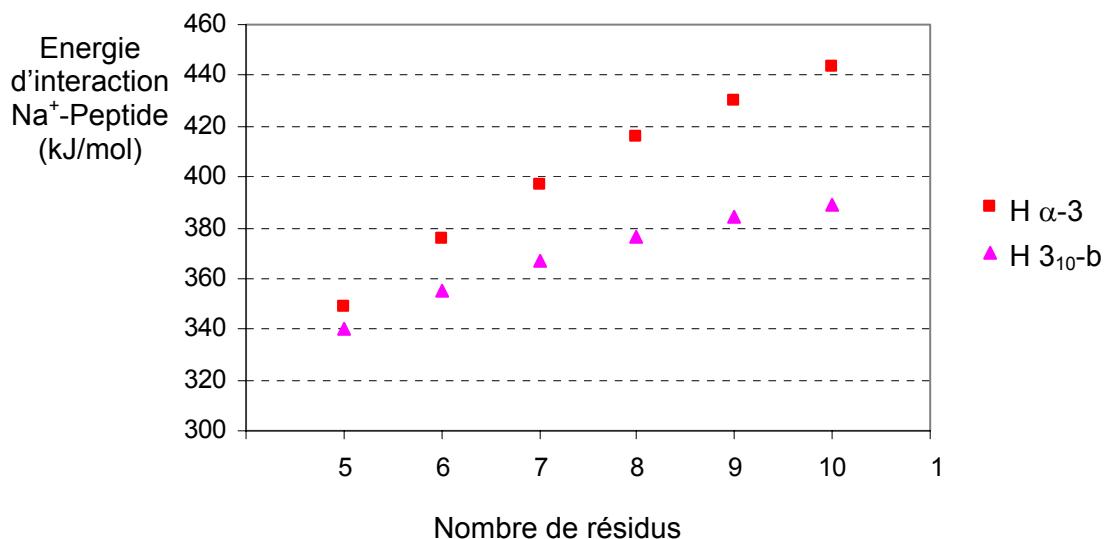


**Figure VI-2 :** Structures d'hélices  $\alpha$  et  $3_{10}$  envisagées, illustrées sur l'exemple de  $\text{Gly}_9\text{-Na}^+$ .

Ainsi pour un nombre faible de résidus, la stabilisation de la structure par les liaisons hydrogène est la contribution la plus importante et l'hélice  $3_{10}$  sera la plus stabilisée. Quand le nombre de résidus augmente, l'interaction électrostatique entre le cation et le macro-dipôle devient la contribution stabilisante la plus importante et l'hélice  $\alpha$  devient alors plus stable.



**Figure VI-3 :** Enthalpies standards de conformation à 298 K des différentes hélices relatives à celle de la structure hélice  $\text{H } \alpha\text{-3}$ , calculées au niveau ri-BLYP/GB2//ri-BLYP/PB2, en fonction du nombre de résidus n de  $\text{Gly}_n\text{-Na}^+$ .



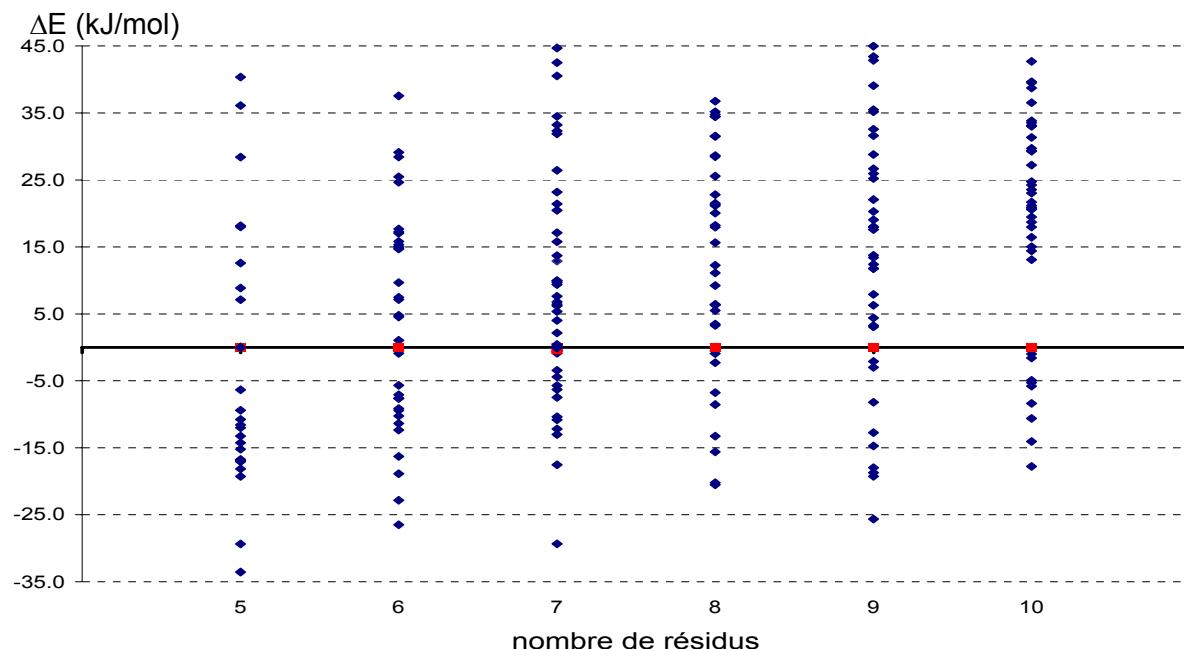
**Figure VI-4 :** Energies d'interaction métal-peptide pour les structures  $\text{H } \alpha\text{-3}$  et  $\text{H } 3_{10}\text{-b}$ , calculées au niveau ri-BLYP/GB2//ri-BLYP/PB2 en fonction du nombre de résidus n de  $\text{Gly}_n\text{-Na}^+$ .

### III. Les différentes structures globulaires

#### III.A. Tendances énergétiques générales

En appliquant la stratégie de recherche conformationnelle décrite plus haut aux structures des complexes  $\text{Gly}_n\text{-Na}^+$ ,  $5 \leq n \leq 10$ , nous avons engendré environ 350 structures dont la plupart sont globulaires, allant de 36 conformations pour  $\text{Gly}_5\text{-Na}^+$  à 72 pour  $\text{Gly}_{10}\text{-Na}^+$ . Toutes ces structures ont été optimisées au niveau ri-BLYP/PB2. Le calcul des fréquences de vibration au niveau ri-BLYP/PB2 et les calculs d'énergies au niveau ri-BLYP/GB2//ri-BLYP/PB2 n'ont été menés que pour les structures les plus stables. La dispersion des énergies relatives à celle du conformère  $\text{H } \alpha\text{-3}$  est présentée sur la figure VI-4, en se limitant à 45 kJ/mol. Il serait très fastidieux de décrire chacune de ces structures. Par la suite nous nous sommes limités aux structures des conformères les plus bas en énergie dans une gamme de 20 kJ/mol.

Cette figure fait apparaître que les structures les plus stables sont toujours des globules, mais que la différence entre le globule le plus stable et la meilleure hélice diminue quand le nombre de résidus augmente. Pour  $n=10$ , la structure  $\text{H } \alpha\text{-3}$  reste 18 kJ/mol plus haute que le meilleur conformère globulaire. Cependant nous verrons que les structures de  $\text{Gly}_{10}\text{-Na}^+$  les plus basses en énergie présentent tout de même localement des structures en hélices.



**Figure VI-5 :** Energies des conformères globulaires relatives à celle de la structure en hélice  $\text{H } \alpha\text{-3}$ , calculées au niveau ri-BLYP/GB2//ri-BLYP/PB2 en fonction du nombre  $n$  de résidus de  $\text{Gly}_n\text{-Na}^+$ .

## III.B. Géométries

### III.B.1. Descripteurs

Afin de décrire les différentes structures des conformères, nous avons choisi d'utiliser deux types de descripteurs. En effet les conformations correspondent à un compromis entre la maximisation de l'interaction électrostatique entre  $\text{Na}^+$  et les groupements carbonyle et éventuellement l'amine terminale du peptide, et la maximisation de l'auto-solvatation du peptide. Cette dernière contribution correspond aux différentes liaisons hydrogène intramoléculaires. Afin de décrire ces liaisons hydrogène, nous avons utilisé la notation précédemment rencontrée «  $\text{C}_n$  » qui correspond au motif cyclique formé par la liaison hydrogène, entre l'hydrogène d'une fonction amidique et un oxygène de carbonyle (cf. figure VI-6). Nous avons déjà rencontré les motifs  $\text{C}_7$  ou  $\gamma$ -turn,  $\text{C}_{10}$  ou  $\beta$ -turn qui forment les hélices  $\text{3}_{10}$ ,  $\text{C}_{13}$  qui forment les hélices  $\alpha$  et enfin les  $\text{C}_{16}$  qui constituent les hélices  $\pi$ . Tous ces motifs impliquent une liaison hydrogène dans laquelle le carbonyle est du côté N-terminal par rapport à la liaison N-H. On peut également construire des  $\text{C}_n$  dans lesquels le carbonyle est du côté C-terminal ; ce sont des  $\text{C}_5$ ,  $\text{C}_8$ ,  $\text{C}_{11}$ , etc.

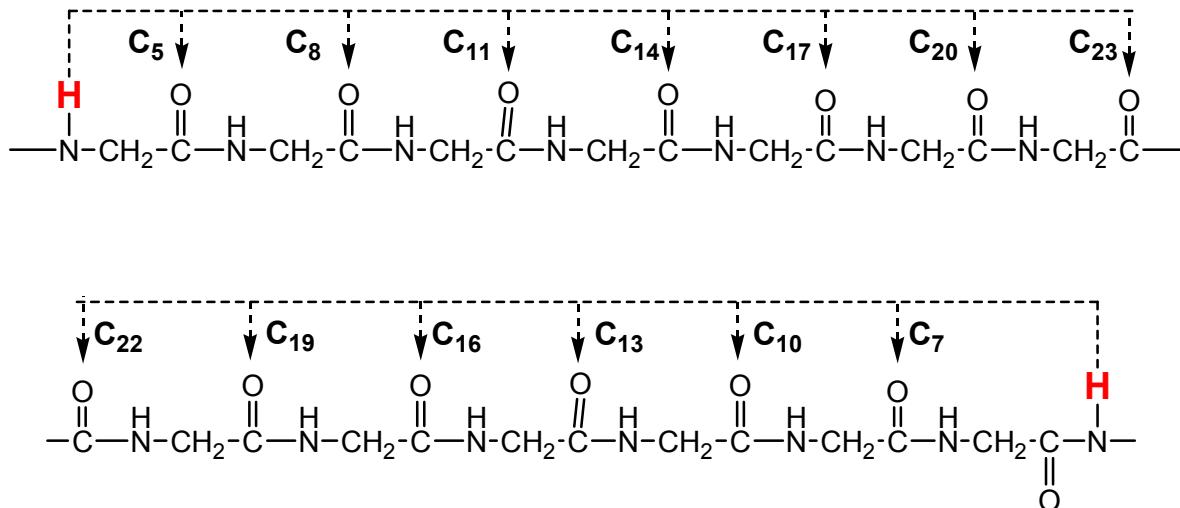


Figure VI-6 : Les différents types de  $\text{C}_n$

Afin de décrire les conformères des complexes  $\text{Gly}_n\text{-Na}^+$ , nous utiliserons une notation comprenant  $n$  éléments spécifiant d'une part les modes de coordination au sodium (en première sphère pour une distance métal-O ou métal-N inférieure à 3.20 Å, en deuxième sphère pour une distance comprise entre 3.20 et 4 Å) et les types de liaisons hydrogène des  $n$  groupements carbonyle. Rappelons que nous avons utilisé la numérotation systématique des atomes des peptides en commençant par le résidu N-terminal. Ainsi lorsque l'oxygène 1

est coordonné au métal en première (deuxième) sphère nous le noterons  $\text{O}^1$  ( $\text{o}^1$ ), s'il est impliqué dans une liaison hydrogène de type  $\text{C}_7$ , il prendra la notation  $\text{C}_7$ , s'il n'est impliqué dans aucune interaction nous le noterons  $\emptyset$ . Par exemple le conformère  $\text{Gly}_5\text{-Na}^+$  **1** qui porte la notation  $\text{O}^1\text{-O}^2\text{-O}^3\text{-O}^4\text{-O}^5$  présente une complexation du sodium par les cinq carbonyles, qui ne sont impliqués dans aucune liaison hydrogène. Dans le conformère  $\text{Gly}_6\text{-Na}^+$  **3** qui porte la notation  $\text{O}^1\text{-}\emptyset\text{-O}^3\text{-o}^4\text{-O}^5\text{-o}^6(+\text{C}_{20})\text{-o}^6\text{H}$ , le sodium est complexé en première sphère par les oxygènes  $\text{O}^1$ ,  $\text{O}^3$  et  $\text{O}^5$ , en seconde sphère par  $\text{O}^4$ ,  $\text{O}^6$  et  $\text{O}^6\text{H}$ . L'oxygène  $\text{O}^6$  est aussi récepteur d'une liaison hydrogène de type  $\text{C}_{20}$ , enfin l'oxygène  $\text{O}^2$  reste sans interaction.

### III.B.2. Géométrie des structures les plus basses en énergie

Les structures des conformères les plus bas en énergie dans une gamme de 20 kJ/mol de  $\text{Gly}_5\text{-Na}^+$  à  $\text{Gly}_{10}\text{-Na}^+$  sont présentées sur les figures VI-7 à VI-12 (respectivement). Leurs énergies relatives ainsi que les descripteurs, le nombre de liaisons hydrogène et le nombre de sites de chélation au sodium sont donnés dans les tableaux VI-1 à VI-6.

