



**HAL**  
open science

# Emergence de biopolymères de complexité contrôlée dans les scénarios d'origine de la vie

Lise Bedoin

► **To cite this version:**

Lise Bedoin. Emergence de biopolymères de complexité contrôlée dans les scénarios d'origine de la vie. Polymères. Sorbonne Université, 2021. Français. NNT : 2021SORUS229 . tel-03475214

**HAL Id: tel-03475214**

**<https://theses.hal.science/tel-03475214v1>**

Submitted on 10 Dec 2021

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

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

# Sorbonne Université

Ecole doctorale 397 - Physique et Chimie des Matériaux

*Laboratoire de Réactivité de Surface (LRS)*

*Institut Parisien de Chimie Moléculaire (IPCM) / Chimie Structurale Organique et  
Biologique (CSOB)*

## **Emergence de biopolymères de complexité contrôlée dans les scénarios d'origine de la Vie**

Par Lise Bedoin

Thèse de Doctorat de Chimie

Dirigée par Jean-François Lambert et Sandra Alves

Présentée et soutenue publiquement le 29 septembre 2021

Devant un jury composé de :

M. Aubriet Frédéric, Professeur, Rapporteur

M. Danger Grégoire, Maître de conférences, Rapporteur

Mme Jaber Maguy, Professeure, Examinatrice

M. Marie Arul, Ingénieur de recherches, Examineur

Mme Alves Sandra, Maîtresse de conférences, Co-directrice de thèse

M. Lambert Jean-François, Professeur, Directeur de thèse



# Table des matières

Table des matières .....	3
Liste des abréviations.....	7
Introduction générale .....	9
Chapitre I. Etude bibliographique .....	13
I. Implication des acides aminés et des surfaces minérales aux origines de la Vie .....	14
I.1. Chimie prébiotique des acides aminés .....	14
I.2. Oligomérisation des acides aminés (et des nucléotides) .....	18
I.3. Etudes expérimentales de la formation de peptides sur surfaces minérales.....	25
II. Outils analytiques pour l'étude de la condensation peptidique.....	31
II.1. Analyses <i>in situ</i> des produits de polymérisation sur surface de silice .....	32
II.2. Méthodes séparatives pour l'analyse de mélanges des peptides .....	33
II.3. Analyse par spectrométrie de masse à ultra haute résolution .....	38
II.4. Analyse de peptides par la spectrométrie de masse en tandem .....	43
II.5. Analyses par spectrométrie de masse et origines de la Vie .....	45
III. Description des systèmes étudiés.....	47
III.1. Choix des acides aminés .....	47
III.2. Spécificités de l'acide glutamique .....	49
III.2.2. La question de la régiosélectivité de la condensation : $\alpha$ - et $\gamma$ -peptides .....	52
Chapitre II. Matériels et méthodes .....	55
I. Préparation des échantillons .....	56
I.1. Dépôt sur surface minérale .....	56
I.2. Description de l'ensemble des échantillons testés .....	60
I.3. Protocole d'activation thermique .....	62
I.4. Etape d'extraction en solution .....	62
II. Techniques de caractérisation <i>in-situ</i> des systèmes AAs/SiO <sub>2</sub> .....	63
II.1. Analyse thermogravimétrique.....	63
II.2. Diffraction des rayons X .....	63



II.3. Spectroscopie infrarouge .....	64
III. Analyses <i>ex situ</i> de la solution de désorption.....	65
III.1. HPLC-UV .....	65
III.2. Spectromètre de masse à haute résolution FTICR .....	65
III.3. Spectromètre de masse TIMS-TOF .....	68
IV. Considérations théoriques sur la polymérisation des acides aminés .....	74
Chapitre III. Sélectivité de polymérisation dans un système binaire (L+E)/SiO <sub>2</sub> .....	77
I. Etude préliminaire du système Leu + Glu .....	78
II. Sélectivité de composition des produits de polymérisation.....	81
II.1. Etude de la stœchiométrie des oligopeptides formés.....	83
II.2. Origine de la sélectivité de stœchiométrie.....	86
II.3. L'entropie de Shannon pour évaluer la sélectivité de polymérisation .....	89
III. Etude de la régiosélectivité de la polymérisation de (L+E)/SiO <sub>2</sub> .....	93
IV. Etude de l'énantiosélectivité de la polymérisation de (L+E)/SiO <sub>2</sub> .....	96
Chapitre IV. Etude du mécanisme de polymérisation : changement des conditions opératoires .....	101
I. Etude cinétique de la polymérisation .....	102
I.1. Activation thermique de (L+E)/SiO <sub>2</sub> à différentes températures et différents temps de réaction .....	102
I.2. Etude de la polymérisation à partir de dipeptides .....	108
II. Variation des quantités d'acides aminés déposés sur la silice .....	115
II.1. Effet de la densité d'acides aminés déposés : importance de la diffusion de surface .....	115
II.2. Variation du rapport Leu :Glu déposé sur la silice.....	118
III. Condensation sur d'autres surfaces minérales.....	122
III.1. Réactivité de Leu + Glu sur montmorillonite.....	123
III.2. Réactivité de Leu + Glu sur alumine .....	124
IV. Evaluation d'autres scénarios de polymérisation.....	127
IV.1. Cycles d'hydratation /déshydratation.....	127
IV.2. Dépôt de Leu et Glu sur des particules de silice différentes .....	129
V. Discussion générale .....	131
V.1. Rappels des données bibliographiques .....	131
V.2. Conclusions de la présente étude .....	132

---

Chapitre V. Vers une généralisation du scénario de polymérisation sur silice .....	137
I. Choix de l'acide aminé trifonctionnel : étude de différents systèmes binaires (L+X)/SiO <sub>2</sub> .....	138
I.1. Présentation des systèmes binaires étudiés .....	138
I.2. Comparaison de différents acides aminés trifonctionnels .....	141
I.3. Comparaison de la réactivité des acides aminés trifonctionnels .....	149
I.4. Conclusion sur les systèmes binaires .....	152
II. Etude de systèmes plus complexes : polymérisation des mélanges GAVD et GAVE .....	154
II.1. Réactivité similaire des deux systèmes : le rôle essentiel du résidu trifonctionnel .....	155
II.2. Différence de réactivité entre les systèmes GAVD et GAVE .....	165
III. Etude de la polymérisation de nucléotides .....	169
Chapitre VI. Etude structurale des oligopeptides formés par condensation sur surface de silice .....	175
I. Présentation des techniques utilisées .....	176
I.1. Détermination des séquences peptidiques par MS/MS .....	176
I.2. Couplage de la spectrométrie de masse avec des techniques de séparation .....	178
II. Elucidation structurale des peptides formés par condensation peptidique dans le système Leu + Glu ....	180
II.1. Etude structurale des oligopeptides intacts .....	180
II.2. Sélectivité de séquence des formes déshydratées .....	186
III. Etude de la sélectivité de séquence des produits de condensation .....	193
III.1. Séquençage par MS/MS et position de l'acide aminé trifonctionnel .....	193
III.2. La question de la régiosélectivité .....	197
IV. Etude structurale des formes déshydratées .....	205
IV.1. Dipeptides déshydratés .....	205
IV.2. Généralisation aux oligopeptides déshydratés .....	212
Conclusion générale .....	217
Références .....	221
Annexes .....	237



# Liste des abréviations

A	
AA	Acide aminé
Asp (ou D)	Acide aspartique
ATG	Analyse Thermogravimétrique
B	
BIRD	Blackbody Infrared Radiative Dissociation Dissociation radiative infrarouge de corps noir
C	
CCS	Collision Cross Section Section efficace de collision
CDI	Carbonyldiimidazole
CID	Collision Induced Dissociation Dissociation induite par collision
D	
DDA	Data Dependent Acquisition Acquisition de spectres MS/MS automatisée
DTIMS	Drift Time Ion Mobility Spectrometry Mobilité ionique dans un tube de dérive
DKP	Diketopiperazine Dicétopipérazine
DRX	Diffraction des Rayons X
E	
ECD	Electron Captured Dissociation Dissociation par capture d'électron
EDD	Electron Detachment Dissociation Dissociation par détachement d'électron
EIC	Extract Ion Chromatogram Chromatogramme d'un ion reconstitué
EID	Electron Induced Dissociation Dissociation induite par un electron
EIM	Extract Ion Mobilogram Mobilogramme d'un ion reconstitué
ESI	Electrospray Ionization Ionisation par électronébulisation
ETD	Electron Transfer Dissociation Dissociation par transfert d'électron
F	
FAIMS	Field Asymmetric Ion Mobility Spectrometry Mobilité ionique à haut champ asymétrique
FIA	Flow Injection Analysis Analyse en flux continu
FTICR	Fourier Transform Ion Cyclotron Resonance Spectromètre à résonance cyclotronique à transformée de Fourier
G	
Gly (ou G)	Glycine
Gln (ou Q)	Glutamine
Glu (ou E)	Acide glutamique
H	
HPLC	High Performance Liquid Chromatography Chromatographie liquide à haute performance

HRMS	High Resolution Mass Spectrometry Spectrométrie de masse à haute résolution
I	
IM-MS	Ion Mobility-Mass spectrometry Mobilité ionique couplée à la spectrométrie de masse
IRMPD	Infrared Multiphoton Dissociation Dissociation par absorption multiple de photons
K	
KM	Kendrick Mass Masse de Kendrick
KMD	Kendrick Mass Defect Défaut de masse de Kendrick
L	
Leu (ou L)	Leucine
LC/MS	Liquid Chromatography – Mass Spectrometry Chromatographie liquide couplée à la spectrométrie de masse
M	
MS	Mass Spectrometry Spectrométrie de masse
MS/MS	Tandem Mass Spectrometry Spectrométrie de masse en tandem
R	
RP-HPLC	Reverse Phase High Performance Liquid Chromatography Chromatographie liquide à haute performance en phase inverse
S	
SORI-CID	Sustained Off-Resonance Irradiation Collision-Induced Dissociation
T	
TIMS-TOF	Trapped Ion Mobility Spectrometry-Time of Flight Mobilité ionique par piégeage d'ions couplée à un spectromètre à temps de vol
TWIMS	Traveling Wave Ion Mobility Spectrometry Mobilité ionique à champ oscillant
U	
UVPD	UltraViolet PhotoDissociation
V	
Val (ou V)	Valine

# Introduction générale

La matière vivante se caractérise par la présence de biopolymères non aléatoires dont la fonction biologique dépend de la séquence de monomères. Pour comprendre les origines de la vie, il est essentiel d'expliquer l'émergence de polymères non aléatoires dans un monde où les cellules vivantes n'existaient pas.

Pour ce faire, il est nécessaire de réussir à synthétiser des oligomères à partir des briques constitutives du vivant disponibles à l'époque Archéenne : des peptides à partir d'acides aminés, ou des acides nucléiques à partir de nucléotides. Parmi tous les scénarios prébiotiques plausibles, nous avons envisagé une polymérisation sur des surfaces minérales. L'objectif initial de ce travail de thèse était de déterminer si une sélectivité de polymérisation peut être mise en évidence lors d'une condensation d'acides aminés sur une surface minérale, dans des conditions cohérentes avec un scénario prébiotique. Des acides aminés déposés sur un support minéral sont activés thermiquement dans des conditions contrôlées. De nombreux travaux de recherche ont prouvé l'efficacité de ce scénario, qui reste pourtant mal compris au niveau fondamental. La principale technique analytique choisie pour étudier les mélanges d'oligopeptides ainsi produits est la spectrométrie de masse à haute résolution, qui a pour principaux avantages d'être spécifique et sensible pour l'analyse de ce type de mélange complexe.

Ce manuscrit s'articule en six chapitres.