	Nom $\text{G}_5\text{Na}^+$	$E^{\text{int}}$ metal-peptide	$\Delta_r H^{\text{GB2}}(298)$ (kJ/mol)	$\Delta_r G^{\text{GB2}}(298)$ (kJ/mol)	notation	nombre de ligands	nombre de liaisons hydrogène
<b>1</b>	2GC-A1-10	427.9	-32.3	-40.7	$\text{O}^1\text{-O}^2\text{-O}^3\text{-O}^4\text{-O}^5$	5O	1
<b>2</b>	2GN-A1-4		-27.3	-32.7	$\text{O}^1\text{-O}^2\text{-O}^3\text{-O}^4\text{-O}^5$	5O	1
<b>3</b>	2GN-A1-5	375.1	-16.5	-18.9	$\text{O}^1\text{-O}^2\text{-C}_7\text{-C}_7\text{-O}^5\text{-O}^5\text{H}$	4O	3
<b>4</b>	2GC-A1-1	403.8	-16.3	-25.5	$\text{O}^1\text{-O}^2\text{-C}_7\text{-O}^4\text{-O}^5$	4O	2
<b>5</b>	2GC-A2-23		-14.8	-24.1	$\text{C}_7\text{-O}^2\text{-O}^3\text{-O}^4\text{-O}^5$	4O	2
<b>6</b>	c4ter-GC		-14.1	-22.2	$\text{N}^1\text{-O}^1\text{-O}^2\text{-O}^3\text{-O}^4\text{-C}_5$	4O,1N <sup>1</sup>	1
<b>7</b>	6sON		-14.0	-20.8	$\text{N}^1\text{-O}^1\text{-O}^2\text{-O}^3\text{-O}^4\text{-O}^5$	5O,1N <sup>1</sup>	0
<b>8</b>	2GC-A2-13	423.0	-11.6	-18.8	$\text{O}^1\text{-O}^2\text{-O}^3\text{-O}^4\text{-O}^5$	5O	1
<b>9</b>	2GN-A1-1		-10.9	-22.0	$\text{O}^1\text{-O}^2\text{-}\emptyset\text{-O}^4\text{-O}^5$	4O	1
<b>10</b>	2GC-A2-22		-10.6	-10.9	$\text{N}^1\text{-O}^1(+\text{C}_7)\text{-C}_7\text{-O}^3\text{-O}^4\text{-O}^5$	4O,1N <sup>1</sup>	1
<b>17</b>	H $\alpha$ -3	348.7	0.0	0.0	$\text{C}_{13}(+\text{C}_{10})\text{-O}^2\text{-O}^3\text{-O}^4\text{-O}^5$	4O	3

**Tableau VI-1 :** Energies d'interaction métal-peptide, grandeurs thermodynamiques à 298 K relatives à la structure **H  $\alpha$ -3** et principales caractéristiques géométriques des complexes  $\text{Gly}_5\text{-Na}^+$ .

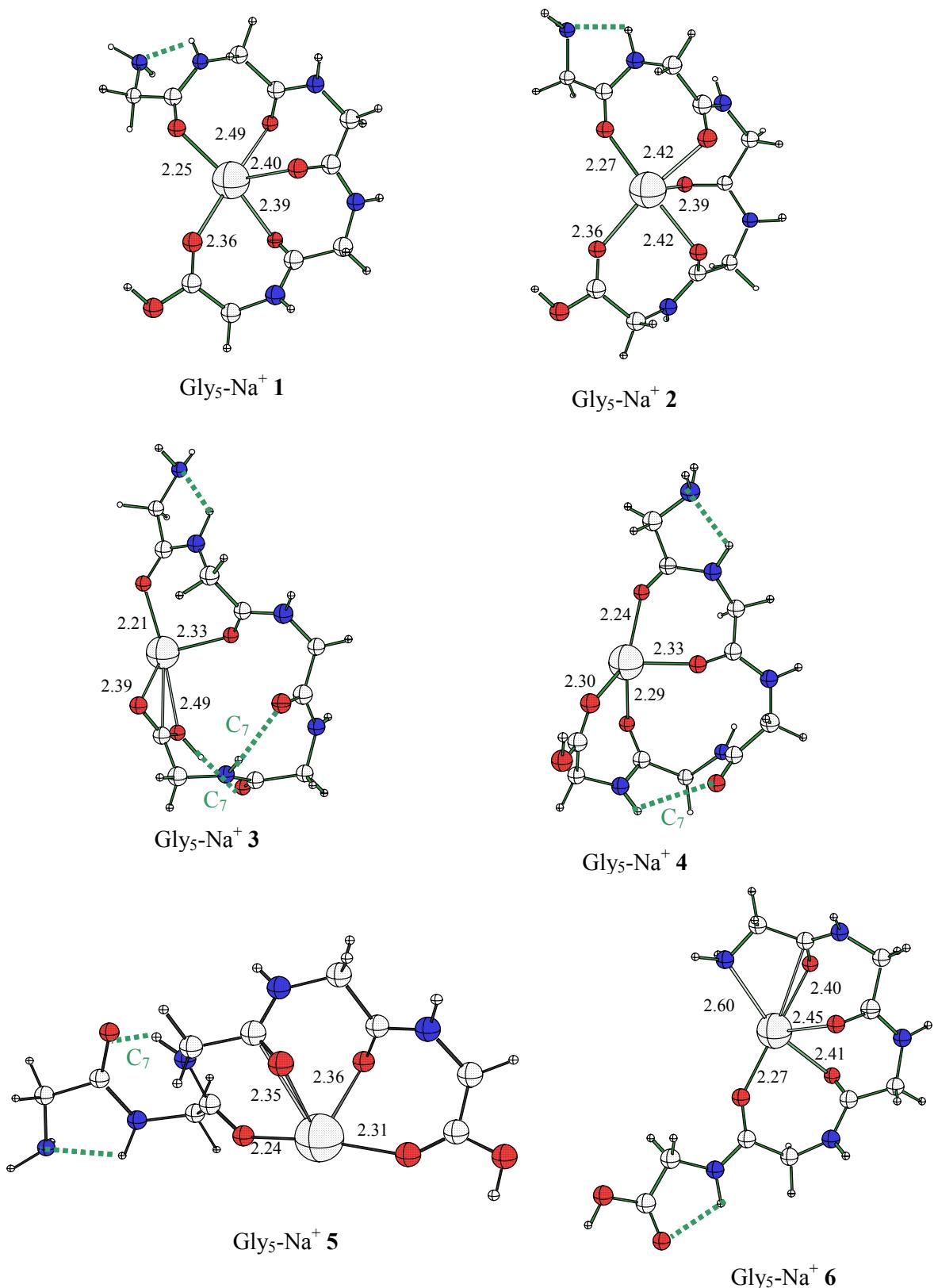
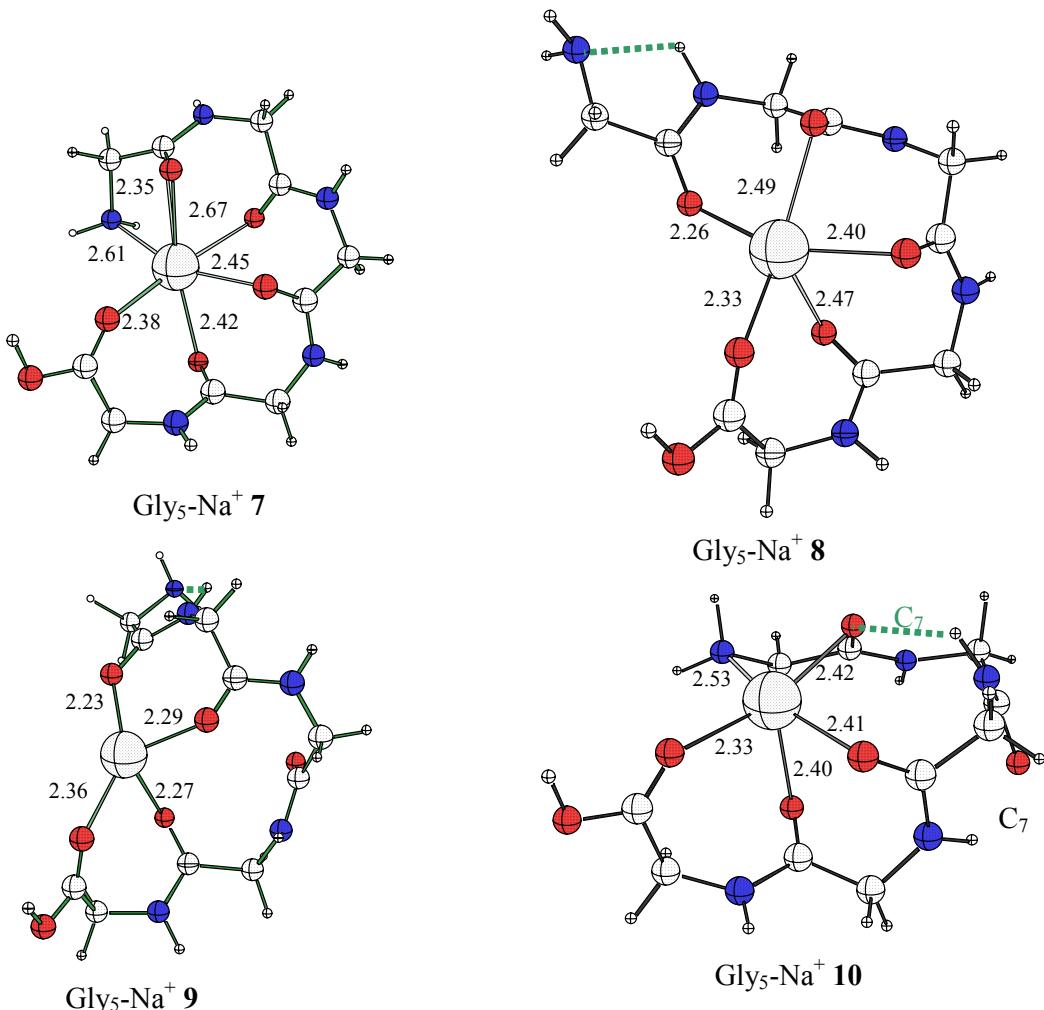


Figure VI-7 : Conformères de plus basse énergie de  $\text{Gly}_5\text{-Na}^+$ .



Suite Figure VI-7 : Conformères de plus basse énergie de  $\text{Gly}_5\text{-Na}^+$ .

Nom $\text{G}_6\text{Na}^+$	$\Delta r\text{H}^{\text{GB2}}$ à 298 K (kJ/mol)	$\Delta r\text{G}^{\text{GB2}}$ à 298 K (kJ/mol)	notation	nombre de ligands	nombre de liaisons hydrogène
1 c4ter	-28.6	-23.3	$\text{O}^1\text{-O}^2\text{-O}^3\text{-O}^4\text{-O}^5\text{-O}^6$	5O+1o	1
2 G <sub>7</sub> -G <sub>8</sub> -A3-29	-23.3	-26.2	$\text{O}^1\text{-C}_7\text{-C}_7\text{-O}^4\text{-O}^5\text{-O}^6$	4O	2
3 A1-24	-21.4	-12.0	$\text{O}^1\text{-}\emptyset\text{-O}^3\text{-o}^4\text{-O}^5\text{-o}^6(+\text{C}_{20})\text{-o}^6\text{H}$	3O+3o	2
4 A2-20	-17.9	-29.7	$\text{O}^1\text{-O}^2\text{-O}^3\text{-O}^4\text{-O}^5\text{-}\emptyset$	5O	1
5 A1-27	-13.6	-10.7	$\text{O}^1\text{-O}^2\text{-O}^3\text{-}\emptyset\text{-O}^5\text{-O}^6$	5O	2
6 G <sub>7</sub> -A2-8	-11.3	-7.4	$\text{O}^1\text{-}\emptyset\text{-O}^3\text{-C}_7\text{-O}^5\text{-O}^6(+\text{C}_{14})$	4O	3
7 A1-13	-10.8	-6.9	$\text{O}^1\text{-}\emptyset\text{-O}^3\text{-C}_7\text{-C}_7\text{-O}^6\text{-O}^6\text{H}$	4O	3
8 A3-1	-9.8	-6.0	$\text{O}^1\text{-C}_7\text{-C}_{10}\text{-O}^4\text{-O}^5\text{-O}^6(+\text{C}_{14})$	4O	4
9 G <sub>7</sub> -c5b	-9.2	-1.2	$\text{o}^1(+\text{C}_{19})\text{-O}^2\text{-O}^3\text{-O}^4\text{-O}^5\text{-O}^6(+\text{C}_{20})$	5O+1o	2
10 A1-23	-7.5	-10.7	$\text{O}^1\text{-O}^2\text{-O}^3\text{-C}_7\text{-O}^5\text{-o}^6$	4O+1o	2
11 A1-25	-6.1	-10.2	$\text{O}^1\text{-O}^2\text{-O}^3\text{-o}^4\text{-O}^5\text{-C}_{20}$	4O+1o	2
16 H $\alpha$ -3	0.0	0.0	$\text{C}_{13}(+\text{C}_{10})\text{-C}_{13}\text{-O}^3\text{-O}^4\text{-O}^5\text{-O}^6$	4O	4

Tableau VI-2 : Energies d'interaction métal-peptide, grandeurs thermodynamiques à 298 K relatives à la structure H  $\alpha$ -3 et principales caractéristiques géométriques des complexes  $\text{Gly}_6\text{-Na}^+$ .

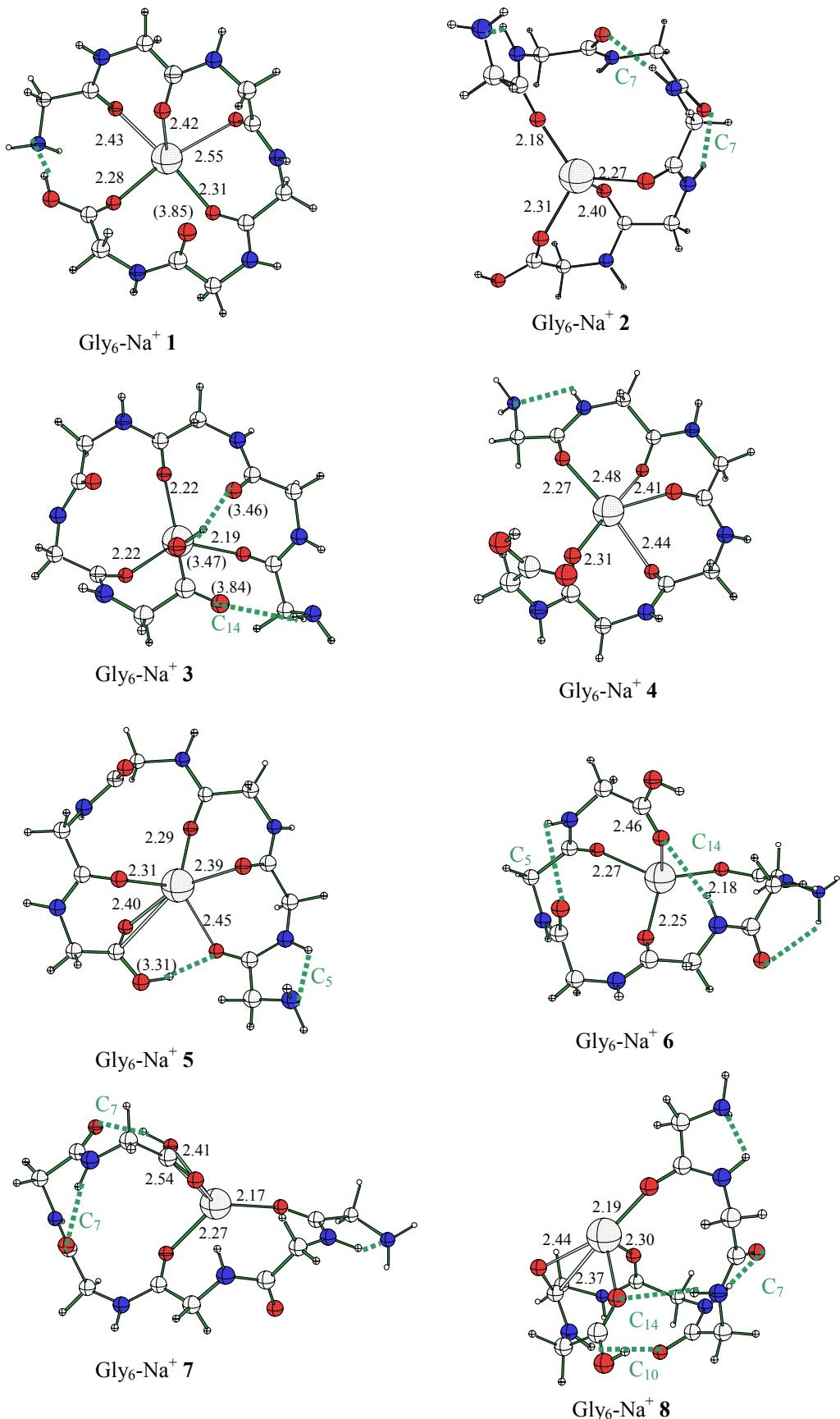
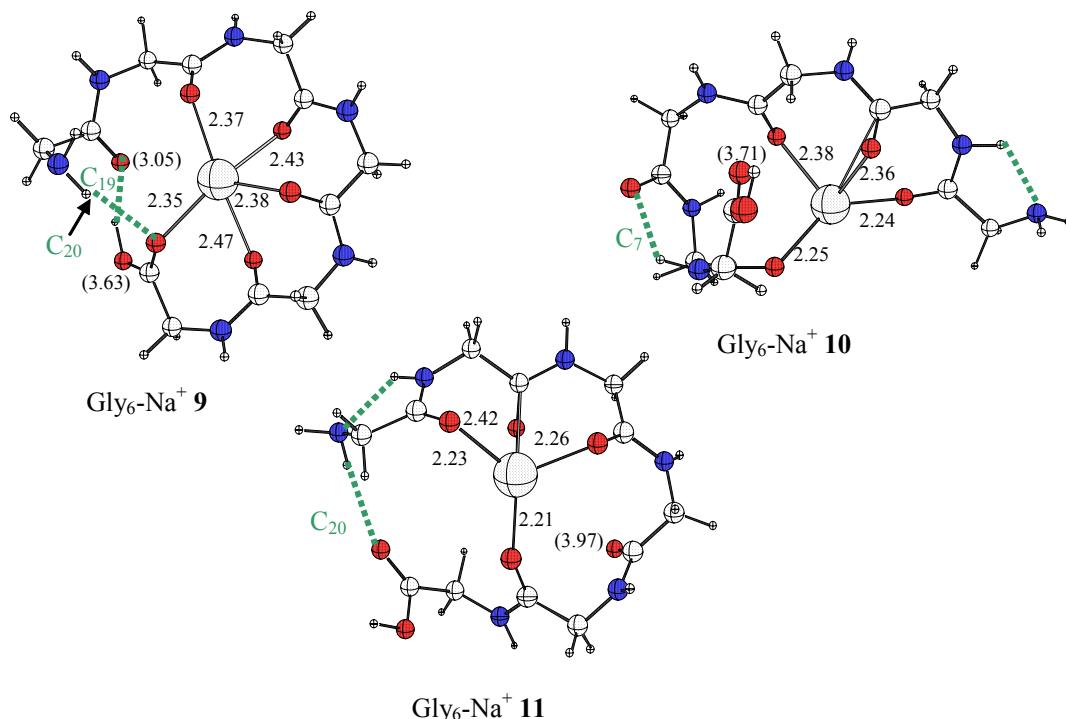


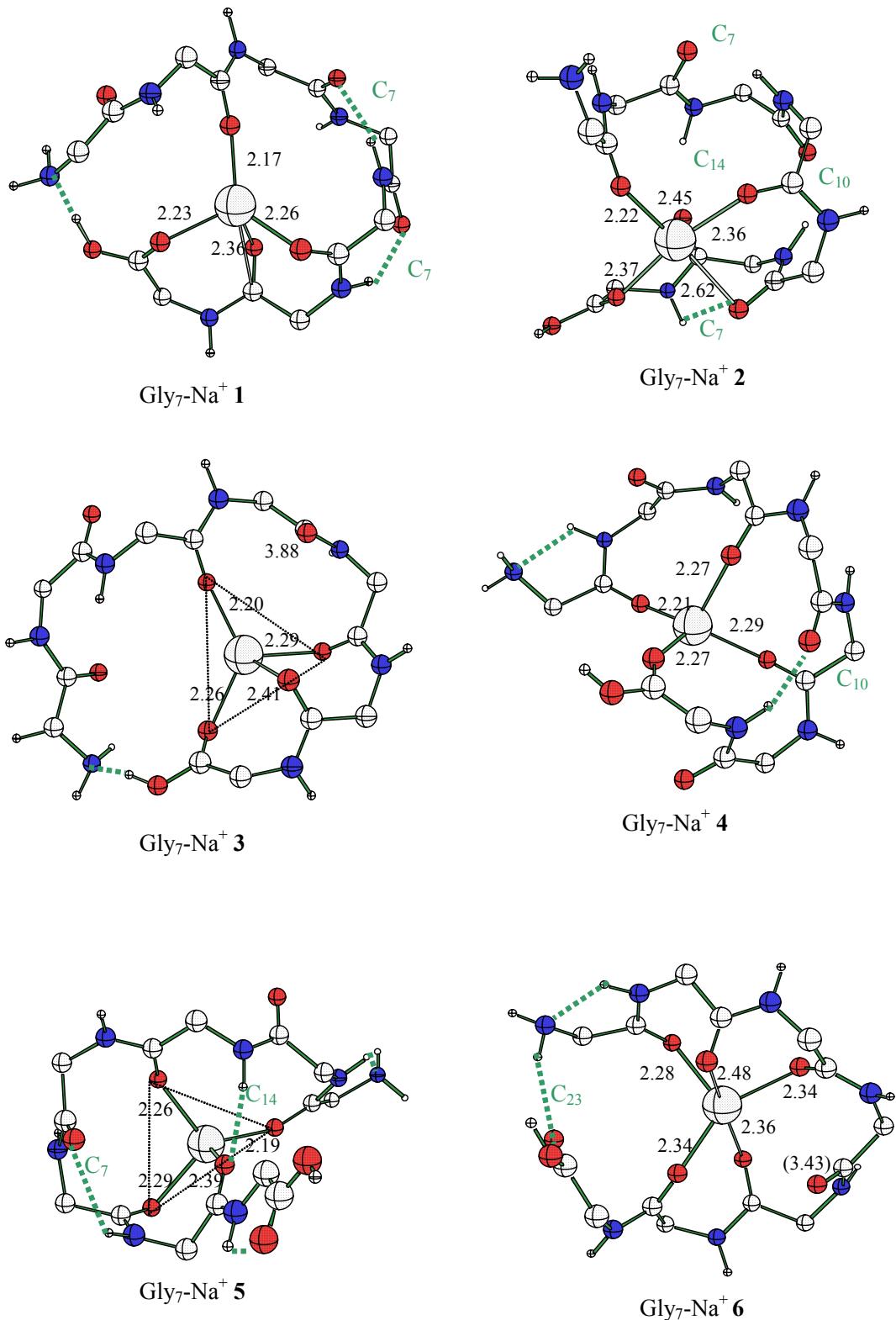
Figure VI-8 : Conformères de plus basse énergie de  $\text{Gly}_6\text{-Na}^+$ .



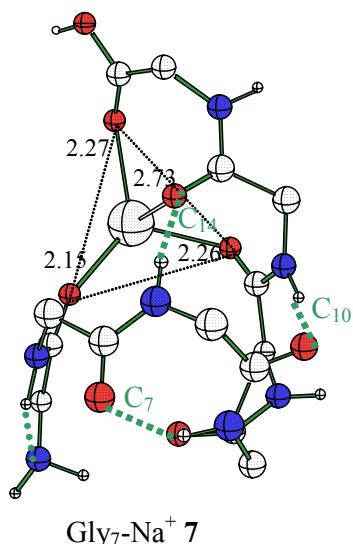
Suite Figure VI-8 : Conformères de plus basse énergie de  $\text{Gly}_6\text{-Na}^+$ .

Nom $\text{G}_7\text{Na}^+$	$\Delta r\text{H}^{\text{GB2}}$ à 298 K (kJ/mol)	$\Delta r\text{G}^{\text{GB2}}$ à 298 K (kJ/mol)	notation	nombre de ligands	nombre de liaisons hydrogène
1 G <sub>8</sub> -A3-29	-32.6	-26.2	$\emptyset\text{-O}^2\text{-C}_7\text{-C}_7\text{-O}^5\text{-O}^6\text{-O}^7$	4O	3
2 A3-2'	-17.3	-17.6	$\text{O}^1\text{-C}_7\text{-C}_{10}\text{-O}^4\text{-O}^5\text{(+C}_7\text{)-O}^6\text{(+C}_{14}\text{)-O}^7$	5O	5
3 c4trans	-15.9	-22.7	$\emptyset\text{-}\emptyset\text{-O}^3\text{-}\emptyset\text{-O}^5\text{-O}^6\text{-O}^7$	4O	1
4 G <sub>8</sub> -A1-11	-14.5	-23.7	$\text{O}^1\text{-}\emptyset\text{-O}^3\text{-C}_{10}\text{-O}^5\text{-}\emptyset\text{-O}^7$	4O	2
5 A2-8	-12.6	-23.7	$\text{O}^1\text{-}\emptyset\text{-O}^3\text{-C}_7\text{-O}^5\text{-O}^6\text{(+C}_{14}\text{)-C}_5$	4O	4
6 c5b	-9.9	-21.6	$\text{O}^1\text{-O}^2\text{-O}^3\text{-O}^4\text{-O}^5\text{-O}^6\text{-C}_{23}$	5O+1o	2
7 A3-1	-9.2	-6.0	$\text{O}^1\text{-C}_7\text{-C}_{10}\text{-}\emptyset\text{-O}^5\text{-O}^6\text{(+C}_{14}\text{)-O}^7$	4O	4
15 H $\alpha$ -3	0.0	0.0	$\text{C}_{13}\text{(+C}_{10}\text{)-C}_{13}\text{-C}_{13}\text{-O}^4\text{-O}^5\text{-O}^6\text{-O}^7$	4O	5

Tableau VI-3 : Energies d'interaction métal-peptide, grandeurs thermodynamiques à 298 K relatives à la structure H  $\alpha$ -3 et principales caractéristiques géométriques des complexes  $\text{Gly}_7\text{-Na}^+$ .