Dans le premier, une description générale du contexte prébiotique sera exposée. Une discussion des différentes études antérieures traitant de la sélectivité de polymérisation justifiera du choix du système étudié et des outils analytiques utilisés.

Dans le second chapitre, les matériaux utilisés et méthodes d'analyse employées dans cette étude seront décrits, avec un accent porté sur la description des appareils.

Le troisième chapitre sera le premier chapitre de résultats et portera sur l'étude de la polymérisation de l'acide glutamique et de la leucine sur une surface de silice, système qui servira de référence pour la suite.

Le quatrième chapitre portera sur une étude mécanistique de la réaction de polymérisation du système exposé au Chapitre III, avec notamment une étude cinétique visant à accéder aux différentes étapes de la condensation.

Dans le cinquième chapitre, d'autres couples d'acides aminés seront explorés, ainsi que des systèmes à quatre acides aminés dans le but de mettre en évidence les principes généraux à l'œuvre dans les polymérisations. La complexité des spectres de masse obtenus nous conduira à utiliser des techniques d'analyse de données basées sur la masse précise telles que les analyses de Kendrick et de Van Krevelen pour comparer les produits de polymérisation formés.

Le sixième chapitre apportera finalement des données supplémentaires sur la structure des oligopeptides formés à l'aide de techniques de caractérisation avancée comme la spectrométrie de masse en tandem, et la spectrométrie de masse à haute résolution couplée avec des techniques de séparation telles que la chromatographie liquide et la mobilité ionique.





# Chapitre I. Etude bibliographique

# I. Implication des acides aminés et des surfaces minérales aux origines de la Vie

## I.1. Chimie prébiotique des acides aminés

### I.1.1. Apparition de la Vie sur Terre – Origine des petites biomolécules

L'étude des origines de la Vie sur Terre s'intéresse aux conditions et processus qui ont permis l'émergence du vivant sur notre planète. Le fonctionnement actuel des organismes vivants implique la synthèse de nombreuses molécules biologiques plus ou moins complexes dont la machinerie cellulaire élaborée n'existait pas aux origines du vivant. Des scénarios de synthèse alternatifs, dits « prébiotiques » ont donc dû se produire pour former les premières molécules organiques. On appelle ainsi « molécule prébiotique » une molécule organique formée sans l'intervention d'êtres vivants. L'abiogénèse caractérise le passage de l'inerte au vivant. Les peptides dits abiogéniques ont donc été produits dans un environnement naturel plausible mais pas au sein d'organismes vivants.

La question de l'Origine de la Vie suscite de nombreuses réflexions. L'univers est vieux de 13,7 milliards d'années et la Terre n'a « que » 4,5 milliards d'années. Les premières traces d'apparition de la vie, des cyanobactéries, sont datées d'il y a environ 3,5 milliards d'années. Si l'on peut identifier l'apparition de formes de vie anciennes, le mécanisme de cette apparition n'est pas encore clair. Il est très probable que la préexistence d'un monde minéral complexe soit une condition nécessaire pour l'émergence de biomolécules simples.

A la fin du XIX<sup>ème</sup> siècle, les premières théories supposent que la vie aurait été apportée par des météorites<sup>1</sup>. Sans aller jusque-là, elles auraient pu apporter certains constituants chimiques nécessaires à la vie. Effectivement, l'analyse de météorites, et notamment celle de Murchison (tombée en Australie en 1969 et récoltée très peu de temps après sa chute, minimisant les possibilités de contamination), a montré qu'elles contenaient une grande variété d'acides aminés<sup>2</sup> et de bases nucléiques<sup>3</sup>). Le milieu interstellaire constitue donc un lieu possible pour la synthèse prébiotique de composés polyatomiques, avec pour source d'énergie le rayonnement cosmique, les irradiations UV et les décharges électriques<sup>4-6</sup>.

Cette synthèse primordiale aurait pu aussi se produire dans les océans ou au creux des sources hydrothermales. De nombreuses expériences ont été menées en laboratoire pour tenter de (re-)créer les entités constitutives du vivant à partir d'espèces chimiques compatibles avec les conditions d'une Terre primitive. La plus connue est l'expérience de Miller, qui a montré la formation d'acides aminés après avoir recréé en laboratoire les conditions régnant sur Terre à l'époque de l'apparition de la vie<sup>5</sup> – du moins telles qu'on les concevait dans les années 1950. On pensait en effet que l'atmosphère primitive était essentiellement composée de CH<sub>4</sub>, NH<sub>3</sub> et d'eau sous forme de vapeur. L'atmosphère primitive aurait ainsi été anoxique et nettement réductrice. Miller a donc soumis un mélange de H<sub>2</sub>, CH<sub>4</sub>, et NH<sub>3</sub> à des décharges électriques mimant les éclairs et à un flux d'eau chauffée mimant la pluie et les océans. Il a alors obtenu un mélange de produits organiques qui a ensuite été baptisé « soupe primordiale ». Les gaz ont notamment formé du HCN et du formaldéhyde (H<sub>2</sub>CO). Ces molécules à liaisons multiples ont alors réagi avec l'eau pour former onze acides aminés (AAs) différents sur les vingt naturels (sans énantiosélectivité, c.a.d. avec autant de molécules énantiomères L que D). En 2011, Parker reprend les échantillons de Miller avec des techniques analytiques plus sophistiquées et réussit à caractériser et quantifier vingt-trois acides aminés, dont certains ne sont pas utilisés par les organismes vivants sur Terre<sup>6</sup>.

Les acides aminés, dont la structure est montrée Figure I-1, sont parmi les molécules emblématiques du vivant. Ils constituent les unités monomériques des peptides/protéines. Également molécules emblématiques du vivant, les nucléotides sont les monomères des acides nucléiques (voir Figure I-1). Les nucléotides sont les briques élémentaires de l'ARN (acide ribonucléique) et de l'ADN (acide désoxyribonucléique). La différence entre l'ARN et l'ADN se situe au niveau de la substitution du sucre : l'ARN contient des résidus de ribose (OH) là où l'ADN contient du désoxyribose (H). L'atome d'oxygène supplémentaire sur chacun des nucléotides de l'ARN entraîne une plus grande flexibilité de l'ARN, permettant la formation d'hélices et le rendant chimiquement moins stable.

Les acides  $\alpha$ -hydroxylés (qui ont des structures proches de celles des acides aminés, voir Figure I-1). ont été produits lors de réactions prébiotiques<sup>5,7-10</sup>. Par ailleurs, en 1960, Oró<sup>11</sup> a mis en évidence l'obtention de nucléobases à partir de cyanure d'ammonium. Sutherland<sup>12,13</sup> a montré la formation de nucléotides, d'acides aminés et de glycérol en

soumettant un mélange d'espèces chimiques compatibles avec la Terre primitive à des rayons UV.

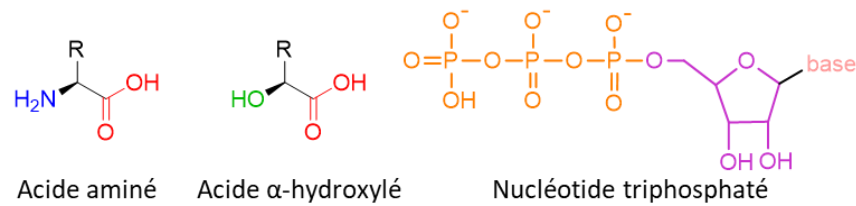


Figure I-1. Structures d'un acide aminé, d'un acide  $\alpha$ -hydroxylé (configuration L) et d'un nucléotide.

Il existe donc un certain nombre d'hypothèses plausibles pour expliquer naturellement la formation des « briques de base » du vivant. Mais leur formation n'est qu'une première étape vers la complexité des structures vivantes. Des plus grandes entités ont dû ensuite être formées à partir de ces molécules élémentaires. En particulier, deux des constituants chimiques les plus typiques des organismes vivants sont des polymères de ces briques simples : les protéines sont de longs peptides, c'est-à-dire des polymères d'acides aminés, et les acides nucléiques (ARN, ADN) des polymères de nucléotides.

### I.1.2. « Monde ARN » ou « monde des peptides » ?

Pour expliquer l'organisation du vivant au-delà des molécules les plus simples, deux modèles principaux ont été proposés : le « monde ARN » et le « monde des peptides »<sup>14</sup>. Le système à l'origine de la Vie a dû être capable de maîtriser le métabolisme et de conserver l'information génétique ; or dans les cellules modernes ces fonctions sont associées à deux types de macromolécules, respectivement les protéines et l'ADN. Chacune dépendant de l'autre pour son fonctionnement, on se retrouve, pour expliquer leur origine, devant un problème typique du type « l'œuf ou la poule ? ».

Dans les années 1980, Walter Gilbert (titulaire d'un prix Nobel pour sa contribution au séquençage de l'ADN) évoque l'hypothèse d'un « monde ARN » dans lequel l'acide ribonucléique serait le précurseur unique des macromolécules biologiques<sup>15</sup>. Il envisage un monde ne contenant que des molécules d'ARN qui catalysent la synthèse les unes des autres et se répliquent sans l'intervention d'autres molécules. Les molécules d'ARN auraient

ensuite commencé à synthétiser des protéines en liant des acides aminés activés. Enfin, l'ADN serait apparu comme le dernier support de l'information, obtenu à une étape ultérieure par transcription inverse à partir des molécules d'ARN. Le principal problème de ce modèle est qu'il n'explique pas comment les molécules initiales d'ARN ont été formées. Cette question est d'autant plus complexe que, si on connaît des scénarios prébiotiques qui aboutissent aux bases nucléiques ou aux sucres (synthèse du formose)<sup>16</sup>, aucun ne permet encore d'unir une base nucléique, un sucre (le ribose) et un groupe phosphate en une seule molécule.

Les protéines, autres molécules de base du vivant, sont quant à elles connues pour être de bons catalyseurs puisque la plupart des enzymes sont des protéines. Certains dipeptides peuvent déjà posséder des propriétés catalytiques intéressantes<sup>17</sup>. Les protéines résultent d'une condensation entre acides aminés, qui, nous l'avons vu, étaient disponibles sur la Terre prébiotique. En revanche, elles ne peuvent généralement pas se répliquer et conserver l'information génétique.

Une collaboration entre les oligopeptides et les protéines est sans doute le scénario le plus probable<sup>14 18</sup>. Brack suppose ainsi que le monde ARN apparaît comme un épisode biologique mais pas au tout début de la vie sur Terre<sup>19</sup>.

### I.1.3. Chiralité de la Vie

La chiralité est un paramètre important pour les systèmes vivants. De nombreuses biomolécules possèdent des énantiomères D ou L. La vie a une sélectivité presque parfaite pour certaines espèces chirales, et notamment les sucres D et les acides aminés L<sup>20</sup>. La genèse de l'homochiralité est le processus qui a mené à la préférence presque exclusive d'une forme énantiomérique dans le vivant. On peut considérer que ce processus a eu lieu en trois étapes : la perte de symétrie, l'enrichissement chiral et la transmission chirale<sup>21</sup>.

En ce qui concerne la première étape, de nombreux agents physiques pouvant induire l'apparition de la chiralité ont été proposés<sup>22</sup>. De Marcellus et al<sup>23</sup> ont montré que la lumière UV circulairement polarisée pouvait induire la formation d'une molécule (l'acide aminé alanine) avec un excès énantiomérique à partir d'un mélange ne contenant pas de substances chirales. Il faut noter que l'intervention des surfaces minérales, au centre de cette thèse, a pu jouer un rôle également dans l'apparition de la chiralité : certaines

surfaces, comme la calcite permettent la séparation chirale d'acides aminés. Ainsi, Hazen et al<sup>24</sup> ont montré que certaines faces de la calcite avait une préférence pour l'adsorption de la forme L de l'acide aspartique L plutôt que de la forme D. L'effet inverse est observé pour d'autres faces. Ce phénomène suggère une concentration locale d'énantiomères L ou D entraînant une polymérisation homochirale<sup>25</sup>. Certaines études montrent un scénario dans lequel l'amplification de l'homochiralité aurait pu évoluer lors de l'étape de polymérisation des petites biomolécules, comme proposé pour les acides nucléiques par Joyce et al<sup>26</sup> et Goldanskii et al<sup>27</sup>.

L'histoire du vivant ne s'arrête pas là. La recherche sur l'origine du métabolisme, c'est-à-dire l'organisation « en réseau » des réactions chimiques et la maîtrise des flux énergétiques est par exemple un domaine très exploré mais le présent travail ne porte pas sur ces aspects.

## 1.2. Oligomérisation des acides aminés (et des nucléotides)

Le but de ce projet doctoral est d'étudier la sélectivité de la réaction de polymérisation d'acides aminés en peptides, selon une réaction de condensation. Cette condensation de deux monomères consiste en une réaction entre la fonction amine de l'un et la fonction acide carboxylique de l'autre, accompagnée de l'élimination d'une molécule d'eau comme montré sur la Figure I-2. Or, la condensation des acides aminés en protéines ne se produit pas spontanément. D'après les données thermodynamiques, elle est défavorable en solution aqueuse. En d'autres termes, la variation d'enthalpie libre standard ( $\Delta_r G^\circ$ ) de cette réaction est positive : cela revient à dire que la liaison peptidique est thermodynamiquement instable<sup>28</sup>, ou que les protéines en solution aqueuse sont des « molécules activées », des « réservoirs d'énergie libre », comme d'ailleurs la plupart des molécules du vivant. Les données disponibles<sup>29</sup> indiquent que chaque étape de condensation peptidique a une constante d'équilibre à température ambiante de  $10^{-3}$  à  $10^{-2}$ . Les concentrations à l'équilibre d'oligopeptides en solution aqueuse, même si la concentration totale d'acides aminés est élevée, deviennent donc très faibles dès que la longueur du peptide dépasse quelques résidus.



Figure I-2. Schéma de condensation de deux acides aminés en peptide

A l'image de la formation de peptides, la condensation de monomères d'ARN n'est pas favorable en solution aqueuse<sup>30</sup>.

Par conséquent, que l'on veuille synthétiser des peptides ou des acides nucléiques, il faut proposer un scénario où la réaction de polymérisation devienne thermodynamiquement favorable, c'est-à-dire s'accompagne d'une diminution de l'énergie libre ( $\Delta_r G^\circ < 0$ ). Il est important de souligner que le problème est de nature thermodynamique et donc qu'une modification de la cinétique de la réaction, par exemple l'ajout d'un catalyseur, ne sera pas suffisante pour réaliser la réaction souhaitée.

### I.2.1. Scénarios prébiotiques pour la production des peptides

Plusieurs scénarios s'affranchissant de cette limitation ont été proposés et testés.

#### I.2.1.1. Polymérisation en solution

Une première voie de recherche a été de modifier les conditions thermodynamiques. Ainsi, des scénarios hydrothermaux ont été proposés pour la condensation des acides aminés. La condensation peptidique est en effet plus favorable dans des conditions de pression et de température élevées. Une étude de Cleaves et al.<sup>31</sup> conclut toutefois qu'il est peu probable que ces scénarios puissent avoir agi comme source d'oligopeptides. L'application de conditions hydrothermales, notamment en combinaison avec la présence de surfaces minérales, constitue malgré tout une direction de recherche intéressante pour l'allongement<sup>32</sup> de chaînes oligopeptidiques déjà formées.

Pour réaliser la synthèse peptidique en solution aqueuse, on peut la coupler avec une seconde réaction thermodynamiquement favorable, de manière à rendre négatif le changement d'enthalpie libre global des deux réactions. C'est en somme ce que réalise la biochimie actuelle, où un certain nombre de molécules « riches en énergie » sont sacrifiées



de façon à réaliser les réactions de l'anabolisme, et notamment la synthèse des protéines. Les biochimistes maîtrisent depuis longtemps un certain nombre de procédures de synthèse fonctionnant *in vitro* et impliquant ce qu'on appelle parfois des activateurs sacrificiels.

Ainsi, le N,N-carbonyldiimidazole (CDI), est un agent activant couramment utilisé en synthèse peptidique pour convertir les acides carboxyliques en amides ou en esters. Son action implique le remplacement du OH en C-terminal de la chaîne peptidique par un meilleur groupe partant, assurant ainsi le couplage entre l'agent activant et la molécule d'acide aminé. Le  $\Delta_r G^\circ$  négatif est fourni par l'hydrolyse des liaisons entre le groupement carbonyl central et les imidazoles, qui est favorable en solution aqueuse. Le CDI a été employé dans plusieurs études de chimie prébiotique<sup>33-36</sup>. Cependant il n'existe aucune étude prouvant le caractère prébiotique de ce composé<sup>37</sup>. Nous l'avons utilisé dans le cours de cette thèse (Chapitre III.II.2), mais uniquement comme point de comparaison pour évaluer la sélectivité de polymérisation dans nos conditions. Le CDI représente une alternative aux carbodiimides, des groupes fonctionnels de formule  $R^1-N=C=N-R^2$  couramment utilisés en synthèse peptidique pour convertir les acides carboxyliques en amides ou en esters. Ils fonctionnent quant à eux selon un mécanisme légèrement différent, dans lequel l'enthalpie libre nécessaire est fournie par l'hydrolyse des liaisons imides.

Une possibilité beaucoup plus envisageable est la formation d'intermédiaires réactionnels particuliers, les N-carboxyanhydrides<sup>38</sup> (NCA), qui peuvent être obtenus entre autres par réaction des acides aminés avec les cyanates, réservoirs d'énergie libre bien connus. Ces NCA peuvent ensuite réagir pour former des oligopeptides. Ce scénario est intéressant prébiotiquement mais nous ne l'avons pas étudié dans le cadre de notre travail.

Il existe des scénarios prébiotiques avec des agents activants inorganiques, comme le trimétaphosphate (« P3m »). Les polyphosphates ont pu être présents à l'ère primitive, notamment formés dans les volcans. Il s'agit de molécules « riches en énergie » en solution aqueuse, métastables, et qui peuvent se coupler avec plusieurs types de molécules par phosphorylation. La polymérisation des acides aminés en solution aqueuse<sup>39</sup> a ainsi pu être réalisée. Une étude réalisée au LRS a par exemple montré que le P3m permet de condenser deux molécules d'acide aspartique en solution (résultats non publiés).

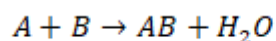
### I.2.1.2. Utilisation de surfaces minérales

Un autre scénario fait appel à l'implication de surfaces minérales afin d'abaisser l'activité de l'eau. Le cristallographe anglais John Desmond Bernal a émis le premier l'hypothèse que les étapes de polymérisation des biomolécules impliquaient des réactions sur des surfaces minérales comme les argiles<sup>33</sup>. Il se basait sur l'argument que l'adsorption locale augmente la concentration effective, favorisant l'association de molécules. Les acides aminés pourraient se fixer entre les feuillets d'argile et leur rapprochement favoriserait la réaction de condensation<sup>34,35</sup>. Cette possibilité est restée longtemps sans vérification, jusqu'aux premières études expérimentales (cf. Chapitre I.1.3).

Il importe de souligner ici que le rôle des surfaces minérales ne peut pas se réduire à celui de catalyseur. En effet, le problème de la formation des peptides est tout d'abord thermodynamique<sup>28</sup>, et par définition un catalyseur ne modifie pas la thermodynamique d'une réaction. L'explication de Bernal était qu'une augmentation par adsorption de l'activité des monomères doit automatiquement déplacer vers la droite l'équilibre de condensation de ces monomères, conformément au principe de Le Châtelier. Cette hypothèse est cohérente, mais le calcul des concentrations des monomères nécessaires pour obtenir une polymérisation significative aboutit à des valeurs impossiblement élevées. Le rôle de la surface ne peut donc pas être uniquement dû à l'augmentation des concentrations locales dans son voisinage.

On pourrait alors penser que le rôle des surfaces est de modifier la constante d'équilibre de la réaction de condensation, qui serait plus favorable en « phase adsorbée » qu'en phase aqueuse. Cela serait le cas si les produits de la réaction (en l'occurrence les peptides) interagissaient plus fortement avec la surface que les réactifs (les acides aminés). Cette hypothèse n'a été testée qu'assez tard<sup>40</sup> – avec des résultats négatifs, l'équilibre de polymérisation n'étant pas sensiblement déplacé en phase adsorbée. De plus, elle implique un paradoxe, souligné par de Duve et Miller. En effet, les biopolymères devraient alors être de plus en plus fortement immobilisés sur la surface au fur et à mesure qu'ils s'allongent, ne les rendant plus disponibles pour une évolution ultérieure.

En fin de compte, la solution repose probablement sur le principe de Le Châtelier, mais appliqué cette fois à l'activité de l'eau. La réaction de condensation s'écrit :



avec une constante d'équilibre  $K = \frac{\alpha_{AB} \cdot \alpha_{H_2O}}{\alpha_A \alpha_B}$ . Il en résulte que, si l'activité de l'eau  $\alpha_{H_2O}$

diminue, le rapport  $\frac{\alpha_{AB}}{\alpha_A \alpha_B}$  doit augmenter pour que le quotient réactionnel reste constant. La

condensation peut donc se réaliser en séchant le système, et le rôle thermodynamique de la surface se réduit à celui d'une « plate-forme de séchage ». Dans ce cas, le séchage d'une solution d'acides aminés favorise la polymérisation, même en l'absence de surface, mais cela nécessite d'activer le système à des températures assez élevées, par exemple 240 °C pour la glycine massique au lieu de 160 °C pour Gly/SiO<sub>2</sub><sup>41</sup>. Par conséquent, le rôle de la surface est double : thermodynamique en permettant de déshydrater extensivement les acides aminés, et cinétique, en catalysant la formation de liaisons peptidiques<sup>42</sup>.

### I.2.1.3. Un scénario mixte : utilisation de cycles hydratation/déshydratation

Plusieurs travaux de recherches ont étudié un scénario impliquant des cycles hydratation/déshydratation, l'objectif étant d'imiter les changements climatiques tels qu'ils ont pu se produire dans l'environnement géochimique réel. Lahav, un des premiers dans cette approche, a ainsi montré que l'augmentation du nombre de cycles appliqués entraînait une augmentation de la longueur des oligopeptides formés pour une polymérisation de glycine déposée sur argile. Les détails de ces cycles ainsi que d'autres conditions expérimentales testées sont reportés dans le Tableau I-1. Ces méthodes ont été utilisées pour la polymérisation d'acides aminés en peptides mais aussi d'hydroxy acides en polyesters, ou encore la production de depsipectides (comprenant une alternance de liaisons amide et ester) en associant des hydroxy acides à des acides aminés.