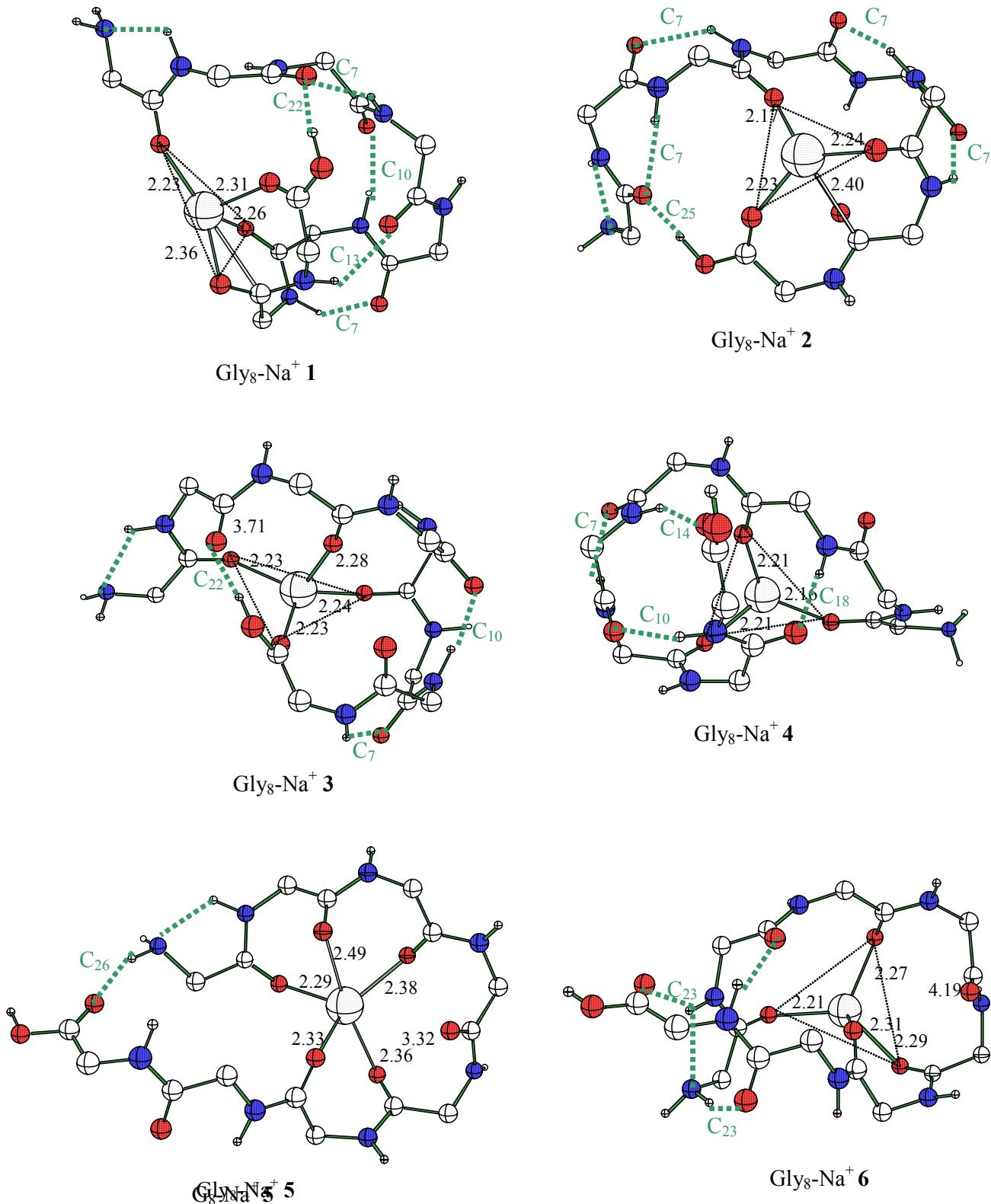
**Figure VI-9 :** Conformères de plus basse énergie de  $\text{Gly}_7\text{-Na}^+$ . Les hydrogènes des groupes  $\text{CH}_2$  ne sont pas représentés.



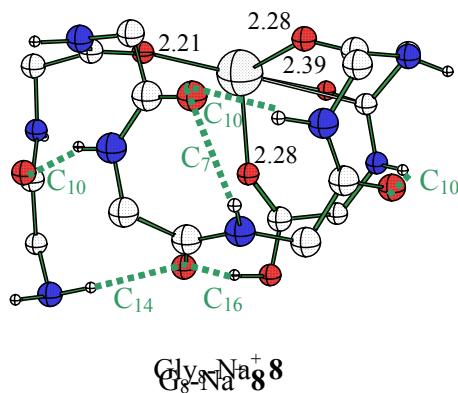
Suite Figure VI-9 : Conformères de plus basse énergie de  $\text{Gly}_7\text{-Na}^+$ .

Nom $\text{G}_8\text{Na}^+$	$E^{\text{int}}$ metal-peptide	$\Delta rH^{\text{GB2}}$ à 298 K (kJ/mol)	$\Delta rG^{\text{GB2}}$ à 298 K (kJ/mol)	notation	nombre de ligands	nombre de liaisons hydrogène
1 G <sub>9</sub> -A4-24	-424.3	-19.5	-5.7	$\text{O}^1\text{-C}_7\text{(+C}_{22}\text{)-C}_{10}\text{-C}_{13}\text{-C}_7\text{-O}^6\text{-O}^7\text{-O}^8$	40	6
2 A3-29		-17.3	-8.4	$\text{C}_7\text{(+C}_{25}\text{)-C}_7\text{-O}^3\text{-C}_7\text{-C}_7\text{-O}^6\text{-O}^7\text{-O}^8$	40	6
3 A1-11		-14.9	-11.1	$\text{O}^1\text{-C}_{22}\text{-O}^3\text{-C}_{10}\text{-O}^5\text{-C}_7\text{-}\emptyset\text{-O}^8$	40	4
4 A1-9		-11.1	-5.4	$\text{O}^1\text{-}\emptyset\text{-O}^3\text{-C}_7\text{-C}_{10}\text{-O}^6\text{-C}_{18}\text{-C}_{14}$	30	5
5 c6go		-6.1	-10.1	$\text{O}^1\text{-O}^2\text{-O}^3\text{-o}^4\text{-O}^5\text{-O}^6\text{-}\emptyset\text{-C}_{26}$	5O+1o	2
6 A1-17		-2.3	0.1	$\text{O}^1\text{-C}_{19}\text{-O}^3\text{-}\emptyset\text{-O}^5\text{-O}^6\text{-C}_{23}\text{-C}_{23}$	40	4
7 H $\alpha$ -3	-415.8	0.0	0.0	$\text{C}_{13}\text{(+C}_{10}\text{)-C}_{13}\text{-C}_{13}\text{-C}_{13}\text{-O}^5\text{-O}^6\text{-O}^7\text{-O}^8$	40	6
8 A4-5		0.5	12.4	$\text{C}_{10}\text{-O}^2\text{-C}_{10}\text{(+C}_7\text{)-C}_{14}\text{(+C}_{16}\text{)-C}_{10}\text{-O}^6\text{-O}^7\text{-O}^8$	40	6

Tableau VI-4 : Energies d'interaction métal-peptide, grandeurs thermodynamiques à 298 K relatives à la structure H  $\alpha$ -3 et principales caractéristiques géométriques des complexes  $\text{Gly}_8\text{-Na}^+$ .



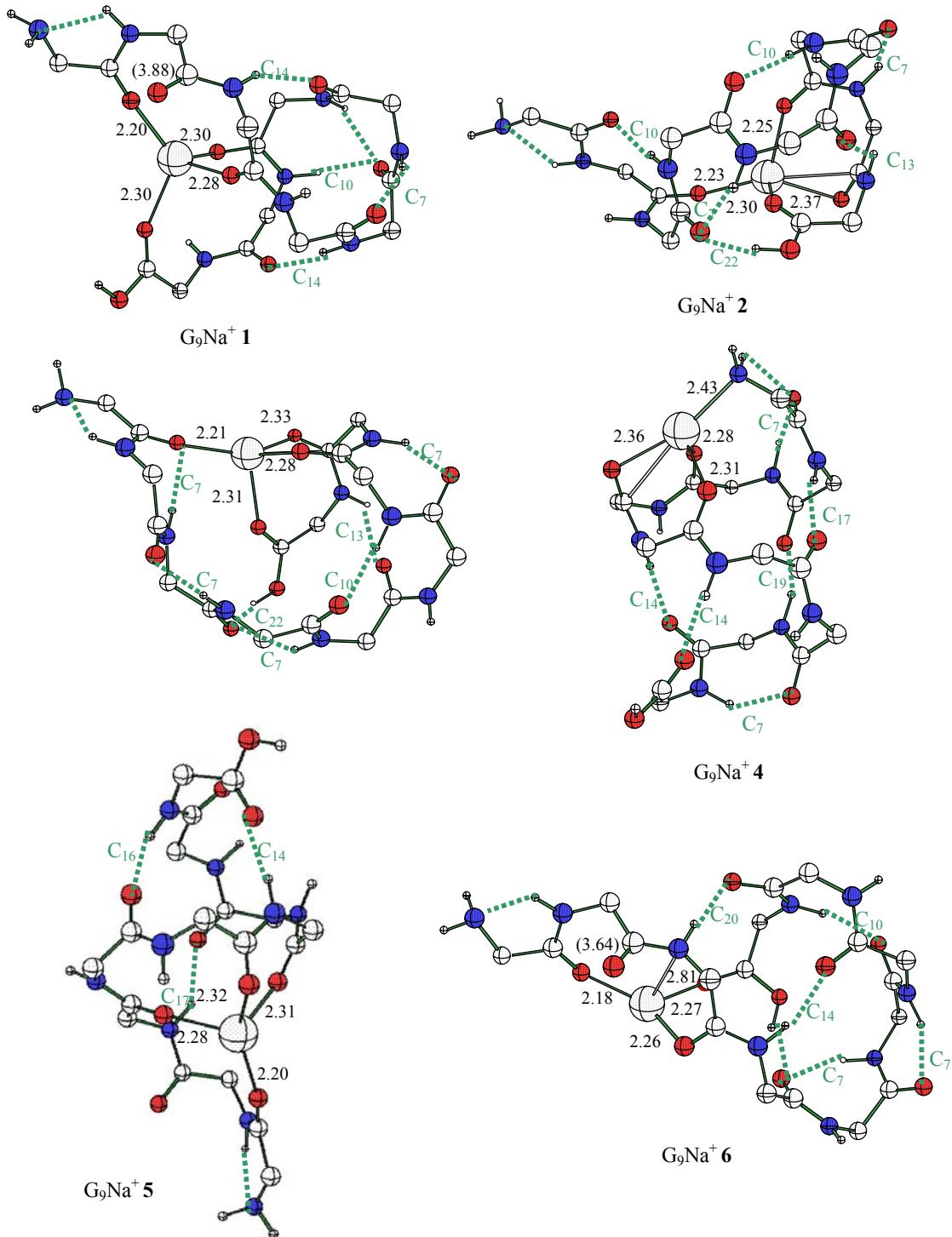
**Figure VI-10 :** Conformères de plus basse énergie de  $\text{Gly}_8\text{-Na}^+$ . Les hydrogènes des groupes  $\text{CH}_2$  ne sont pas représentés.



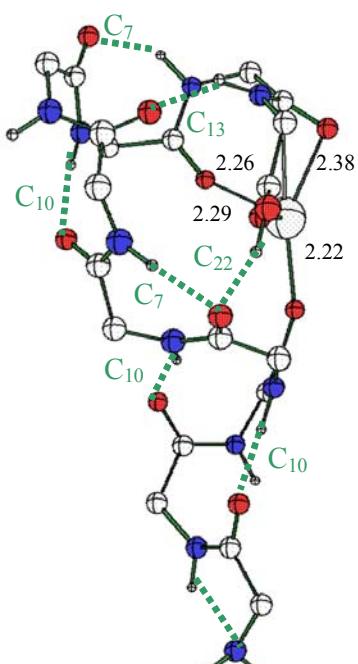
Suite Figure VI-10 : Conformères de plus basse énergie de  $\text{Gly}_8\text{-Na}^+$ .

	Nom $\text{G}_9\text{Na}^+$	$\Delta r\text{H}^{\text{GB2}}$ à 298 K (kJ/mol)	$\Delta r\text{G}^{\text{GB2}}$ à 298 K (kJ/mol)	notation	nombre de ligands	nombre de liaisons hydrogène
1	$\text{G}_{10}\text{-A3-2}$	-23.8	-21.1	$\text{O}^1\text{-o}^2\text{-O}^3\text{-C}_7\text{-C}_{10}(+\text{C}_7)\text{-C}_{14}\text{-O}^7\text{-C}_{14}\text{-O}^9$	4O+1o	6
2	$\text{G}_8\text{-A3-8}$	-18.0	-3.5	$\text{C}_{10}\text{-O}^2\text{-C}_7(+\text{C}_{22})\text{-C}_{10}\text{-C}_{13}\text{-C}_7\text{-O}^7\text{-O}^8\text{-O}^9$	4O	7
3	A4-25	-16.8	-1.3	$\text{O}^1(+\text{C}_7)\text{-C}_7\text{-C}_7(+\text{C}_{22})\text{-C}_{10}\text{-C}_{13}\text{-C}_7\text{-O}^7\text{-O}^8\text{-O}^9$	4O	8
4	$\text{G}_{10}\text{-A2-12}$	-8.7	-0.2	$\text{N}^1\text{-C}_7(+\text{C}_5)\text{-C}_{19}\text{-O}^3\text{-O}^4\text{-O}^5\text{-C}_{17}\text{-C}_7\text{-C}_{14}\text{-C}_{14}$	4O,1N <sup>1</sup>	7
5	$\text{G}_{10}\text{-A4-5}$	-6.9	-5.7	$\text{O}^1\text{-}\emptyset\text{-O}^3\text{-C}_{16}\text{-O}^5\text{-O}^6\text{-C}_{17}\text{-}\emptyset\text{-C}_{14}$	4O	4
6	A1-14	-1.8	4.7	$\text{O}^1\text{-o}^2\text{-O}^3\text{-C}_7(+\text{C}_{19})\text{-C}_7\text{-C}_{10}\text{-C}_{14}\text{-C}_{20}\text{-O}^9$	3O,1N <sup>3</sup> +1o	7
10	H $\alpha$ -3	0.0	0.0	$\text{C}_{13}(+\text{C}_{10})\text{-C}_{13}\text{-C}_{13}\text{-C}_{13}\text{-C}_{13}\text{-O}^6\text{-O}^7\text{-O}^8\text{-O}^9$	4O	7

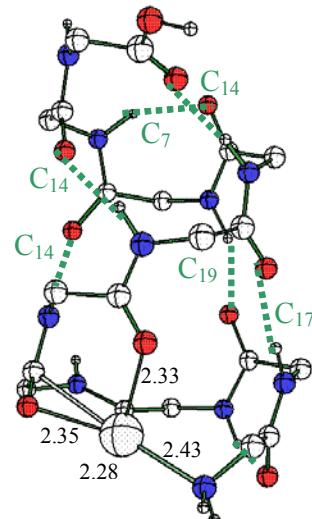
Tableau VI-5 : Energies d'interaction métal-peptide, grandeurs thermodynamiques à 298 K relatives à la structure H  $\alpha$ -3 et principales caractéristiques géométriques des complexes  $\text{Gly}_9\text{-Na}^+$ .