*Tableau I-1 : Exemples de conditions d'oligomérisation impliquant des cycles hydratation/déshydratation*

Articles	Système	Etape sèche	Etape humide
Lahav et al. <sup>43</sup>	Glycine sur argile	Déshydratation à 60 °C pendant 1 à 2 jours. Chauffage à 94 °C pendant 2 à 3 jours.	Réhydratation à l'eau
White et al. <sup>44</sup>	Acides aminés sur argile	Evaporation dans four sous vide à 40 °C. Chauffage à 80-94 °C	Réhydratation à l'eau
Bujdak et al. <sup>45</sup>	Glycine sur montmorillonite	80 °C pendant 24h	Réhydratation à l'eau
Forsythe et al. <sup>7</sup>	Acide lactique	85 °C pendant 18h	Réhydratation à l'eau pendant 30 min à température ambiante et chauffage à 65 °C pendant 5h30.
Ervin et al. <sup>46</sup>	Acide aminé + acide $\alpha$ -hydroxylé	85 °C pendant 24h	Réhydratation à l'eau

Les procédures d'hydratation et de déshydratation sont toutes différentes et loin d'être standardisées et les résultats sont donc difficiles à interpréter. Elles suggèrent toutefois une réflexion intéressante sur les conditions d'apparition de la vie, qui est un état de la matière hors d'équilibre et riche en molécules de haute énergie libre. La source de cette énergie libre pourrait résider dans les déséquilibres physiques, manifestés par l'alternance de conditions d'humidité faible et élevée, et l'énergie du vivant trouverait son origine en partie dans celle du système météorologique de la planète.

### I.2.2. Formation de nucléotides et d'oligonucléotides

Puisque la condensation de monomères d'ARN n'est pas favorable en solution aqueuse, certaines études<sup>47-49</sup> utilisent des nucléotides activés. Zagorevskii et al.<sup>49</sup> ont, par exemple, montré l'oligomérisation d'adénosine 5'-phosphorimidazolide (ImpA) et d'uridine 5'-phosphorimidazolide (ImpU) sur montmorillonite (utilisée comme catalyseur). Ces travaux sont souvent cités comme un exemple de polymérisation « on the rocks », mais il faut souligner que dans le cas présent, les surfaces minérales ne jouent pas de rôle thermodynamique, puisque la force motrice thermodynamique est fournie par la dégradation d'une molécule activée.

Il n'existe pas à notre connaissance d'étude prouvant une polymérisation de nucléotides non activés sur surfaces minérales, contrairement à la polymérisation des acides aminés.

Mungi et al.<sup>50</sup> ont montré la possibilité de former des oligomères de nucléotides avec une polymérisation « assistée par lipides » avec des cycles hydratation/déshydratation à partir du nucléotides non activés. Le travail ne se prononce pas sur le mécanisme de cette polymérisation, d'ailleurs accompagnée de déglycosylation partielle.

### I.2.3. La question de la sélectivité de polymérisation

Dans le cadre d'une étude sur les origines de la vie, la principale interrogation est de comprendre comment des biomolécules fonctionnelles ont pu être obtenues à partir de molécules simples. Or, l'obtention préférentielle de certaines séquences plutôt que d'autres est cruciale pour mener à une structuration des chaînes peptidiques et à l'apparition d'une fonction. Même si l'on peut admettre une certaine « tolérance à l'erreur », un degré important de sélectivité d'incorporation des acides aminés dans le polymère en croissance doit être invoqué. Cette notion de sélectivité est d'un grand intérêt, de même que celle de l'homochiralité, c'est-à-dire l'apparition d'une sélection pour la forme L des résidus d'acides aminés<sup>25</sup>.

Eu égard à l'importance de la question, étonnamment peu d'études ont été réalisées concernant la sélectivité de stœchiométrie lors de la polymérisation des peptides. Fox et Harada ont obtenu des oligopeptides de séquences non aléatoires en chauffant un mélange d'acides aminés Glu et Asp<sup>51,52</sup>. Leyton et al.<sup>53,54</sup> ont observé des oligopeptides avec une séquence régulière Gly-Glu(-Gly-Glu)<sub>n</sub>, à partir d'un mélange de Gly et Glu et de nanoparticules d'oxyde métallique. A partir d'un mélange de Glu, Tyr et Gly activé à 180 °C pendant 24 h, Nakashima et al.<sup>55</sup> ont montré, en comparant les quantités de peptides obtenus avec des calculs de polymérisation aléatoire, la formation préférentielle de tripeptides pyroGlu-Tyr-Gly et pyroGlu-Gly-Tyr. Hartmann et al.<sup>56</sup> ont repris ces expériences et proposé un mécanisme expliquant cette sélectivité : lors du chauffage, la première réaction serait la cyclisation interne de Glu en pyroGlu, suivie de la formation d'un dipeptide cyclique entre les deux autres acide aminés (cyclo(Gly-Tyr)). Ensuite, l'acide pyroglutamique réagirait avec le dipeptide cyclique pour donner un tripeptide linéaire. Mosqueira et al.<sup>57</sup> ont

ensuite tenté d'expliquer par des modèles théoriques les résultats observés. Le modèle est appliqué à un mécanisme de polymérisation qui se fait *via* l'addition de dipeptides cycliques, les oligomères formés ont donc un nombre impair de résidus.

Quelques études sur surface ont montré l'obtention préférentielle de certaines séquences peptidiques en partant de mélanges d'acides aminés. Bujdák et al.<sup>45</sup> ont montré qu'à partir d'un mélange de Ala et Gly-Gly adsorbé sur montmorillonite, une séquence Ala-Gly-Gly est formée préférentiellement. Bujdák et Rode ont également prouvé la formation préférentielle des séquences d'acides aminés sur un support d'alumine non poreuse<sup>58</sup>. Les dipeptides Gly-AA et Ala-AA sont formés préférentiellement quand l'AA est Leu ou Val tandis que Ala-Gly et Pro-Gly sont préférentiellement formés. Cette sélectivité est expliquée par un effet inductif que les groupes de la chaîne latérale exercent sur la nucléophilie du groupe amine et l'électrophilie du groupe carboxyl (Gly-Val serait préféré car le groupe isopropyl de Val a un fort effet inductif). Implicitement, ils postulent donc que les proportions de différentes séquences seraient imposées par leurs stabilités thermodynamiques respectives, et donc que la polymérisation serait à l'équilibre chimique. Cela n'est pas certain du tout : l'origine des sélectivités pourrait très bien être cinétique.

Il existe donc des indications d'une possible sélectivité dans les scénarios de polymérisation des acides aminés sur surfaces minérales, mais elles sont restées « anecdotiques » et les explications d'un tel phénomène restent assez hypothétiques.

### I.3. Etudes expérimentales de la formation de peptides sur surfaces minérales

L'utilisation de surfaces minérales permet donc la polymérisation de biomolécules. Plusieurs surfaces prébiotiquement plausibles ont été envisagées<sup>59,60</sup>. Parmi elles, celle de la silice amorphe a été majoritairement utilisée dans ce travail et très souvent dans la littérature.

#### I.3.1. Les surfaces de silice

La silice (SiO<sub>2</sub>) possède de très nombreux polymorphes, la plupart basés sur des tétraèdres (SiO<sub>4</sub>) connectés par des liaisons siloxanes (Si-O-Si). Certains sont des solides cristallins non

poreux (quartz, tridymite, eddingtonite...), d'autres des solides cristallins microporeux (les zéolithes) ou des solides mésoporeux dont les parois sont amorphes mais les pores périodiquement disposés. Plusieurs de ces polymorphes ont pu être présents sur la Terre archéenne, et par ailleurs la silice est souvent utilisée comme modèle des silicates qui composent la majeure partie de la croûte terrestre. De nombreuses phases de la silice ont fait l'objet d'une caractérisation extensive<sup>61</sup>, notamment en ce qui concerne leurs propriétés d'adsorption de petites molécules organiques. Elles possèdent de nombreux points communs. Par exemple, l'interruption par la surface du réseau de liaisons siloxanes crée des « liaisons pendantes »  $\equiv\text{Si}-\text{O}^\cdot$  ou  $\equiv\text{Si}^\cdot$  qui, en conditions hydratées (comme c'est normalement le cas à la surface de la Terre), sont « guéries » par réaction avec des molécules d'eau, aboutissant à la formation de groupements silanols  $\equiv\text{Si}-\text{OH}$ . Une surface de silice expose donc des groupements siloxanes et silanols.

Une caractérisation plus détaillée révèle l'existence de plusieurs groupements différents à la surface (voir Figure I-3). Du point de vue de la connectivité des liaisons covalentes, on peut distinguer les silanols géminaux (deux groupes  $-\text{OH}$  portés par le même silicium) et les silanols terminaux et, parmi ces derniers les silanols isolés et les silanols vicinaux (deux groupes  $-\text{OH}$  sur deux Si connectés par une liaison siloxane  $\text{Si}-\text{O}-\text{Si}$ ). Du point de vue du voisinage spatial, les silanols peuvent être en mesure d'effectuer des liaisons H avec leurs plus proches voisins (on parlera de silanols associés) ou non (cette situation inclut la plupart des silanols isolés).

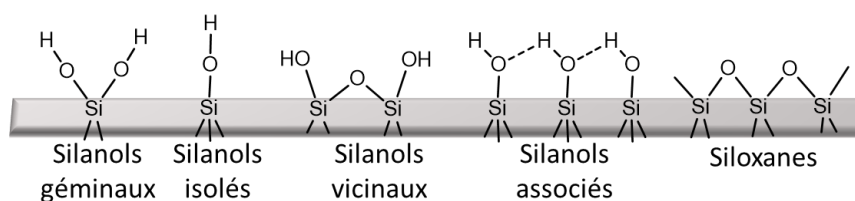


Figure I-3. Différents groupes de surface d'une silice amorphe.

La surface utilisée dans nos travaux est une silice amorphe non poreuse commercialisée sous le nom d'Aerosil380. Il s'agit d'une silice « fumée », obtenue par calcination à la flamme de  $\text{SiCl}_4$ , de surface spécifique élevée ( $380 \text{ m}^2/\text{g}$ ). Ces silices, par rapport à d'autres formes telles que les silices de précipitation, se caractérisent par une densité de silanols relativement

faibles (environ 4 par nm<sup>2</sup>)<sup>62</sup>, et une quasi absence de silanols internes. Lorsqu'elles sont calcinées, on observe une déshydroxylation due à la condensation progressive des silanols en siloxanes, donc une diminution de la quantité de silanols résiduels, mais ce phénomène ne se manifeste qu'à partir de 250 °C. A température et sous humidité ambiantes, la surface des silices Aerosil est recouverte de plusieurs couches d'eau. Cette eau physisorbée peut être éliminée par chauffage dès 100 °C (la température précise varie selon la vitesse de chauffage, et selon que celui-ci est réalisé sous vide ou sous flux de gaz).

### I.3.2. Historique des investigations sur les systèmes acides aminés/silice

Comme dit précédemment, la possibilité théorique que les surfaces minérales aient joué un rôle dans la polymérisation des acides aminés a été évoquée par Bernal dès les années 1950. Ce n'est qu'une vingtaine d'années plus tard qu'elle a été testée expérimentalement, d'abord par Paecht-Horowitz et al.<sup>63,64</sup>, sur des adénylates d'acides aminés, puis sur des acides aminés non activés<sup>65</sup>. Dans les deux cas, les phases minérales utilisées étaient des argiles, respectivement des montmorillonites et kaolinites.

Pendant la décennie 1990-2000, l'étude de l'adsorption-polymérisation des acides aminés sur surface minérale a principalement été poursuivie par deux groupes. Nous pouvons citer tout d'abord les travaux de Basiuk, le premier à réaliser la polymérisation d'acides aminés sur silice<sup>66-72</sup> (et dans une moindre mesure sur alumine). Ces auteurs ont déposé une douzaine d'acides aminés différents sur des surfaces de silice pressées en pastilles auto-cohérentes par sublimation sous vide suivie de dépôt en phase vapeur (CVD). Une telle approche semble moins pertinente pour la chimie prébiotique que l'adsorption des molécules depuis une phase aqueuse<sup>28</sup>, mais elle leur permettait notamment de caractériser les produits de réaction par spectroscopie IR *in situ*. Le produit majoritaire observé dans la plupart des cas consiste en un dipeptide cyclique, une dicétopipérazine (DKP, voir structure sur la Figure I-4), avec un rendement supérieur à 80 %. Sur alumine, de petits peptides linéaires sont aussi observés.



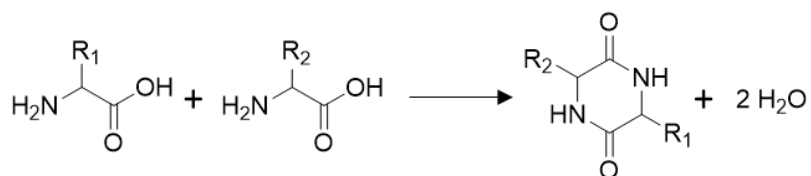


Figure I-4. Formation d'une dicétopipérazine (DKP) par condensation cyclique de deux acides aminés.

A la même époque, Bujdák et Rode ont étudié la formation de liaisons peptidiques sur argiles, en les comparant occasionnellement avec la silice et l'alumine<sup>73-75</sup>. Ces études ont été réalisées avec des scénarios complexes combinant cycles hydratation/déshydratation et variations de température, dans des conditions où le rendement global en peptides restait très faible (températures  $\leq 80^{\circ}\text{C}$ ). Malgré leur caractère empirique, les données obtenues par ces auteurs sur les sélectivités de polymérisation sont intéressantes, comme nous l'avons mentionné plus haut (cf Chapitre I.1.2.3). Contrairement aux travaux de Basiuk et al. (mais conformément à ceux de Lahav sur argiles), les acides aminés étaient déposés à partir de solutions aqueuses, un scénario assez raisonnable par rapport aux conditions de la Terre archéenne.

A partir des années 2000, des études visant à caractériser au niveau moléculaire l'interaction entre acides aminés et groupes de surface de silice apparaissent. Ces études incluent les premiers travaux réalisés au LRS sur les systèmes glycine/SiO<sub>2</sub>, acide glutamique/SiO<sub>2</sub>, et dans une moindre mesure lysine/SiO<sub>2</sub>, utilisant des techniques analytiques comme la spectroscopie UV-visible, l'analyse thermogravimétrique et les spectroscopies vibrationnelles<sup>76,77</sup>, ainsi que la RMN du solide<sup>78-81</sup>. On peut aussi citer les articles de Holland et al. portant sur la lysine<sup>82</sup> et l'alanine<sup>83,84</sup> supportées sur silice et pour l'usage de techniques de RMN avancées, ceux de Ben Shir et al.<sup>85-87</sup>. Tous ces travaux, contrairement à ceux des années 1990, se concentrent sur des systèmes simples comportant un seul acide aminé soumis à une seule activation thermique, afin d'arriver à une compréhension précise de la nature de l'adsorption et de la formation des liaisons peptidiques. En revanche, ces systèmes ne sont pas très prometteurs du point de vue prébiotique. Bien que la formation de liaisons peptidiques ait lieu à températures modérées (120 à 160 °C), les produits obtenus sont presque exclusivement des dipeptides cycliques.

Pour les objectifs de notre travail, il est particulièrement important de mettre en évidence les cas où la formation de peptides linéaires plus ou moins longs a été observée, évitant

« l'impasse DKP ». L'équipe de Martra et al.<sup>88</sup> a rapporté en 2014 la formation d'oligomères de glycine jusqu'aux 11-mères après dépôt par CVD sur silice – les raisons de la différence de comportement de la glycine avec les travaux de Basiuk et al. cités plus haut ne sont pas encore claires. Par ailleurs, Sakhno et al.<sup>89</sup> ont observé, par désorption suivie de spectrométrie de masse (ESI), la formation d'oligopeptides jusqu'aux heptapeptides pour le système mixte (leucine + acide glutamique) sur une surface SiO<sub>2</sub>, après activation thermique simple. Cette étude, qui est l'une des premières utilisant un système à plusieurs acides aminés depuis Bujdák et Rode, a ainsi montré la possibilité de générer des biomolécules plus intéressantes à partir de l'hétérogénéité chimique, plus réaliste d'un point de vue prébiotique. Ces résultats préliminaires ont constitué la raison principale du choix de ce système pour nos premières expériences.

### I.3.3. Mécanisme d'adsorption des acides aminés sur la silice

*A priori*, le mécanisme d'interaction d'un acide aminé avec la silice pourrait se faire de manière électrostatique, par formation de liaisons covalentes, par liaison hydrogène ou par une combinaison de ces liaisons, impliquant éventuellement une forme de reconnaissance moléculaire<sup>28</sup>. L'adsorption électrostatique dépend de la spéciation acido-basique de l'acide aminé et de la surface. Comme les surfaces de silice ne peuvent pas être chargées positivement, l'adsorption électrostatique ne peut concerner qu'une surface négative (donc à pH  $\geq 3,5$ ) et un acide aminé positivement chargé tel que l'arginine ou la lysine, à un pH inférieur à son second pKa, ce qui restreint les conditions.

Plusieurs auteurs ont proposé que les acides aminés sont adsorbés par formation de liaisons acyl ou ester avec les silanols de surface (Si-O-CO-R), notamment White et al.<sup>90</sup>. La question reste très controversée. Pour ce qui concerne la leucine et l'acide glutamique, il a été montré par RMN que l'adsorption impliquait plutôt la formation d'un réseau de liaisons hydrogène entre les fonctions organiques de l'acide aminé, les silanols de la surface de silice, mais aussi un certain nombre de molécules d'eau coadsorbées<sup>91</sup>. Quand l'activité de l'eau diminue, il semble que la fonction ammonium interagisse directement avec les silanols tandis que la fonction carboxylate est liée avec une molécule d'eau « pontante », elle-même liée à la surface. L'effet coopératif de nombreuses liaisons hydrogène aboutit à une adsorption assez forte, et peut d'autre part activer significativement la molécule adsorbée,

d'une façon analogue à ce qu'on observe sur le site réactionnel de certains enzymes.

Une fois déposés sur la surface à partir d'une solution aqueuse, les acides aminés sont sous forme zwitterionique, comme ils le sont en solution<sup>89</sup>. Cela a pu être déterminé par analyse des spectres infrarouge, les principaux groupes fonctionnels (élongations des COO<sup>-</sup> et déformations des NH<sub>3</sub><sup>+</sup> et CH<sub>2</sub>) des acides aminés déposés sur surface ayant des modes de vibration observables. La modélisation moléculaire aboutit par ailleurs à la conclusion que la forme zwitterionique est plus stable que la forme neutre dès qu'un nombre minimum de molécules d'eau sont coadsorbées<sup>91,92</sup>, alors que la forme neutre est plus stable sur une surface complètement anhydre.

### 1.3.4. Adsorption et polymérisation d'acides aminés sur d'autres surfaces

Plusieurs types de supports minéraux ont été utilisés pour les réactions de condensation peptidique.

Les premiers, et les plus fréquemment utilisés, sont des argiles. D'après Bujdák et Rode, la silice et l'alumine sont plus efficaces que les argiles pour la dimérisation ; les argiles seraient quant à elles plus efficaces pour l'élongation des chaînes peptidiques<sup>73</sup>. Les structures argileuses étant extrêmement diverses et par nature hétérogènes, cette généralisation est sans doute prématurée. En effet, les acides aminés intercalés entre les feuillets argileux se trouvent dans un environnement très différent de ceux qui sont adsorbés sur les bords des feuillets, et on peut supposer qu'ils ne réagissent pas de la même façon malgré le peu de données existantes sur ce sujet. La caractérisation au niveau moléculaire, qui progresse lentement pour les systèmes AA/SiO<sub>2</sub>, reste presque inexistante pour les systèmes AA/argile. Deux études ont été réalisées au LRS sur ces derniers. Dans la première<sup>93</sup>, l'acide glutamique et l'arginine ont été coadsorbés sur une montmorillonite. Les deux acides aminés étaient apparemment adsorbés séparément, l'un (Arg) dans les feuillets, et l'autre (Glu) sur les bords, aboutissant à une ségrégation spatiale. Les deux acides aminés ne pouvaient alors réagir que dans la « zone de contact ». La polymérisation avait lieu à des températures nettement plus élevées que sur silice, une tendance confirmée pour d'autres systèmes AA/argile<sup>94</sup>. Dans la seconde étude<sup>95</sup>, l'alanine a été activée après adsorption sur hectorite

(argile TOT) et sur kaolinite. La formation de liaisons peptidiques a été observée sur hectorite, à température élevée également, mais ne s'est pas produit sur kaolinite.

Dans le présent travail, la polymérisation d'acides aminés a été testée sur deux autres surfaces, la montmorillonite et l'alumine  $\gamma$ , discutées plus bas. L'efficacité de l'alumine pour la formation de liaisons peptidiques est établie, mais peu d'études se sont focalisées sur ce support.

Plusieurs autres surfaces minérales ont été testées pour la polymérisation des AA. Sans vouloir être exhaustif, elles incluent le  $\text{TiO}_2$ <sup>88,96</sup>, les oxyhydroxydes de fer<sup>97,98</sup>, certains sulfures<sup>99</sup>... Tout porte à croire qu'un même acide aminé sur différentes surfaces peut subir des évolutions très différentes.

Le choix d'une surface particulière est dicté par plusieurs considérations différentes, dont la facilité d'étude, mais il est évidemment préférable que la surface en question ait été effectivement présente sur la Terre primordiale. Les minéraux présents sur la terre primitive étaient surtout des silicates, des oxydes et des sulfures. Une liste extensive de « proto-minéraux » a été établie par Hazen<sup>59</sup>. Elle contient plusieurs polymorphes de la silice, qui répond donc à ce « critère prébiotique ».

Certains minéraux actifs pour la polymérisation pourraient avoir joué aussi d'autres rôles dans l'origine de la vie. Wächtershäuser a émis l'hypothèse que la pyrite peut servir de source d'énergie pour la formation des premiers cycles métaboliques<sup>36</sup>.

## II. Outils analytiques pour l'étude de la condensation peptidique

La polymérisation d'acides aminés sur une surface minérale étudiée dans ce travail résulte en la formation d'un mélange complexe d'oligopeptides. Les méthodes *in situ* les plus couramment utilisées pour caractériser ces produits sont la spectroscopie infrarouge, la spectroscopie RMN du solide et l'analyse thermogravimétrique. Une fois désorbés de la surface, les mélanges d'oligopeptides en solution peuvent être analysés avec les méthodes analytiques plus puissantes, développées en protéomique comme la chromatographie, la spectrométrie de masse ou la spectroscopie de résonance magnétique en solution.

## II.1. Analyses *in situ* des produits de polymérisation sur surface de silice

Différentes méthodes de caractérisation ont été utilisées dans ce projet telles que la diffraction des rayons X (DRX), l'analyse thermogravimétrique (ATG) et la spectroscopie infrarouge (IR).

La diffraction des rayons X sur poudre (DRX) permet d'étudier la cristallinité d'un matériau. Une source de rayons X envoie un rayonnement monochromatique. Les électrons des atomes de l'échantillon diffusent le rayonnement incident. Si l'échantillon comporte des phases cristallines, les rayonnements diffusés par les atomes des plans réticulaires interfèrent et créent un phénomène de diffraction. La loi de Bragg permet d'attribuer un pic à une famille de plans atomiques selon la relation  $2d_{hkl}\sin(\theta) = n\lambda$  où  $d_{hkl}$  est la distance entre les plans d'indices de Miller (h,k,l), n l'ordre de diffraction,  $\theta$  l'angle de diffraction et  $\lambda$  la longueur d'onde du rayonnement utilisé.