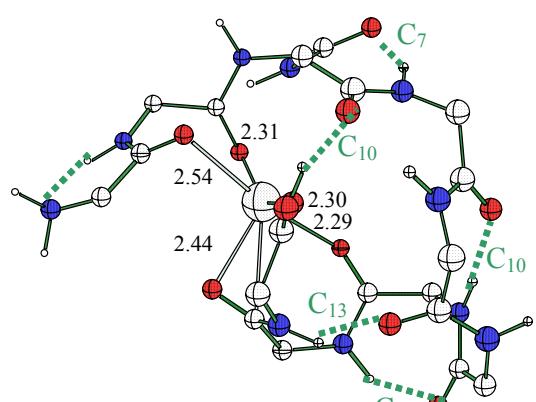
**Figure VI-11 :** Conformères de plus basse énergie de  $\text{Gly}_9\text{-Na}^+$ . Les hydrogènes des groupes  $\text{CH}_2$  ne sont pas représentés.



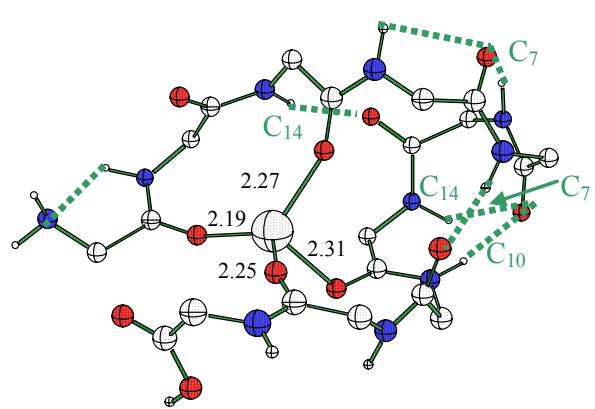
$\text{Gly}_{10}\text{-Na}^+ 1$



$\text{Gly}_{10}\text{-Na}^+ 2$

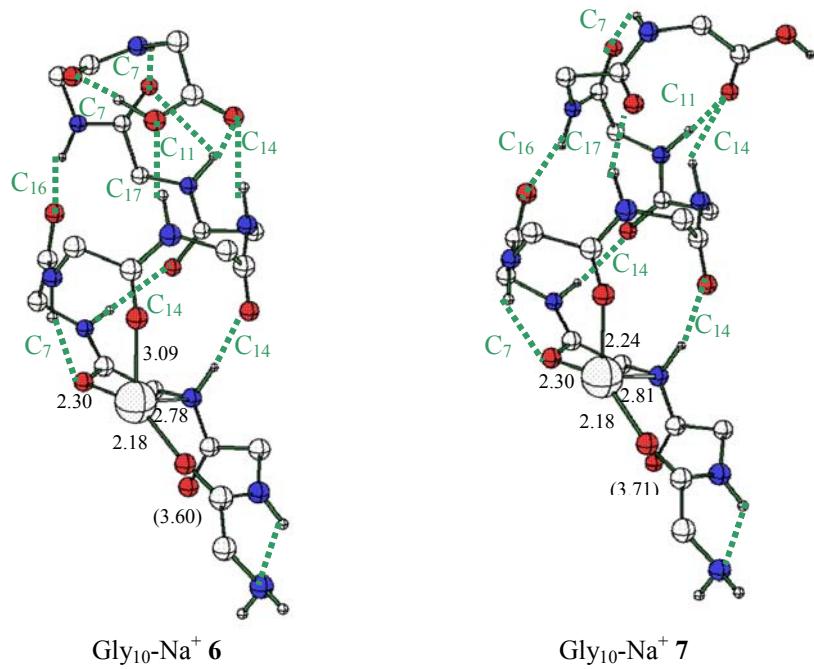


$\text{Gly}_{10}\text{-Na}^+ 3$



$\text{Gly}_{10}\text{-Na}^+ 4$

**Figure VI-12 :** Conformères de plus basse énergie de  $\text{Gly}_{10}\text{-Na}^+$ . Les hydrogènes des groupes  $\text{CH}_2$  ne sont pas représentés.



**Suite Figure VI-12 :** Conformères de plus basse énergie de  $\text{Gly}_{10}\text{-Na}^+$ .

Nom $\text{G}_{10}\text{Na}^+$	$E^{\text{int}}$ metal-peptide	$\Delta rH^{\text{GB2}}$ à 298 K (kJ/mol)	$\Delta rG^{\text{GB2}}$ à 298 K (kJ/mol)	notation	nombre de ligands	nombre de liaisons hydrogène
1 G <sub>9</sub> -G <sub>8</sub> -A3-8	451.4	-17.2	-2.0	$\text{C}_{10}\text{-C}_{10}\text{-O}^3\text{-C}_7(+\text{C}_{22})\text{-C}_{10}\text{-C}_{13}\text{-C}_7$ $\text{-O}^8\text{-O}^9\text{-O}^{10}$	4O	8
2 A2-12ter		-5.7	8.7	$\text{N}^1\text{-C}_7\text{-C}_{19}\text{-O}^3\text{-O}^4\text{-O}^5\text{-C}_{17}\text{-C}_7$ $\text{-C}_{14}\text{-C}_{14}\text{-C}_{14}$	3O, 1N <sup>1</sup>	7
3 G <sub>9</sub> -A4-25	430.9	-4.9	-5.2	$\text{O}^1(+\text{C}_{10})\text{-O}^2\text{-C}_7\text{-C}_{10}(+\text{C}_{22})\text{-C}_{10}\text{-C}_{13}\text{-C}_7$ $\text{-O}^8\text{-O}^9\text{-O}^{10}$	5O	8
4 A3-2	467.0	-4.6	-5.5	$\text{O}^1\text{-}\emptyset\text{-O}^3\text{-C}_7(+\text{C}_5)\text{-C}_{10}(+\text{C}_7)\text{-C}_{14}\text{-O}^7$ $\text{-C}_{14}\text{-O}^9\text{-}\emptyset$	4O	7
5 H $\alpha$ -3	443.5	0.0	0.0	$\text{C}_{13}(+\text{C}_{10})\text{-C}_{13}\text{-C}_{13}\text{-C}_{13}\text{-C}_{13}\text{-C}_{13}\text{-C}_7$ $\text{-O}^8\text{-O}^9\text{-O}^{10}$	4O	8
6 A4-5		0.4	13.5	$\text{O}^1\text{-}\emptyset\text{-O}^3(+\text{C}_7)\text{-C}_{16}\text{-O}^5\text{-C}_{14}\text{-C}_{14}$ $\text{-C}_7(+\text{C}_5)\text{-C}_7\text{-C}_{14}(+\text{C}_{11})$	3O, 1N <sup>3</sup>	11
7 A3-1		3.1	3.5	$\text{O}^1\text{-}\emptyset\text{-O}^3(+\text{C}_7)\text{-C}_{16}\text{-O}^5\text{-C}_{14}\text{-C}_{14}$ $\text{-C}_7(+\text{C}_5)\text{-C}_{17}\text{-C}_{14}(+\text{C}_{11})$	3O, 1N <sup>3</sup>	10

**Tableau VI-6 :** Energies d'interaction métal-peptide, grandeurs thermodynamiques à 298 K relatives à la structure H  $\alpha$ -3 et principales caractéristiques géométriques des complexes  $\text{Gly}_{10}\text{-Na}^+$ .

### III.B.3. Coordination au métal

Les tableaux VI-1 à VI-6 indiquent le nombre et les positions des carbonyles coordinés au sodium. On constate que dans toutes les structures les plus basses en énergie, le nombre de sites de coordination varie de 3 à 6. La majeure partie des complexes présente 4 ou 5 carbonyles ou amine chélatés au sodium, seules les structures  $\text{Gly}_6\text{-Na}^+$  **3** et  $\text{Gly}_8\text{-Na}^+$  **4** sont tridentates, et seule  $\text{Gly}_5\text{-Na}^+$  **7** est hexadentate. Ceci reste en bon accord avec les résultats expérimentaux en phase condensée. En effet, l'étude de banques de données basées sur la structure cristallographique de protéines montre que les protéines sodiées ont un nombre de coordination qui varie de 4 à 8 avec une part plus importante pour les protéines hexadentates puis tétradentates.<sup>17-19</sup> Les différences observées entre cette étude statistique et nos résultats peuvent être dues à plusieurs facteurs, parmi lesquels la taille de nos peptides modèles, limitée à 10 résidus, ainsi que l'absence de chaînes latérales oxygénées qui contribuent à la coordination de  $\text{Na}^+$  dans les protéines sodiées. Par ailleurs, l'étude en phase gazeuse des complexes  $\text{Na}^+-(\text{H}_2\text{O})_n(\text{C}_6\text{H}_4\text{F}_2)_n$  par spectroscopie IRPD (dissociation par absorption photonique infrarouge) par l'équipe de J. M. Lisy montre que la première sphère de coordination de  $\text{Na}^+$  est constituée de quatre molécules d'eau contrairement à la phase aqueuse, où le nombre de coordination est six.<sup>20</sup> Ce comportement a été confirmé par des calculs de chimie quantique menés par Kim *et al.* et Feller *et al.* sur des clusters hydratés de sodium. Ces calculs montrent que dans le complexe  $\text{Na}^+-(\text{H}_2\text{O})_6$ , la structure 4+2, comprenant quatre molécules en première sphère de coordination et deux en seconde, est énergétiquement favorisée par rapport aux structures 5+1 et 6+0 (la structure 5+1 étant plus stable que 6+0).<sup>21, 22</sup> Dans le cas de  $\text{Na}^+-(\text{H}_2\text{O})_5$ , les structures 4+1 et 5+0 sont iso-énergétiques. Les complexes  $\text{Na}^+-(\text{MeOH})_n$  ont des structures similaires, où le complexe 6+(n-6) est défavorisé par rapport aux formes 5+(n-5) ou 4+(n-4).<sup>21</sup>

La plupart des longueurs de liaison métal-oxygène en première sphère sont dans une gamme de 2.20 à 2.50 Å, ce qui est en accord avec les valeurs observées par diffraction des rayons X : 2.25 Å pour la distance minimale  $\text{Na}^+$ -O carboxylate et 2.33 Å pour la distance moyenne  $\text{Na}^+$ -O carbonyle.<sup>18</sup> On peut par ailleurs remarquer que lorsque le peptide est tétradentate, les longueurs de liaison sont plus faibles que dans le cas de peptides pentadentates. Ainsi dans le cas de  $\text{Gly}_5\text{-Na}^+$ , les conformères pentadentates **1**, **2**, **7** et **8** présentent au moins trois distances  $\text{Na}^+$ -O supérieures à 2.39 Å, alors que dans les conformères **4** et **9** tétradentates, toutes les liaisons métal-oxygène carbonyle sont inférieures à 2.33 Å. Ce comportement est semblable pour les peptides plus longs. Ceci est dû d'une part à la rigidité du squelette peptidique, mais aussi à l'encombrement stérique pour les peptides de grandes tailles.

### III.B.4. Liaisons hydrogène

Le nombre et les variétés de liaisons hydrogène croissent avec le nombre de résidus. Cette tendance était prévisible car plus le nombre de résidus augmente, plus la chaîne peptidique est flexible, d'autant que le nombre de carbonyles complexés au sodium reste de 4 à 5 pour presque tous les  $\text{Gly}_n\text{-Na}^+$ . Chaque structure de  $\text{Gly}_5\text{-Na}^+$  présente 0 à 3 liaisons hydrogène et seul le motif  $\text{C}_7$  est rencontré. Pour  $\text{Gly}_6\text{-Na}^+$ , on rencontre de 1 à 4 liaisons hydrogène, pour  $\text{Gly}_7\text{-Na}^+$  : de 1 à 5, pour  $\text{Gly}_8\text{-Na}^+$  : de 2 à 6. A partir de  $\text{Gly}_9\text{-Na}^+$ , le nombre de liaisons hydrogène augmente plus rapidement : 4 à 8 et enfin 7 à 11 pour  $\text{Gly}_{10}\text{-Na}^+$ . Ainsi la structure  $\text{Gly}_{10}\text{-Na}^+$  **6** possède un enchevêtrement de 11 liaisons hydrogène, avec des motifs  $\text{C}_5$ ,  $\text{C}_7$ ,  $\text{C}_{16}$ , et 7 liaisons successives de type  $\text{C}_{14}$ ,  $\text{C}_7$ ,  $\text{C}_5$ ,  $\text{C}_{17}$  et  $\text{C}_{11}$ .

On peut remarquer que la plupart des structures globulaires sont relativement compactes, jusqu'à  $\text{Gly}_8\text{-Na}^+$ . Ces structures globulaires sont principalement déterminées par l'enroulement du peptide autour de  $\text{Na}^+$ . Par contre à partir de  $\text{Gly}_9\text{-Na}^+$ , il existe des structures qui présentent localement des motifs hélicoïdaux. La structure  $\text{Gly}_{10}\text{-Na}^+$  **1** qui présente une complexation des huit résidus du côté C-terminal identique à celle de  $\text{Gly}_8\text{-Na}^+$  **1**, forme localement un motif en hélice  $3_{10}$  avec les cinq résidus du côté N-terminal. Quant aux structures  $\text{Gly}_9\text{-Na}^+$  **4**,  $\text{Gly}_{10}\text{-Na}^+$  **2**,  $\text{Gly}_{10}\text{-Na}^+$  **6** et  $\text{Gly}_{10}\text{-Na}^+$  **7**, elles ont une forme globale en hélice qui n'est pas périodique, où le sodium est complexé du côté N terminal. Ceci peut paraître contradictoire avec la discussion de la partie I.A.2.b où nous avons mentionné que les hélices  $\alpha$ ,  $3_{10}$  ou  $\pi$  pouvaient être stabilisées par interaction électrostatique avec un cation du côté C-terminal. Cependant dans le cas de ces quatre structures, la majorité des carbonyles sont orientés vers l'amine terminale (contrairement aux hélices périodiques habituelles) et forment des liaisons hydrogène de type  $\text{C}_{14}$  et  $\text{C}_{17}$  :  $\text{N}^i \rightarrow \text{O}^{i+3,4}$ . Leur macro-dipôle étant donc orienté dans le sens inverse de celui des hélices classiques, il est normal que l'interaction avec le cation soit du côté N-terminal. On peut alors imaginer des structures en hélices périodiques avec des liaisons hydrogène orientées de manière à favoriser la complexation de  $\text{Na}^+$  du côté N-terminal (Figure VI-6). Nous avons pu localiser sur la surface de potentiel une structure en hélice régulière qui ne présente que des motifs  $\text{C}_{14}$ , cependant pour  $\text{G}_{10}\text{-Na}^+$ , elle se situe 84.8 kJ/mol au dessus de l'hélice  $\alpha$ -**3**. Une hélice régulière à motif  $\text{C}_{17}$  a également été décrite,<sup>23</sup> mais nos essais pour la localiser ont conduit à la structure en  $\text{C}_{14}$ .

### III.B.5. Conclusion

En conclusion, les structures les plus stables des  $\text{Gly}_n\text{-Na}^+$  ( $n=5-10$ ) possèdent les caractéristiques suivantes :

- pour les petites valeurs de  $n$ , la stabilité est contrôlée par la maximisation des interactions métal-peptide, principalement de type  $\text{Na}^+-\text{O}$ .
- lorsque la coordination du sodium tend à être saturée, la construction d'un réseau de liaisons hydrogène dans le peptide devient le facteur principal de stabilité.
- la quasi-totalité des structures les plus stables présentent 4 ou 5 sites de chélation au sodium. Il est probable que la capacité du peptide à lier au sodium des oxygènes de plus en plus distants dans la séquence, donc impliquant une énergie de déformation plus faible du peptide, est aussi un facteur de stabilité.
- A partir de neuf résidus, le réseau de liaisons hydrogène peut se construire à distance du métal, et on voit apparaître des fragments de structure en hélice.

On peut enfin remarquer que la « propagation » de structures en ajoutant (ou enlevant) un résidu à  $\text{Gly}_n\text{-Na}^+$  pour construire une structure de  $\text{Gly}_{n+1}\text{-Na}^+$  (ou  $\text{Gly}_{n-1}\text{-Na}^+$ ), fait apparaître qu'aucun type de structure n'est le plus stable pour plusieurs valeurs de  $n$ . Les modèles considérés sont probablement trop petits, vue l'importance de la construction progressive de la coordination du sodium. On peut néanmoins constater une tendance nette pour  $\text{Gly}_9\text{-Na}^+$  et  $\text{Gly}_{10}\text{-Na}^+$  : les quatre structures les plus stables de ces deux espèces sont des mêmes types, mais avec des énergies relatives différentes dans les deux cas. En particulier, la structure la plus stable de  $\text{Gly}_{10}\text{-Na}^+$ , qui est la seconde pour  $\text{Gly}_9\text{-Na}^+$ , présente un début d'hélice  $3_{10}$  à distance de  $\text{Na}^+$  ; il serait intéressant de savoir si ce type de structure composite peut rester plus stable que l'hélice  $\alpha$  pour de plus grandes tailles de peptide.

### III.C. Energies relatives des conformères globulaires

Afin de comprendre les tendances des énergies relatives des différents conformères de  $\text{Gly}_n\text{-Na}^+$  pour chaque valeur de  $n$ , nous avons fait l'hypothèse « minimale » que leur stabilité ne dépend que du nombre de liaisons métal-peptide et du nombre de liaisons hydrogène, sans aucune autre précision. Ces critères nous permettent de comprendre qualitativement l'ordre énergétique des sept conformères les plus bas de  $\text{Gly}_5\text{-Na}^+$  : les structures à cinq sites de chélation sont les plus stables, et parmi les groupes possédant un même nombre de sites, les structures les plus stables sont celles qui comportent le plus grand nombre de liaisons hydrogène. Par contre, il n'est pas possible d'expliquer la différence d'énergie de 21.4 kJ/mol entre les complexes **1** et **8**, qui possèdent la même liaison hydrogène et tous les oxygènes des carbonyles complexés. De même pour  $\text{Gly}_7\text{-Na}^+$ , le conformère **1**, 15.3 kJ/mol plus stable que le **2**, possède une interaction  $\text{Na}^+-\text{O}$  et deux liaisons hydrogène en moins. Il existe de nombreux autres cas qui mettent en défaut ce modèle.

Pour mieux décrire l'interaction métal-ligand, nous avons évalué cette contribution par la différence d'énergie entre le complexe et la somme de l'énergie du peptide dans la conformation du complexe et celle du cation  $\text{Na}^+$  ; quelques valeurs sont disponibles dans les tableaux VI-1 à VI-6. Dans le cas de  $\text{Gly}_5\text{-Na}^+$ , les conformères **1** et **8** avec cinq oxygènes coordinés au sodium, ont des énergies d'interaction de 428 et 423 kJ/mol, valeurs nettement supérieures à celles des conformères **3** et **4** qui possèdent seulement quatre oxygènes en interaction avec  $\text{Na}^+$ . Dans le cas de  $\text{Gly}_{10}\text{-Na}^+$ , cette tendance est inversée puisque le conformère **3** pentadentate a une énergie d'interaction nettement plus faible que celle des conformères tétradentates **1**, **4** et **5**. La simple connaissance du nombre de sites de chélation n'est donc pas nécessairement révélatrice de la stabilisation due à l'interaction  $\text{Na}^+$ -peptide. Il faut notamment tenir compte des interactions à distance entre le cation et la partie du peptide qui ne lui est pas liée. Il s'agit d'interactions de polarisation et électrostatiques, qui décroissent assez lentement quand la distance au cation croît. Nous ne disposons malheureusement d'aucune méthode d'évaluation qualitative de ces interactions à distance. De plus, ces termes ne permettent pas d'expliquer dans le cas de  $\text{Gly}_5\text{-Na}^+$ , la différence d'énergie de 21.4 kJ/mol entre les complexes **1** et **8**, qui possèdent la même liaison hydrogène et une énergie d'interaction métal-peptide similaire.

Pour comprendre l'énergie relative des différents conformères, il faut donc tenir compte, en plus de l'énergie d'interaction entre le peptide et le cation, de l'énergie de déformation du peptide pour l'amener de sa structure optimale isolée à celle qu'il adopte dans le complexe. Pour essayer d'en obtenir une approche qualitative, on peut séparer cette énergie de déformation en deux contributions, d'une part l'énergie de torsion du squelette carboné et d'autre part l'énergie de stabilisation due aux liaisons hydrogène. La première contribution ne peut être évaluée simplement. Nous pourrions en première approximation considérer qu'elle est constante pour tous les conformères. Pour la deuxième contribution, on peut faire l'hypothèse que l'énergie d'une structure peut être évaluée par rapport à une référence fictive, par la somme de énergies des liaisons hydrogène qu'elle comporte. Cependant, cette contribution est difficile à évaluer numériquement. On sait que l'énergie d'une liaison hydrogène se situe entre 5 et 20 kJ/mol, nous avons évalué l'énergie d'interaction du dimère de la N-méthylacétamide, orienté de manière à former une liaison N-H $\cdots$ O=C, comme étant de 12 kJ/mol, ce qui donne un ordre de grandeur d'énergie d'une liaison hydrogène optimale dans un peptide. Cependant les liaisons hydrogène de type C<sub>5</sub>, C<sub>7</sub>, C<sub>10</sub> et C<sub>13</sub> ne peuvent être considérées comme énergétiquement équivalentes car dans les trois premiers cas, elles n'ont pas une orientation optimale. Actuellement, nous n'avons donc pas pu trouver d'explication systématique de l'ordre énergétique des différents conformères globulaires au sein d'une même série  $\text{Gly}_n\text{-Na}^+$ . Cependant une interprétation qualitative permet d'appréhender l'évolution entre les séries  $\text{Gly}_n\text{-Na}^+$ .

On peut remarquer d'une part que les conformères de plus basses énergies de Gly<sub>5</sub>-Na<sup>+</sup> et Gly<sub>6</sub>-Na<sup>+</sup>, présentent une coordination pentadentate avec une seule liaison hydrogène, alors que pour les complexes Gly<sub>7</sub>-Na<sup>+</sup>, Gly<sub>8</sub>-Na<sup>+</sup> et Gly<sub>9</sub>-Na<sup>+</sup> la coordination est tétradentate avec un nombre croissant de liaisons hydrogène. Ceci semble indiquer que dans le cas de petits peptides, l'interaction métal-ligand est plus stabilisante que la formation de liaisons hydrogène. Dans le cas des peptides plus grands, le nombre de coordination moins important permet de minimiser les contraintes de torsion de la chaîne peptidique, tout en maximisant le nombre de liaisons hydrogène. On peut observer une transition à partir de Gly<sub>10</sub>-Na<sup>+</sup>. En effet, même si l'hélice  $\alpha$ -3 reste 18.2 kJ/mol plus haute en énergie que la structure **1**, cinq des sept conformères de plus basse énergie présentent des structures localement en hélice. La structure globale du conformère Gly<sub>10</sub>-Na<sup>+</sup> **1** peut être décomposée en deux fragments, d'une part une structure globulaire du côté C-terminal identique à celle de Gly<sub>8</sub>-Na<sup>+</sup> **1**, d'autre part les cinq premiers résidus du côté N-terminal qui forment une hélice 3<sub>10</sub>. Quant aux structures Gly<sub>9</sub>-Na<sup>+</sup> **4** et Gly<sub>10</sub>-Na<sup>+</sup> **2, 6 et 7**, elles ont une forme globale en hélice qui n'est pas périodique, où le sodium est complexé du côté N-terminal et présente un nombre important de motifs C<sub>14</sub>.

#### IV. Conclusion et perspectives

L'un des buts de cette étude était de déterminer le nombre  $n_H^{Gly}$  de résidus, tel que la structure  $Gly_{n_H^{Gly}} - Na^+$  soit le minimum global sur la surface de potentiel. Nous avons pour cela étudié les différentes conformations, à la fois globulaires et hélicoïdales, des complexes Gly<sub>n</sub>-Na<sup>+</sup>, n variant de 5 à 10, en couplant des recherches conformationnelles de type Monte Carlo basées sur le champ de force AMBER et la fonctionnelle de la densité ri-BLYP. Cette méthode montre que  $n_H^{Gly}$  est supérieur à 10. Cependant il est intéressant de noter que pour Gly<sub>10</sub>-Na<sup>+</sup>, le conformère le plus bas en énergie présente une structure globulaire autour du sodium et un domaine de cinq résidus en hélice 3<sub>10</sub>. Nous avons observé que la structure en hélice  $\alpha$  était moins stable que l'hélice 3<sub>10</sub> pour Gly<sub>5</sub>-Na<sup>+</sup>, situation inversée pour Gly<sub>6</sub>-Na<sup>+</sup>. Il est donc probable qu'à partir de Gly<sub>11</sub>-Na<sup>+</sup>, il existe une conformation relativement basse en énergie qui présente un domaine en hélice  $\alpha$ .

A ce stade, trois perspectives sont envisagées. Premièrement, nous n'avons pas pu tirer de tendance systématique pour comprendre la stabilité des différents conformères. Il serait donc intéressant de déterminer une méthode d'analyse des deux grandes contributions énergétiques : l'énergie d'interaction entre le peptide et le cation, et l'énergie de déformation du peptide, en définissant une référence appropriée qui pourrait être le brin  $\beta$ , et en établissant une valeur standard pour chaque type de liaison hydrogène présente.

D'autre part il serait intéressant de poursuivre l'étude des conformères de Gly<sub>n</sub>-Na<sup>+</sup> avec une taille croissante, afin de déterminer  $n_H^{Gly}$ .

Enfin, l'étude des conformères de Ala<sub>n</sub>-Na<sup>+</sup>, 5≤n≤10 est en cours. En effet, comme nous l'avons vu au début de ce chapitre, la poly-alanine peut être considérée comme un système modèle de peptide ayant une forte propension à former des hélices. Le nombre  $n_H^{Ala}$  devrait être inférieur à  $n_H^{Gly}$ . Ceci semble confirmé par les expériences de M. Jarrold qui montrent que Ala<sub>12</sub>-Na<sup>+</sup> a une structure en hélice et qui restent équivoques pour n<12, donc  $n_H^{Ala} \leq 12$ . Sur la base des expériences d'hydratation des peptides Ac-Ala<sub>n</sub>-Lys+H<sup>+</sup>, qui indiquent une structure en hélice à partir de n=8 (paragraphe I-b), ces auteurs proposent une transition vers l'hélice  $\alpha$  au voisinage de n=9 pour Ala<sub>n</sub>-Na<sup>+</sup>. Nos résultats préliminaires sur Ala<sub>n</sub>-Na<sup>+</sup> sont compatibles avec une transition vers l'hélice  $\alpha$  pour n=8 ou 9. De plus, si on arrive à évaluer les différentes contributions énergétiques qui stabilisent les différents conformères, on peut espérer comprendre comment le rôle d'un simple méthyle peut transformer une poly-glycine qui possède la plus petite propension à former les hélices, en une poly-alanine qui possède la plus grande.

## **V. Annexe 1 : Calibration des méthodes de chimie quantique sur Gly<sub>6</sub>-Na<sup>+</sup>**

### **V.A. Méthodes et Bases de calculs employés**

Nous avons évalué les performances de plusieurs types de méthodes et de bases atomiques pour la détermination de la géométrie et de l'énergie relative des différents conformères dans le cas de G<sub>6</sub>-Na<sup>+</sup>. Ce système est suffisamment grand pour que les principaux types de structures, en particulier les hélices, soient représentés, et restent de taille raisonnable pour des calculs à des niveaux donnant accès à une bonne précision. Nous avons choisi de calibrer quatre méthodes : HF, MP2, B3LYP et ri-BLYP et quatre bases : PB1 : 6.31G\*, PB2 : SV(P), MB<sup>\*\*</sup> : 6-31+G\*, MB<sup>\*\*\*</sup> : 6-31+G\*\*, GB1 : 6-311+G(2d,2p) et GB2 : TZVPP. Les valeurs d'énergie relative sont rassemblées dans le tableau VI-7. Les bases PB2 et GB2 sont spécifiques de l'approximation de « résolution de l'identité », utilisée en ri-BLYP.