La diffraction des rayons X est utilisée pour vérifier l'adsorption des acides aminés à la surface de la silice. En effet, dans les cas les plus simples, chaque acide aminé présente un « recouvrement de saturation » qui lui est propre : une capacité d'adsorption maximale au-delà de laquelle les acides aminés n'interagissent plus avec la surface mais précipitent en masse. Dans le cas de l'acide glutamique, la situation est plus complexe ; des cristallites de petites tailles qui ne sont pas détectables en DRX, semblent se former dès l'adsorption des acides aminés<sup>91</sup>. Dans la mesure du possible, on veut éviter la précipitation afin d'avoir un état homogène dans lequel tous les acides aminés interagissent avec la surface et réagissent *a priori* de la même façon.

L'analyse thermogravimétrique (ATG) consiste en la mesure de la variation de masse d'un échantillon en fonction du temps avec application d'une rampe de température contrôlée (généralement linéaire). La condensation des acides aminés en peptides impliquant la formation d'une liaison amide, l'ATG permet d'observer cette réaction en mesurant la libération d'une molécule d'eau en phase gazeuse en fonction de la température appliquée.

La spectroscopie infrarouge (IR) est une technique basée sur l'absorption du rayonnement infrarouge par le matériau analysé. Pour être active en infrarouge, une vibration doit avoir un moment dipolaire permanent non nul. C'est une technique permettant de déterminer les fonctions chimiques présentes dans un matériau, elle est appropriée pour confirmer la

présence de la liaison peptidique formée après activation thermique par l'observation des bandes d'absorption de la fonction amide. La spectroscopie infrarouge permettant uniquement d'identifier des groupes fonctionnels, elle ne donne pas d'indications facilement interprétables au sujet de la longueur des oligopeptides formés ni au sujet de leur stœchiométrie, deux questions fondamentales pour notre étude.

Par ailleurs, la spectroscopie de résonance magnétique nucléaire (RMN) est une technique très puissante pour déterminer les structures des composés organiques. Cette technique consiste à observer les transitions entre deux niveaux d'énergie d'un noyau soumis à un champ magnétique. Pour qu'il y ait une interaction avec le champ magnétique, le noyau atomique observé doit avoir un spin nucléaire non nul<sup>100</sup>. Elle est couramment utilisée pour la détermination structurale de molécules biologiques mais est difficile à interpréter pour les mélanges complexes et dans ce dernier cas, nécessite l'utilisation de composés standards comme référence. Elle peut être utilisée pour des échantillons à l'état solide (ce qui a l'avantage d'éviter une étape de désorption pour une polymérisation sur surface) mais a une résolution plus limitée qu'à l'état liquide. Cette technique analytique est par ailleurs peu sensible et nécessite souvent de longs temps d'accumulation. Nous n'avons pas utilisé cette technique mais privilégié l'utilisation de la spectrométrie de masse à haute résolution pour l'analyse de nos systèmes complexes après polymérisation et désorption en solution.

## II.2. Méthodes séparatives pour l'analyse de mélanges de peptides

Aucune technique de caractérisation *in situ* n'est donc vraiment appropriée pour identifier précisément la composition d'un mélange de peptides adsorbés sur une surface minérale. Pour ce faire, il est nécessaire de passer par une étape de désorption de la matière organique. Cette étape supplémentaire peut entraîner une modification de la spéciation des peptides supportés, mais elle ouvre la possibilité d'utiliser des techniques plus sensibles comme la chromatographie et la spectrométrie de masse. Par ailleurs, d'un point de vue prébiotique, l'évolution chimique ultérieure de peptides formés sur une surface minérale doit inévitablement passer par leur désorption.

## II.2.1. Les techniques de chromatographie

Différentes techniques chromatographiques sont utilisées pour la séparation de peptides en général et applicables au contexte de la condensation d'oligomères. Par ailleurs, les méthodes séparatives peuvent être couplées à la spectrométrie de masse, permettant une analyse exhaustive de tels échantillons complexes.

La chromatographie liquide en phase inverse (RP-HPLC) est la technique la plus couramment utilisée. En outre, le couplage avec la MS, notamment l'ESI comme méthode d'ionisation, constitue la méthode de choix pour l'analyse de mélanges complexes. La séparation est fondée sur des interactions hydrophobes entre la molécule et la phase stationnaire. La phase stationnaire étant apolaire et la phase mobile polaire et hydrophile, la séparation des composés se fait en fonction de leur polarité : les composés les plus polaires, moins retenus sur la colonne sont élués en premier. La force d'éluion augmente en augmentant le pourcentage de solvant organique. La RP-HPLC s'applique généralement à la séparation de petites molécules apolaires (lipides, peptides...). Le type de greffage de la silice de la colonne peut influencer les propriétés de rétention. Ainsi certaines colonnes C8 sont plus rétentives que des colonnes C18. Il est nécessaire de tamponner la phase mobile pour l'étude de certains composés ionisables. L'acide formique et l'acétate d'ammonium sont des tampons usuels utilisables en couplage avec la spectrométrie de masse. Le pH choisi doit être suffisamment différent du pKa de la molécule analysée pour analyser l'espèce soit sous sa forme ionisée, soit sous sa forme libre, et ne doit pas être trop basique pour ne pas abimer la colonne.

La chromatographie de phase inverse a été utilisée par Yu et al. en couplage avec la spectrométrie de masse pour des mélanges de depsipeptides produits avec valine/acide lactique<sup>101</sup>. Forsythe et al.<sup>102</sup> ont également utilisé la phase inverse avec un couplage LC/IM-MS/MS pour identifier les oligomères mixtes de Gly, Ala et d'acide glycolique allant jusqu'à 8 résidus acides aminés.

D'autres conditions de séparation peuvent être utilisées. Par exemple, la chromatographie à interaction hydrophile (HILIC, Hydrophobic Interaction Liquid Chromatography) est une méthode permettant de pallier la forte polarité des composés, qui ne sont pas retenus en phase inverse : la phase stationnaire est polaire et la phase mobile est composée d'eau et

d'un solvant organique miscible dans l'eau<sup>103</sup>. Cette technique a été utilisée pour l'étude de systèmes proches des nôtres, notamment pour séparer des mélanges de Gly, Gly<sub>2</sub> et Gly<sub>3</sub><sup>104</sup> ou des depsipeptides<sup>105</sup>. La chromatographie d'échange d'ions, basée sur un processus d'échange entre la phase mobile et des groupements fonctionnels ionisés greffés sur le support solide, permet de séparer des espèces avec de très faibles différences de propriétés. Cette technique (et en particulier de type échange de cations) a par exemple été utilisée pour étudier la rétention de peptides courts contenant de deux à neuf résidus acides aminés ayant des propriétés physicochimiques variées<sup>106</sup>.

Enfin, la chromatographie d'appariement d'ions (ou chromatographie de paires d'ions, IP-HPLC) est une combinaison de chromatographie de partage inverse et de chromatographie d'échanges d'ions. La phase stationnaire utilisée est de type C8 ou C18 et un agent d'appariement d'ions est ajouté à la phase mobile, ce qui permet aux molécules ioniques de former des paires d'ions dans la phase mobile pour devenir électriquement neutres. La rétention des composés se fait alors par interactions électrostatiques et hydrophobes. Cette technique a notamment été utilisée pour séparer des oligomères de lysine<sup>107,108</sup>, de glycine<sup>109-113</sup>, ou pour séparer la glutamine du glutamate<sup>114</sup>.

## II.2.2. Utilisation de la mobilité ionique pour une analyse rapide de mélanges complexes

La mobilité ionique est une technique permettant de séparer des molécules en fonction des mesures de section efficace de collision (collision cross section, CCS, aussi notée  $\Omega$ ) d'ions en phase gazeuse. Cette technique séparative peut être couplée à d'autres techniques analytiques comme la spectrométrie de masse qui sert alors à la détection des ions. Le couplage IM-MS (développé par Clemmer<sup>115</sup> en 1998) permet une séparation des analytes, en fonction de leur taille et leur charge dans un premier temps par mobilité ionique, suivie d'une séparation en fonction de m/z par SM, ce qui permet d'accéder à une analyse rapide d'analytes en deux dimensions simultanément<sup>116</sup>. Cette technique a démontré une grande capacité de séparation d'espèces isobariques et isomères, que la spectrométrie de masse seule ne peut pas distinguer. Un autre avantage est la possibilité de coupler cet instrument IM-MS avec la chromatographie liquide en amont (soit une analyse multidimensionnelle 3D).



### II.2.2.1. Principe général de la spectrométrie de mobilité ionique

La spectrométrie par mobilité ionique (IMS) est une technique séparative d'ions en phase gazeuse. Elle peut être assimilée à la séparation électrophorétique, à la différence que la séparation s'effectue en phase gazeuse. La Figure I-5 montre un schéma du principe.

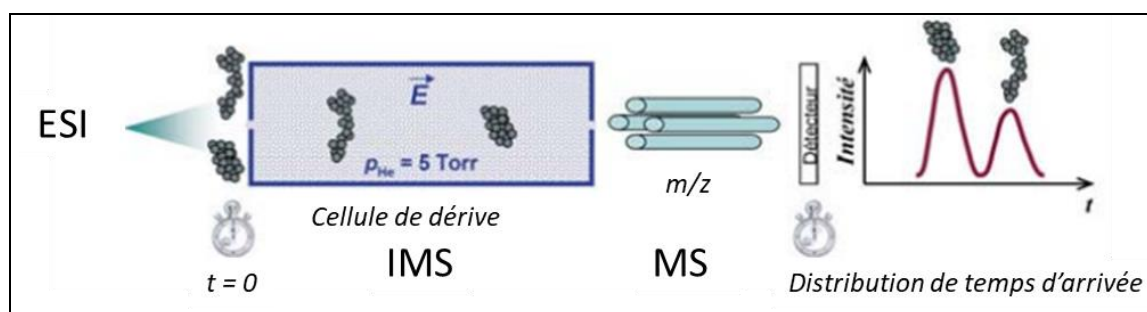


Figure I-5. Principe de la méthode de spectrométrie de mobilité ionique (IMS) couplée à la spectrométrie de masse (MS). D'après Gabelica et al.<sup>117</sup>, modifié.

Les molécules sont d'abord ionisées (électrospray dans notre cas, mais en théorie toutes techniques d'ionisation peuvent être utilisées) puis introduites dans la cellule de dérive afin d'être séparées sous l'action combinée d'un champ électrique et des collisions avec le gaz tampon. Le mobilogramme (aussi appelé spectre de mobilité) d'un échantillon rapporte l'intensité en fonction de la mobilité  $K$  (ou  $1/K$ , la mobilité réduite). La mobilité  $K$  d'un ion en phase gazeuse (exprimée en  $\text{cm}^2 \cdot \text{V}^{-1} \cdot \text{s}^{-1}$ ) est la mesure de son parcours au travers d'un gaz tampon sous l'action d'un champ électrique. Elle se calcule par la relation (1), dans laquelle  $t_d$  est le temps de dérive de l'ion,  $l$  la longueur du tube de dérive et  $E$  le champ électrique appliqué.

$$K = \frac{v_d}{E} = \frac{l}{t_d E} \quad (1)$$

L'équation de Mason-Schamp<sup>118</sup> relie la mobilité  $K$  et la section efficace de collision  $\Omega$  :

$$K = \frac{3\sqrt{2\pi}}{16} \cdot \frac{q}{N} \cdot \left(\frac{1}{m} + \frac{1}{M}\right)^{1/2} \cdot \left(\frac{1}{k_B T}\right)^{1/2} \cdot \frac{1}{\Omega} \quad (2)$$

où  $q$  est la charge ionique,  $N$  la densité numérique du gaz utilisé,  $m$  la masse moléculaire de l'ion,  $M$  la masse moléculaire du gaz neutre,  $k_B$  la constante de Boltzmann et  $T$  la température. Cette équation n'est valable que pour des champs électriques faibles, car lorsque le champ électrique est élevé,  $K$  dépend de la valeur de  $E$ , et  $v_d$  n'est plus proportionnelle à  $E$ . La densité du gaz et la température étant des paramètres qui varient le long de la cellule de dérive (en raison de phénomènes de diffusion), la mobilité  $K$  d'un ion change également le long de la cellule.