## V.B. Energies

### V.B.1. Influence de la base

#### V.B.1.a. Sur la fonctionnelle de la densité B3LYP

Les résultats au niveau B3LYP/PB1 montrent une surestimation des énergies relatives à celle de la structure la plus stable, comparés à ceux obtenus au niveau B3LYP/MB<sup>\*\*</sup>//B3LYP/PB1. Par contre il existe un très bon accord entre les énergies relatives obtenues aux niveaux B3LYP/MB<sup>\*\*</sup>//B3LYP/PB1 et B3LYP/GB1//B3LYP/PB1 puisque la différence maximale est de 4.3 kJ/mol. Ainsi l'ajout de fonctions diffuses sur les atomes autres que les hydrogènes suffit à faire converger les énergies relatives des conformères.

#### V.B.1.b. Sur la fonctionnelle de la densité ri-BLYP

Dans le cas des calculs ri-BLYP, nous avons utilisé les bases SV(P), de même type que 6-31G\*, et TZVPP, analogue à 6-311+G(2d,2p). Les résultats montrent une surestimation des énergies relatives en l'absence d'orbitales diffuses, comparable à celle observée avec la fonctionnelle B3LYP.

#### V.B.1.c. Sur la méthode *ab initio* MP2

Il est connu que les calculs MP2 convergent beaucoup moins vite avec l'extension de base que les calculs DFT. Nous avons donc réalisé une étude systématique de cette extension. L'ajout de fonctions s et p diffuses sur les atomes lourds (de la base PB1 à la base MB<sup>\*\*</sup>), influence peu les énergies relatives des formes compactes (excepté pour la structure **8**) puisque la déviation maximale est de 5.3 kJ/mol. Par contre les structures qui présentent le plus de liaisons hydrogène : structures hélicoïdales et **8** (cf. figures VI-8 et VI-13) sont fortement stabilisées par l'ajout d'orbitales diffuses. Ainsi l'hélice **H**  $\alpha$ -**1** voit son énergie relative diminuée de 18.1 kJ/mol. L'utilisation de la base MB<sup>\*\*\*</sup>, dans laquelle des fonctions de polarisation sont ajoutées sur les atomes d'hydrogène, ne montre pas de différence significative avec les énergies relatives MP2/MB<sup>\*\*</sup>//HF/PB, la différence d'énergie maximale étant de 3.7 kJ/mol.

La comparaison des énergies relatives obtenues au niveau MP2/GB1//HF/PB1 et MP2//MB<sup>\*\*</sup>//HF/PB1 montre qu'avec la grande base, d'une part les structures compactes (excepté **8**) sont déstabilisées (de 2.7 à 7.8 kJ/mol), d'autre part les énergies des structures hélicoïdales et **8** sont fortement abaissées : pour l'hélice **H**  $\alpha$ -**3**, de 31.5 kJ/mol, pour **8**, de 20.9 kJ/mol. Les hélices **H**  $\alpha$ -**1** et **H**  $\alpha$ -**2** semblent être des cas pathologiques (stabilisation de 140.4 et 132.7 kJ/mol, respectivement). Ceci semble indiquer que l'utilisation de GB1 stabilise fortement d'une part les structures en hélice, d'autre part les structures qui présentent une fonction acide trans (cas de **8**, **H**  $\alpha$ -**1** et **H**  $\alpha$ -**2**). Afin d'évaluer l'influence de

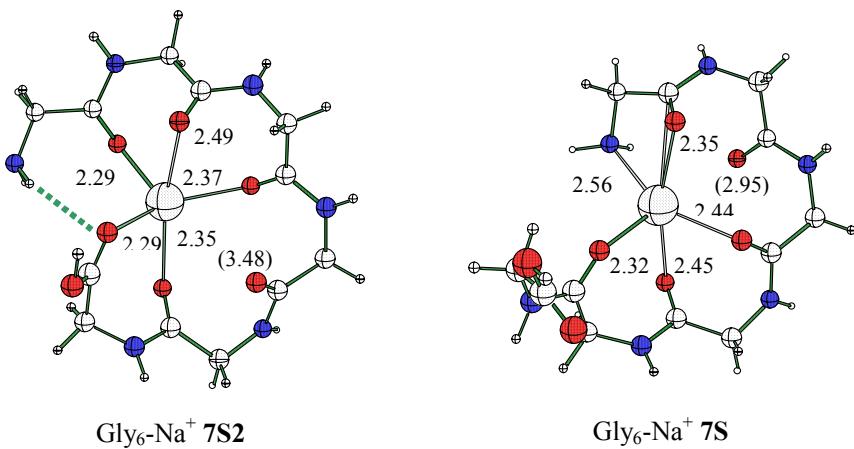
la base sur les calculs d'énergie MP2 de structures présentant des acides carboxyliques cis et trans, nous avons effectué des tests sur deux molécules modèles : l'acide acétique et  $\text{Gly-Na}^+ \mathbf{3}$ . Nous avons observé une légère tendance à la stabilisation de l'énergie des formes acide trans avec la grande base, cependant cette stabilisation n'excède pas 10-20 kJ/mol. Des calculs sont en cours afin de comprendre cet effet de GB1.

### V.B.2. Influence de la méthode

La comparaison des calculs riBLYP et B3LYP montre un bon accord entre les valeurs d'énergies relatives obtenues avec ces deux méthodes et avec un même niveau de base pour les différents conformères. La déviation moyenne sur les calculs PB et GB est de 5.7 et 3.7 kJ/mol pour une valeur maximale de 11.5 kJ/mol et 13.0 kJ/mol, respectivement.

Les calculs HF/PB, même s'ils conduisent à un ordre énergétique légèrement différent de celui de B3LYP/GB1//B3LYP/PB1, permettent d'identifier les conformères les plus bas en énergie. La déviation entre les niveaux HF/PB1 et B3LYP/GB1//B3LYP/PB1 est de 4.7 kJ/mol en moyenne, pour une valeur maximale de 14 kJ/mol.

Les calculs MP2/GB//HF/PB présentent quant à eux des cas pathologiques qui semblent être associés au motif hélicoïdal et à la fonction acide trans, cependant la compréhension de ces cas demande des calculs complémentaires.



**Figure VI-13 : Structures des conformères complémentaires de  $\text{Gly}_6\text{-Na}^+$ .**

Nom	HF/PB1	MP2/PB1/ HF/PB1		MP2/ MB <sup>++*</sup> / HF/PB1		MP2/GB/ HF/PB1		B3LYP/ MB <sup>++*</sup> / PB1		B3LYP/ GB1/ B3LYP/ PB1		B3LYP/ MB <sup>++*</sup> / PB1		B3LYP/ GB1/ B3LYP/ PB1		riBLYP/ GB2/ riBLYP/ PB2	
		MP2/ MB <sup>++*</sup> / HF/PB1	MP2/ MB <sup>++*</sup> / HF/PB1	MP2/ MB <sup>++*</sup> / HF/PB1	MP2/GB/ HF/PB1	B3LYP/ PB1	B3LYP/ MB <sup>++*</sup> / PB1	B3LYP/ MB <sup>++*</sup> / PB1	B3LYP/ PB1	B3LYP/ MB <sup>++*</sup> / PB1	B3LYP/ GB1/ B3LYP/ PB1	riBLYP/ PB2	riBLYP/ PB2	riBLYP/ PB2	riBLYP/ PB2	riBLYP/ PB2	riBLYP/ PB2
1	c4ter	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	A1-24	-0.8	2.5	-2.8	-2.3	5.0	13.6	6.4	6.2	19.4	19.4	7.2	7.2	7.2	7.2	7.2	7.2
8	A3-1	12.7	12.0	4.2	2.9	10.4	24.7	12.9	12.9	33.8	33.8	18.8	18.8	18.8	18.8	18.8	18.8
7	A1-13	34.5	45.3	39.9	39.8	19.0	35.2	21.4	20.3	35.2	35.2	17.8	17.8	17.8	17.8	17.8	17.8
5	A1-27	15.2	21.2	23.6	22.6	27.6	17.5	13.1	12.3	23	23	15	15	15	15	15	15
6	G <sub>7</sub> -A2-8	10.9	33.5	31.5	28.6	34.2	28.3	13.3	11.3	37.7	37.7	18	18	18	18	18	18
4	A2-20	6.7	28.5	29.0	26.4	35.1	23.5	10.1	5.8	34.1	34.1	18.8	18.8	18.8	18.8	18.8	18.8
20	7s-2	18.9	33.7	36.6	33.9	41.1	36.3	32.5	27.5	47.8	47.8	32.1	32.1	32.1	32.1	32.1	32.1
14	7s	22.4	42.9	44.6	42.3	49.3	38.7	28.1	25.1	43.3	43.3	27.2	27.2	27.2	27.2	27.2	27.2
16	H $\alpha$ -3	25.8	48.2	30.9	33.4	-0.6	39.3	22.0	24.6	46.3	46.3	28.6	28.6	28.6	28.6	28.6	28.6
19	H 3 <sub>10</sub> -b	35.0				-81.9	49.4	25.0	28.3	32.6	32.6	21.6	21.6	21.6	21.6	21.6	21.6
28	H $\alpha$ -2	52.5	45.1	36.0	36.3	-96.7	52.3	41.4	43.6	56.6	56.6	44	44	44	44	44	44
36	H $\alpha$ -1	70.6	38.1	19.9	23.7	-120.4	70.4	56.6	61.4	70.2	70.2	62.4	62.4	62.4	62.4	62.4	62.4
37	H 3 <sub>10</sub>	96.3	74.2	47.1	49.5	-98.3	100.5	70.5	73.6	104.1	104.1	77.1	77.1	77.1	77.1	77.1	77.1

Tableau VI-7 : Energies relatives à 298K des structures les plus basses en énergie de  $\text{G}_6\text{-Na}^+$ , avec différentes méthodes : HF, MP2, B3LYP et ri-BLYP et différentes bases : PB1 : 6.31G\*, PB2 : SV(P), MB<sup>++\*</sup> : 6-31+G\*, MB<sup>++\*</sup> : 6-311+G(2d,2p) et GB2 : TZVPP.

## Bibliographie

- [1] DeGrado, W. F.; Summa, C. M.; Pavone, V.; Nastri, F.; Lombardi, A., De novo design and structural characterization of proteins and metalloproteins. *Annu. Rev. Biochem.* **1999**, 68, 779-819.
- [2] Avbelj, F.; Luo, P. Z.; Baldwin, R. L., Energetics of the interaction between water and the helical peptide group and its role in determining helix propensities. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, 97, (20), 10786-10791.
- [3] Pace, C. N.; Scholtz, J. M., A helix propensity scale based on experimental studies of peptides and proteins. *Biophys. J.* **1998**, 75, (1), 422-427.
- [4] Impronta, R.; Barone, V.; Kudin, K. N.; Scuseria, G. E., Structure and conformational behavior of biopolymers by density functional calculations employing periodic boundary conditions. I. The case of polyglycine, polyalanine, and poly-alpha-aminoisobutyric acid in vacuo. *J. Am. Chem. Soc.* **2001**, 123, (14), 3311-3322.
- [5] Topol, I. A.; Burt, S. K.; Deretey, E.; Tang, T. H.; Perczel, A.; Rashin, A.; Csizmadia, I. G., alpha- and 3<sub>10</sub>-helix interconversion: A quantum-chemical study on polyalanine systems in the gas phase and in aqueous solvent. *J. Am. Chem. Soc.* **2001**, 123, (25), 6054-6060.
- [6] Venkatraman, J.; Shankaramma, S. C.; BalaRama, P., Design of folded peptides. *Chem. Rev.* **2001**, 101, (10), 3131-3152.
- [7] Kemp, D. S.; Allen, T. J.; Oslick, S. L., The Energetics of Helix Formation by Short Templated Peptides in Aqueous-Solution .1. Characterization of the Reporting Helical Template Ac-He1(1). *J. Am. Chem. Soc.* **1995**, 117, (25), 6641-6657.
- [8] Kemp, D. S.; Boyd, J. G.; Muendel, C. C., The Helical S-Constant for Alanine in Water Derived from Template-Nucleated Helices. *Nature* **1991**, 352, (6334), 451-454.
- [9] Maison, W.; Kennedy, R. J.; Miller, J. S.; Kemp, D. S., C-terminal helix capping propensities in a polyalanine context for amino acids bearing nonpolar aliphatic side chains. *Tetrahedron Lett.* **2001**, 42, (30), 4975-4977.
- [10] Ruan, F. Q.; Chen, Y. Q.; Hopkins, P. B., Metal-Ion Enhanced Helicity in Synthetic Peptides Containing Unnatural, Metal-Ligating Residues. *J. Am. Chem. Soc.* **1990**, 112, (25), 9403-9404.
- [11] Kelso, M. J.; Hoang, H. N.; Appleton, T. G.; Fairlie, D. P., The first solution structure of a single alpha-helical turn. A pentapeptide alpha-helix stabilized by a metal clip. *J. Am. Chem. Soc.* **2000**, 122, (42), 10488-10489.