Une CCS reflète la conformation d'une molécule et est, dans des conditions d'analyse déterminées, spécifique d'une paire ion-gaz donnée<sup>119</sup>. Cette grandeur peut être utilisée pour identifier la molécule, si bien que des travaux se sont penchés sur la possibilité de prédire pour une structure donnée les valeurs théoriques de CCS.

Il existe différents types de mobilité ionique reposant sur ce principe, parmi lesquels ceux utilisant une séparation temporelle et spatiale (DTIMS-mobilité ionique dans un tube de dérive, TWIMS-mobilité à champ oscillant), des filtres en mobilité (FAIMS-mobilité ionique à haut champ asymétrique) ou le piégeage et la libération successive d'ions (TIMS-mobilité ionique à piégeage d'ions). La technique utilisée dans le présent travail est le TIMS (voir le principe de la cellule TIMS dans le Chapitre II).

### II.2.2.2. Etude des peptides par mobilité ionique

Bien que moins utilisée que la chromatographie liquide, la mobilité ionique est une technique relativement courante en tant que méthode de séparation de mélanges de peptides et/ou de protéines<sup>116,120</sup>. Par ailleurs, plusieurs études ont utilisé la mobilité ionique pour différencier des composés énantiomères. Domalain et al.<sup>121,122</sup> ont ainsi décrit la séparation par TWIMS d'un mélange d'énantiomères de phénylalanine, tyrosine et tryptophane par complexation en présence de Cu(II). Dwivedi et al.<sup>123</sup> ont rapporté la séparation en phase gazeuse de deux énantiomères R et S du tryptophane en utilisant un gaz de dérive dopé par un gaz chiral.

Si on ne dispose pas de standards, plusieurs méthodes sont utilisées pour prédire les CCS. La procédure pour comparer les valeurs de CCS calculées théoriquement et les valeurs déterminées expérimentalement est la suivante : génération d'une liste de conformères ;

minimisation des structures en prenant les structures de plus faible énergie ; calcul des CCS théoriques ; comparaison entre les valeurs CCS théoriques et expérimentales<sup>124</sup>.

## II.3. Analyse par spectrométrie de masse à ultra haute résolution

Dans le présent travail, les mélanges de peptides obtenus par condensation d'acides aminés sont relativement complexes, en ce sens qu'ils comportent jusqu'à plusieurs centaines de composés différents. La technique analytique la plus appropriée pour permettre la distinction des composés présents dans de tels mélanges est la spectrométrie de masse à haute résolution, seule ou en couplage avec des méthodes séparatives telles que celles qui ont été présentées au § II.

### II.3.1. Description du spectromètre de masse FTICR

Un spectromètre de masse est composé d'une source d'ions, d'un analyseur en masse sous vide et d'un détecteur. Les ions produits dans la source sont transmis à l'analyseur en masse. L'analyseur en masse sépare les ions selon leur rapport  $m/z$ . Le détecteur permet de compter le nombre d'ions qui l'atteignent.

L'appareil principalement utilisé dans ce projet doctoral est de type FTICR (Fourier Transform Ion Cyclotron Resonance). L'instrument est un appareil hybride QhFTICR 7 T (SolarixXR, BrukerDaltonics, Germany) (voir Chapitre II). Les oligopeptides sont ionisés à l'aide de la source électrospray (ESI).

#### II.3.1.1. Source électrospray

L'ionisation électrospray (ESI) est une méthode de désorption/ionisation dite « douce », en ce sens qu'elle permet de produire des ions moléculaires intacts à partir de molécules polaires, de masses importantes et thermiquement labiles en solution. Elle est donc parfaitement adaptée à l'analyse d'oligomères d'acides aminés. Dans le cas d'une source électrospray, la transmission des ions formés à pression atmosphérique se fait par l'intermédiaire d'une interface avec un vide intermédiaire. L'ionisation a lieu en plusieurs

étapes : la nébulisation de la solution en gouttelettes électroniquement chargées sous l'action d'un champ électrique, la libération des ions à partir des gouttelettes et le transport des ions de la région à pression atmosphérique jusqu'au vide de l'analyseur *via* une région de vide intermédiaire<sup>125</sup>. Le nombre d'ions qui arrivent au détecteur est dépendant de l'efficacité d'ionisation ESI, mais aussi de l'efficacité de transmission des ions au travers de systèmes d'optiques et jusqu'à l'analyseur en masse et au détecteur.

### II.3.1.2. Analyseur en masse à haute résolution à transformée de Fourier

Il existe deux grands types d'instruments à transformée de Fourier : l'Orbitrap et le FTICR (Fourier Transform Ion Cyclotron Resonance). Dans les deux cas, les ions sont piégés dans l'analyseur en masse sous ultra-vide ( $10^{-10}$  mbar). Les paquets d'ions adoptent un mouvement harmonique qui peut être mesuré sous forme de signal transitoire (Figure I-6). Le signal obtenu dans le domaine temporel est transformé dans le domaine des fréquences (combinaison de toutes les fréquences caractérisant les mouvements dans une direction de l'espace des ions acquis en même temps) par transformée de Fourier, puis converti en valeurs de  $m/z$  (inversement proportionnelles à la fréquence)<sup>126</sup>.

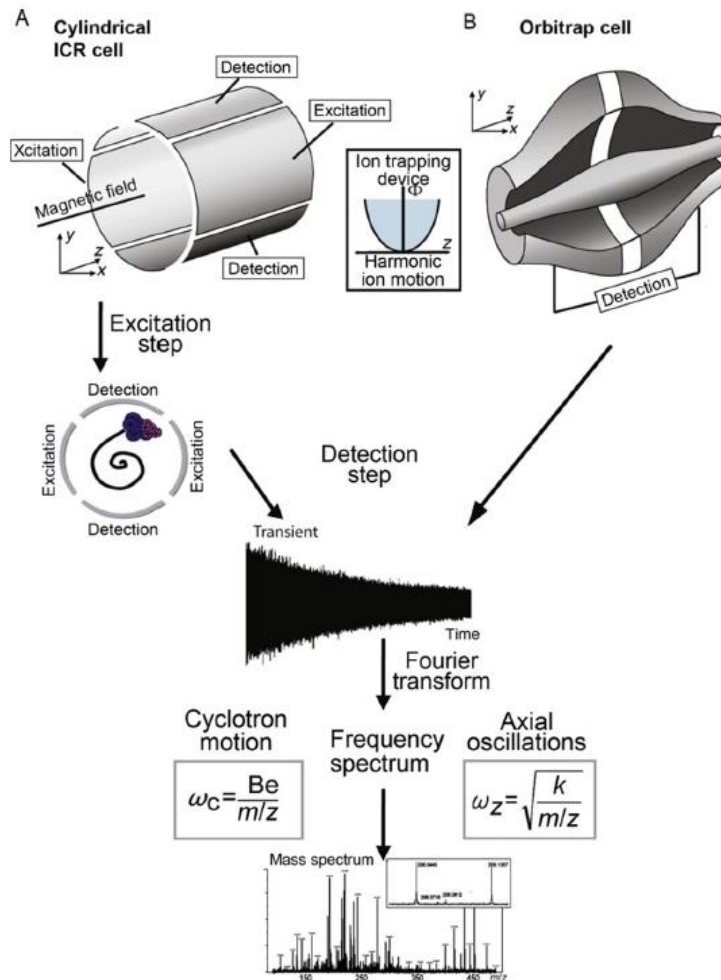


Figure I-6. Principes de fonctionnement des instruments à transformée de Fourier (A) cellule ICR, (B) Orbitrap. D'après Alves et al.<sup>126</sup>

Le principe de piégeage des ions est différent entre les deux technologies FTMS. L'Orbitrap n'utilise qu'un champ électrostatique alors que dans une cellule ICR, les ions sont piégés sous l'action d'un champ magnétique et d'un champ électrostatique avant d'être excités à l'aide d'un potentiel électrique alternatif. Au sein de la cellule ICR, le mouvement cyclotronique est mesuré. Il consiste en la fréquence de rotation de l'ion autour de son orbite dans un champ magnétique en l'absence de champ électrique<sup>127</sup>. Cette fréquence, notée  $\omega_c$  et exprimée en rad/s, est caractéristique du rapport masse/charge de l'ion ; elle se calcule par la formule  $\omega_c = \frac{qB}{m}$ , où  $q$  est la charge électrique (en Coulombs),  $B$  le champ magnétique (en Tesla) et  $m$  la masse de l'ion (en kg).

Le courant produit par le mouvement cyclotronique des ions est ensuite détecté avant d'être transformé dans le domaine fréquentiel. La durée de détection doit être grande devant la

période de rotation des ions pour obtenir une bonne précision. Les performances de l'appareil sont fortement liées aux caractéristiques de l'aimant ; les deux paramètres clés sont l'intensité et l'homogénéité du champ magnétique.

Les détails de la cellule utilisée sont expliqués dans le Chapitre II.

### II.3.2. Critères d'une analyse en haute résolution

Les critères permettant de définir les performances d'un analyseur en masse sont le pouvoir de résolution en masse, la précision en masse, la sensibilité et la gamme dynamique.

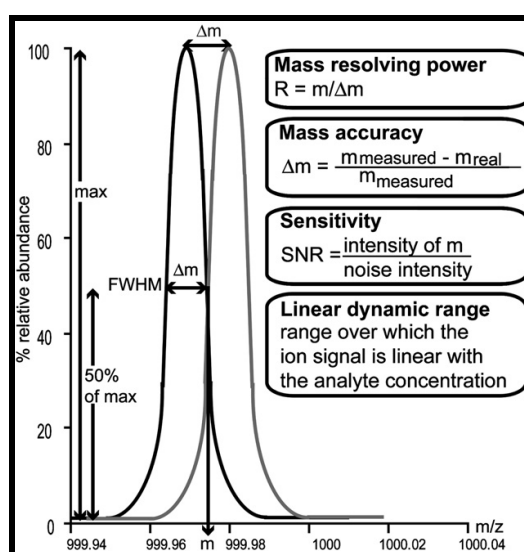


Figure I-7. Paramètres utilisés pour décrire les performances d'un analyseur<sup>128</sup>.

La résolution est la capacité d'un spectromètre de masse à distinguer deux ions ayant de rapports  $m/z$  proches. Pour  $z=1$ , le pouvoir de résolution en masse ( $R_p$ ) est le rapport  $m/\Delta m$  où  $m$  est la masse de l'ion donné.  $\Delta m$  peut être déterminé par deux méthodes différentes : l'écart entre les deux *maxima* de deux pics consécutifs (définition de la vallée) ou la largeur à mi-hauteur (FWHM : full-width half-height maximum) (voir Figure I-7).

La précision en masse est le calcul de l'erreur expérimentale dans la mesure des rapports  $m/z$ . En détection à haute résolution, elle est souvent exprimée en parties-par-million (ppm) et se calcule par la formule  $(\Delta M/M) \times 10^6$  avec  $\Delta M$  la différence entre la masse théorique (notée  $M$ ) et celle mesurée expérimentalement. La précision en masse est la capacité du spectromètre de masse à mesurer expérimentalement la valeur  $m/z$  exacte d'un composé

donné ; il est souhaitable qu'elle ait une précision de 3 à 5 décimales. La précision en masse est directement liée au pouvoir de résolution en masse accessible pour un analyseur en masse donné.