- [12] Ghadiri, M. R.; Choi, C., Secondary Structure Nucleation in Peptides - Transition-Metal Ion Stabilized Alpha-Helices. *J. Am. Chem. Soc.* **1990**, 112, (4), 1630-1632.
- [13] Kohtani, M.; Jarrold, M. F.; Wee, S.; O'Hair, R. A. J., Metal ion interactions with polyalanine peptides. *J. Phys. Chem. B* **2004**, 108, (19), 6093-6097.
- [14] Hudgins, R. R.; Jarrold, M. F., Helix formation in unsolvated alanine-based peptides: Helical monomers and helical dimers. *J. Am. Chem. Soc.* **1999**, 121, (14), 3494-3501.
- [15] Hudgins, R. R.; Mao, Y.; Ratner, M. A.; Jarrold, M. F., Conformations of Gly<sub>n</sub>H<sup>+</sup> and Ala<sub>n</sub>H<sup>+</sup> peptides in the gas phase. *Biophys. J.* **1999**, 76, (3), 1591-1597.
- [16] Kohtani, M.; Jarrold, M. F., The initial steps in the hydration of unsolvated peptides: Water molecule adsorption on alanine-based helices and globules. *J. Am. Chem. Soc.* **2002**, 124, (37), 11148-11158.
- [17] Harding, M. M., The architecture of metal coordination groups in proteins. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **2004**, 60, 849-859.
- [18] Glusker, J. P., Structural Aspects of Metal Liganding to Functional-Groups in Proteins. *Adv. Protein Chem.* **1991**, 42, 1-76.
- [19] <http://tanna.bch.ed.ac.uk/arch/>.
- [20] Patwari, G. N.; Lisy, J. M., Mimicking the solvation of aqueous Na<sup>+</sup> in the gas phase. *J. Chem. Phys.* **2003**, 118, (19), 8555-8558.
- [21] Kim, J. S.; Lee, S.; Cho, S. J.; Mhin, B. J.; Kim, K. S., Structures, Energetics, and Spectra of Aqua-Sodium(I) - Thermodynamic Effects and Nonadditive Interactions. *J. Chem. Phys.* **1995**, 102, (2), 839-849.
- [22] Glendening, E. D.; Feller, D., Cation Water Interactions - the M<sup>+</sup>(H<sub>2</sub>O)<sub>n</sub> Clusters for Alkali-Metals, M=Li, Na, K, Rb, and Cs. *J. Phys. Chem.* **1995**, 99, (10), 3060-3067.
- [23] Son, H. S.; Hong, B. H.; Lee, C. W.; Yun, S.; Kim, K. S., A new type of helix pattern in polyalanine peptide. *J. Am. Chem. Soc.* **2001**, 123, (3), 514-515.

## **Conclusions et Perspectives**

## Bilan

Nous avons étudié la structuration de molécules d'intérêt biologique : acides aminés et peptides par le cation sodium, en combinant approches théorique et expérimentale.

Grâce à la mise en place d'une nouvelle technique couplant un spectromètre de masse FTICR et un laser à électrons libres (LEL), il est possible de déterminer la structure des ions par spectroscopie IRMPD. Cette expérience n'ayant pas de source externe de formation d'ions biologiques de type MALDI ou ESI, nous avons développé la production d'ions par MALDI interne. Nous avons utilisé cette expérience afin de déterminer la structure non-zwitterionique ou zwitterionique des acides aminés Gly et Pro complexés à  $\text{Na}^+$ . Les spectres expérimentaux nous ont permis d'identifier la présence exclusive de la forme zwitterionique dans le cas de Pro- $\text{Na}^+$  et la présence de la forme non-zwitterionique dans le cas de Gly- $\text{Na}^+$ , conformément aux résultats de chimie quantique. Ainsi nous avons fourni la première démonstration directe de la présence d'un zwitterion d'acide aminé en phase gazeuse. Il s'agissait des premiers spectres infrarouge d'ions biologiques en phase gazeuse.

Ces expériences par spectroscopie IRMPD ont été étendues à l'étude d'oligomères de glycine  $\text{Gly}_n\text{-Na}^+$  et d'alanine  $\text{Ala}_n\text{-Na}^+$ , pour  $2 \leq n \leq 3$ . Nous avons dans un premier temps déterminé par des calculs ab initio les structures et les énergies des différents conformères des complexes et des peptides isolés. Les conformères de plus basse énergie sont ceux qui maximisent l'interaction électrostatique du métal avec les groupements carbonyles, avec ou sans l'amine terminale. La comparaison entre les résultats théoriques et expérimentaux montre que la technique IRMPD se révèle efficace pour identifier différentes classes de conformères, d'une part celles qui possèdent des nombres différents de carbonyles complexés au sodium, d'autre part celles dont le motif de complexation est tel que le peptide est enroulé autour du sodium de celles où le cation coiffe le peptide. Cependant cette technique reste limitée pour identifier les conformères dans une même classe ; ceci est notamment dû à la résolution des bandes IRMPD dont la largeur est d'environ  $50 \text{ cm}^{-1}$ . En parallèle, l'équipe de Chrys Wesdemiotis a déterminé les énergies de liaison des complexes  $\text{Gly}_n\text{-Na}^+$  et  $\text{Ala}_n\text{-Na}^+$ , pour  $2 \leq n \leq 4$ , par la méthode cinétique de Cooks. La comparaison avec nos valeurs calculées montre un excellent accord pour tous les peptides (déviations de 1 à 4  $\text{kJ/mol}$ ), excepté pour  $\text{Ala}_2$  et  $\text{Gly}_3$  (déviations de  $8 \text{ kJ/mol}$ ). Dans ce dernier cas, ce désaccord pourrait être dû à une mauvaise évaluation de la température effective. Pour ces petits peptides, l'accord entre expériences et calculs s'est révélé excellent, donnant accès à une description précise des structures et des énergies d'interaction.

Dans une dernière partie, nous nous sommes intéressés aux structures des complexes  $\text{Gly}_n\text{-Na}^+$ ,  $5 \leq n \leq 10$ , par des méthodes de chimie quantique. A partir de  $n=5$ , on peut distinguer deux grandes classes de conformères, d'une part celle où le peptide est globulaire, ne présentant pas de motif régulier et d'autre part celle où le peptide adopte une conformation plus "linéaire" telle que l'hélice  $\alpha$  ou  $3_{10}$ . Il existe vraisemblablement un nombre de résidus caractéristique de l'acide aminé qui compose l'homo-peptide, à partir duquel la structure en hélice  $\alpha$  devient le minimum global de la surface de potentiel. Nos résultats indiquent que ce nombre caractéristique est supérieur à 10, dans le cas de la glycine. Cependant il est intéressant de noter que pour  $\text{Gly}_{10}\text{-Na}^+$ , le conformère le plus bas en énergie présente une structure globulaire autour du sodium et un domaine de cinq résidus en hélice  $3_{10}$ . Comme l'espace conformationnel accessible augmente très rapidement avec le nombre de résidus, nous avons couplé des recherches conformationnelles de type Monte-Carlo basées sur des calculs de champ de forces de type AMBER à des optimisations par calculs de chimie quantique de toutes les structures précédemment engendrées. Pour les calculs de chimie quantique, nous avons utilisé l'approximation « résolution de l'identité » et la fonctionnelle de la densité BLYP. Cette approche a permis d'explorer en détail des surfaces de potentiel très complexes.

## Perspectives

La technique de spectroscopie IRMPD a fourni des informations structurales importantes, mais a aussi montré ses limites pour la différenciation de conformères ayant le même nombre de carbonyles complexés au sodium et adoptant un mode de complexation similaire, dans la gamme 1000 à 2000  $\text{cm}^{-1}$ . La caractérisation des complexes  $\text{Gly}_n\text{-Na}^+$  et  $\text{Ala}_n\text{-Na}^+$   $n=4-6$ , par spectroscopie IRMPD, prévue à très court terme, devrait confirmer ces résultats. Par ailleurs, les simulations de spectres infrarouge, dans la gamme 2500-4000  $\text{cm}^{-1}$  (vibration d'elongation X-H), indiquent que nous pourrions identifier les structures qui présentent une complexation de l'amine avec  $\text{Na}^+$  et les différents atomes impliqués dans une liaison hydrogène intramoléculaire. Cependant cette gamme de fréquence n'est pas accessible avec le Laser à Electrons Libres CLIO. Une solution serait d'utiliser un laser OPO et la technique de « tagging » des ions pour les dissocier par absorption à un seul photon.

Trois perspectives sont envisagées pour donner suite à l'étude des complexes  $\text{Gly}_n\text{-Na}^+$   $n \geq 5$ . Premièrement, il serait intéressant de mettre au point une méthode d'évaluation qualitative de la stabilité des différents conformères. Par ailleurs il serait intéressant de poursuivre l'étude des conformères de  $\text{Gly}_n\text{-Na}^+$  avec une taille croissante, afin de déterminer à partir de quel nombre de résidus l'hélice  $\alpha$  est le conformère le plus bas en

## Conclusions et Perspectives

---

énergie. Enfin, l'étude des conformères de  $\text{Ala}_n\text{-Na}^+$ ,  $5 \leq n \leq 10$  est en cours. En effet, la poly-alanine peut être considérée comme un système modèle de peptide ayant une forte propension à former des hélices ; son nombre caractéristique de résidu associé à la transition globule→hélice a été déterminé expérimentalement comme étant inférieur à 12. De plus, si on arrive à évaluer les différentes contributions énergétiques qui stabilisent les différents conformères, on peut espérer comprendre comment le rôle d'un simple méthyle peut transformer une poly-glycine qui possède la plus petite propension à former les hélices, en une poly-alanine qui possède la plus grande.

## RESUME

Il existe, aujourd'hui, de nombreuses et efficaces méthodes de caractérisation structurale tridimensionnelle de composés biologiques en phase condensée (RMN, diffraction des rayons X ou dichroïsme circulaire). Depuis une dizaine d'années, ce champ d'études s'est étendu à la phase gazeuse. Ce travail s'inscrit dans ce contexte et concerne le rôle structurant de  $\text{Na}^+$  sur les acides aminés Gly et Pro et sur des oligopeptides de Gly et Ala, en combinant approches théoriques et expérimentales de spectroscopie infrarouge d'ions gazeux par dissociation multiphotonique (IRMPD) couplée à la spectrométrie de masse.

Les spectres IRMPD expérimentaux des complexes sodiés d'acide aminé, nous ont permis d'identifier la présence exclusive de la forme zwitterionique dans le cas de  $\text{Pro}-\text{Na}^+$  et la présence de la forme non-zwitterionique dans le cas de  $\text{Gly}-\text{Na}^+$ , conformément aux résultats de chimie quantique. Ainsi nous avons fourni la première démonstration directe de la présence d'un zwitterion d'acide aminé en phase gazeuse. Il s'agissait des premiers spectres infrarouge d'ions biologiques en phase gazeuse.

L'étude théorique des complexes  $\text{Gly}_n-\text{Na}^+$  et  $\text{Ala}_n-\text{Na}^+$  a montré que, pour  $n \leq 5$ , les conformères de plus basse énergie maximisent

l'interaction électrostatique du métal avec les  $n$  groupements carbonyles, avec ou sans l'amine terminale. Ce comportement a été confirmé d'une part, par des expériences de spectroscopie IRMPD pour  $n=2,3$  et d'autre part, par la détermination des énergies de liaison de ces complexes par la méthode cinétique de Cooks ( $n=2-4$ ). Pour l'étude théorique de  $\text{Gly}_n-\text{Na}^+$ ,  $5 \leq n \leq 10$ , nous avons couplé des recherches conformationnelles Monte-Carlo basées sur des calculs de champ de forces AMBER, à des optimisations par calculs de type ri-BLYP utilisant l'approximation « résolution de l'identité ». Cette approche a permis d'explorer en détail des surfaces de potentiel très complexes. On peut distinguer deux classes limites de conformères, celle où le peptide est globulaire et celle où il adopte une conformation en hélice  $\alpha$  ou  $3_{10}$ . Nous avons montré que les structures les plus basses en énergie présentent le plus souvent une complexation tétradentate avec une forte auto-solvatation. Ces structures sont toutes globulaires pour  $n < 10$ . Dans le cas de  $\text{Gly}_{10}-\text{Na}^+$ , le conformère le plus bas en énergie a une structure globulaire autour du sodium et un domaine de cinq résidus en hélice  $3_{10}$ .

**Mots clés :** Spectroscopie infrarouge, ions, phase gazeuse,  $\text{Na}^+$ , acides aminés, peptides, zwitterion, globule, hélice, recherche conformationnelle, mécanique moléculaire, calculs quantiques.

## ABSTRACT

There exists, today, many and effective methods of three-dimensional structural characterization of biological compounds in the condensed phase (NMR, X-ray diffraction or circular dichroism). In the last ten years, this field has extended to the gas phase. This work falls in this context and considers the structuring role of  $\text{Na}^+$  on the amino acids Gly and Pro and on oligopeptides of Gly and Ala, by combining theoretical calculations and infra-red spectroscopy of gaseous ions by multiphoton dissociation (IRMPD), coupled to mass spectrometry.

Experimental IRMPD spectra of the sodiated amino acid complexes, enabled us to identify the exclusive presence of the zwitterionic form in the case of  $\text{Pro}-\text{Na}^+$  and the presence of the non-zwitterionic form in the case of  $\text{Gly}-\text{Na}^+$ , in accordance with quantum chemical results. Thus we provided the first direct demonstration of the presence of a zwitterion of an amino acid in gas phase. These were the first infra-red spectra of biological ions in the gas phase.

The theoretical study of the  $\text{Gly}_n-\text{Na}^+$  and  $\text{Ala}_n-\text{Na}^+$  complexes showed that for  $n \leq 5$ , the lower energy conformers maximize the

electrostatic interaction of the metal ion with the  $n$  carbonyl groups, with or without the terminal amine. This behavior was confirmed on the one hand by IRMPD spectroscopy for  $n=2,3$  and on the other hand by the determination of the binding energies of these complexes, by the Cooks kinetic method ( $n=2-4$ ). For the theoretical study of  $\text{Gly}_n-\text{Na}^+$ ,  $5 \leq n \leq 10$ , we coupled Monte-Carlo conformational searches using the AMBER force field, to geometry optimizations using the ri-BLYP density functional, using the approximation of the "resolution of the identity". This approach made it possible to explore in detail very complex potential surfaces. We can distinguish two limiting conformer classes, one in which the peptide is globular and another where it adopts an  $\alpha$  or  $3_{10}$  helical conformation. We showed that the lowest energy structures generally display a tetradentate complexation with a strong auto-solvation. All of these structures are globular for  $n < 10$ . In the case of  $\text{Gly}_{10}-\text{Na}^+$ , the most stable conformer has a globular structure around sodium plus five  $3_{10}$ -helical residues

**Keywords:** Infra-red spectroscopy, ions, gas phase,  $\text{Na}^+$ , amino acids, peptides, zwitterion, globule, helix, conformational research, molecular mechanics, quantum calculations.