La sensibilité d'un instrument correspond à son rapport signal/bruit. Sa gamme dynamique correspond à sa capacité à distinguer les signaux de faibles intensités par rapport à ceux ayant une intensité plus importante.

Les instruments FTICR présentent les meilleures performances en termes de pouvoir résolutif et de gamme dynamique. C'est la raison pour laquelle une stratégie analytique basée sur des analyses directes et à haut débit utilisant la HRMS a été mise en œuvre dans ce travail.

### II.3.3. Phénomènes de discriminations en spectrométrie de masse

Plusieurs phénomènes ayant lieu dans les différentes parties du spectromètre de masse sont à l'origine de discriminations entre les réponses d'un mélange de molécules : discrimination à l'ionisation, effets de matrice, effets de saturation au détecteur etc. Ces limitations impactent l'analyse de mélanges complexes en ce sens qu'elles limitent l'exhaustivité de la caractérisation, et surtout la possibilité de quantification des composés détectés.

Le premier facteur est la variabilité de l'efficacité d'ionisation entre les différents composés. Ainsi, les analytes hydrophobes ont une meilleure réponse électrospray que les composés polaires. Cela est dû à leur grande affinité pour la surface des gouttelettes de l'électrospray<sup>129</sup>. Ce phénomène de discrimination à l'ionisation est accentué pour les mélanges complexes, où l'on observe une compétition à l'ionisation. La présence de composés moins volatils peut modifier l'efficacité de la formation de gouttelettes<sup>130</sup>, donnant lieu à des phénomènes de suppression d'ions de certains composés lorsque qu'ils sont en présence d'autres dans un mélange (« effets de matrice »). De plus, la présence de matériaux non volatils (sels, agents d'appariement d'ions, métabolites dans le cas d'échantillons biologiques) dans la solution limite la qualité spectrale (et peut d'autre part endommager les composants du détecteur)<sup>129</sup>. Cette variabilité de l'efficacité d'ionisation induit une absence de proportionnalité entre l'intensité du signal et la concentration, ce qui rend la spectrométrie de masse non quantifiable directement.

En outre, d'autres phénomènes perturbent la résolution des pics et l'exactitude en masse dans un FTICR<sup>126</sup>, notamment le phénomène d'espace de charge. Cet effet, dû au système de piégeage, se produit lorsqu'un trop grand nombre d'ions sont piégés en même temps, entraînant des phénomènes de répulsion qui altèrent leurs trajectoires<sup>131</sup>. La qualité de l'étalonnage dépend donc du contrôle du nombre d'ions avant l'injection dans la cellule ICR. Par ailleurs, l'étalonnage interne permet de s'affranchir partiellement de l'effet d'espace de charge d'espace et d'obtenir une précision de mesure inférieure à 1 ppm.

## II.4. Analyse de peptides par la spectrométrie de masse en tandem

### II.4.1. Théorie de la spectrométrie de masse en tandem (MS/MS)

La spectrométrie de masse tandem (MS/MS) consiste en l'activation d'ions précurseurs, suivi de leur dissociation en phase gazeuse et de l'analyse en masse des ions fragments résultants.

Il existe plusieurs types d'expériences MS/MS en fonction des catégories d'instruments : instrument hybride par couplage d'analyseurs dans l'espace (ex. triple quadripôle) et/ou analyseurs dits temporels. Dans le cas des analyseurs temporels, un seul analyseur en masse est nécessaire et les expériences MS/MS procèdent par séquences temporelles. L'appareil sélectionne d'abord les ions dits précurseurs avant de les fragmenter dans un second temps. Les pièges ioniques et les FTICR en sont des exemples. Pour des analyses dans l'espace, deux analyseurs en masse et une cellule de collision sont couplés pour réaliser des expériences MS/MS.

Plusieurs méthodes d'activation des ions existent ; elles peuvent être induites par un (ou des) photon(s) (UVPD (UltraViolet PhotoDissociation), IRMPD (InfraRed MultiPhoton Dissociation), BIRD (Blackbody Infrared Radiative Dissociation)), par l'intermédiaire d'électrons (ECD (Electron-Capture Dissociation), ETD (Electron-Transfer Dissociation), EDD (Electron-Detachment Dissociation) et EID (Electron-Induced Dissociation)) ou plus fréquemment *via* des collisions avec un gaz inerte (mode CID (Collision-Induced Dissociation))<sup>132</sup>. Dans un instrument FTICR, le mode le plus couramment utilisé est le SORI-



CID (Sustained Off-Resonance Irradiation Collision-Induced Dissociation)<sup>133</sup> qui permet d'activer les ions précurseurs à l'intérieur de la cellule ICR. Dans le cas, les expériences de type MS/MS ont été préférentiellement opérées en amont de la cellule ICR (instrument QhFTICR), selon le mode CID.

La justification du développement des différentes méthodes d'activation se base essentiellement sur la production d'informations structurales complémentaires selon le type de composés étudiés.

## II.4.2. Application de la MS/MS au séquençage de peptides

Notre étude se focalise particulièrement sur la sélectivité de polymérisation. Il est donc important d'avoir accès aux séquences des peptides formés. La MS/MS est un outil utilisé en protéomique pour le séquençage de biopolymères.

Pour le séquençage des peptides, la liaison qui se rompt le plus facilement selon un mode d'activation collisionnelle (mode CID) est la plus faible, c'est-à-dire la liaison amide<sup>134</sup>. Les ions produits sont donc essentiellement les ions fragments de série b et y selon la nomenclature de Roepstorff<sup>135</sup> (explicitée dans la Figure I-8). Les ions de type b peuvent ensuite se transformer en oxazolone<sup>136</sup>. Les ions de série b montrent souvent des ions satellites à 28 u de moins. Ces pics correspondent à la perte d'un monoxyde de carbone neutre et correspondent aux ions de série a<sup>134</sup>.

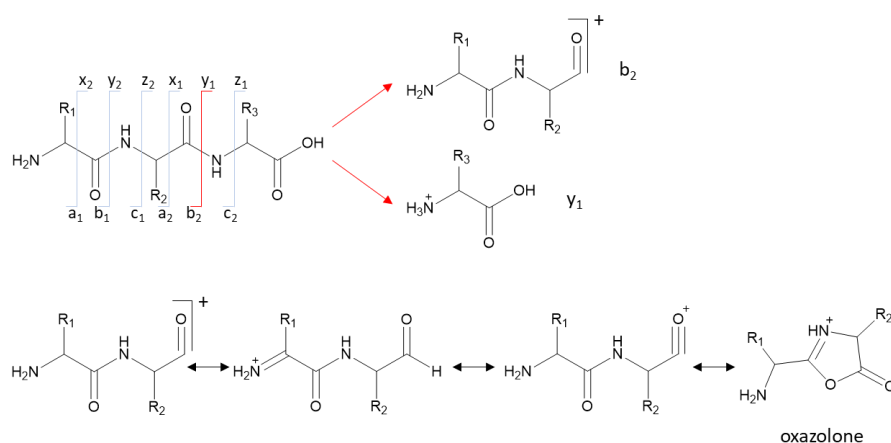


Figure I-8. Schéma de fragmentation d'un peptide<sup>135</sup>, illustré dans le cas d'un tripeptide. Les ions a, b, c sont ceux qui contiennent le groupe N-terminal, les ions x, y et z contenant le C-terminal.

De nombreux paramètres des expériences MS/MS contribuent à l'allure du spectre MS/MS obtenu. En effet, le temps de résidence des ions et la manière de fournir l'énergie ont une influence sur les voies de fragmentation. Ainsi, les systèmes de piégeage sont des méthodes de « chauffage doux », qui favorisent les fragmentations de faible énergie et les processus induits par la charge. Pour le mode CID, les abondances des ions produits varient en fonction de l'appareil utilisé.

## II.5. Analyses par spectrométrie de masse et origines de la Vie

La spectrométrie de masse à ultra haute résolution (HRMS) est un outil clé pour l'élucidation de compositions élémentaires d'échantillons complexes organiques. C'est une technique très utilisée en pétrologie, pour l'analyse de matière organique naturelle<sup>137</sup> et en protéomique. Elle est de plus en plus utilisée dans les études de chimie prébiotique, notamment dans les études de la matière organique contenue dans certaines météorites, les chondrites carbonées. Les mélanges extrêmement complexes qui résultent de l'extraction de ces météorites nécessitent l'utilisation d'appareils très sensibles et résolutifs. La météorite la plus connue est celle dite de Murchison, trouvée en Australie en 1961. Elle a été largement étudiée notamment par le groupe de Schmitt-Kopplin qui a utilisé la détection à ultra haute résolution avec une approche analytique non ciblée pour analyser la matière organique<sup>138-140</sup>. Leurs analyses ont permis l'identification de plus de 150 000 pics, menant à plus de 14 000 compositions moléculaires identifiées. La météorite de Paris a également été analysée par FTICR<sup>141</sup>. Les analyseurs Orbitrap peuvent être utilisés comme instruments pour les missions spatiales pour caractériser l'environnement des planètes *in situ*<sup>142</sup>.

L'usage de la HRMS n'est pas limité à l'analyse d'échantillons naturels, la technique pouvant être aussi appliquée aux synthèses prébiotiques en laboratoire. Par exemple, Wollrab et al ont reproduit les expériences de Miller et ont utilisé un instrument FTICR pour analyser le mélange résultant<sup>143</sup>. Depuis quelques années la HRMS est également utilisée par l'équipe de Forsythe pour étudier la polymérisation de protopeptides<sup>7,144</sup>. A partir de bibliothèques de données de depsipeptides, ils utilisent des approches basées sur les mesures de masses précises, notamment les traitements de Van Krevelen et Kendrick, pour visualiser la diversité

chimique des composés produits. Ces traitements ont été également utilisés pour l'étude de nos systèmes (voir Chapitre V).

Par ailleurs, les résultats de synthèse prébiotique sont souvent contestés parce que les rendements en composés cibles sont faibles et que l'utilisation de réactifs d'origine biologique peut causer de faux positifs. Il a été proposé que toute synthèse prébiotique ne soit considérée comme confirmée que si elle a été reproduite avec des réactifs marqués isotopiquement, tâche pour laquelle la spectrométrie de masse est évidemment la technologie de choix.

Indépendamment de l'instrument utilisé, la spectrométrie de masse permet également d'étudier la réactivité en phase gazeuse. Dans le contexte de la chimie prébiotique, Wincel et al<sup>145</sup> ont étudié la formation de liaisons peptidiques par réaction ion/molécule en phase gazeuse pour un système contenant de la méthionine et de l'acide glutamique.

De surcroît, plusieurs travaux en spectrométrie de masse ont étudié la genèse de l'homochiralité *via* l'analyse des clusters moléculaires. Ces derniers, formés en solution ou en phase gazeuse par des liaisons non covalentes, sont des structures intermédiaires entre les cristaux et les molécules isolées, formées. Le terme « nombre magique » est utilisé pour désigner un cluster d'une taille définie qui lui confère une stabilité inhabituelle<sup>146</sup>. La plupart des études liées aux clusters biomoléculaires se sont focalisées sur la sérine. Plusieurs études ont mis en évidence la capacité de la sérine à former des clusters de nombres magiques, avec de fortes préférences chirales<sup>146 147 148</sup>. La sérine peut aussi transmettre la chiralité vers d'autres acides aminés par substitution de certains acides aminés dans les octamères de sérine<sup>146 147 148 149</sup>.

Par ailleurs, Illos et al<sup>150</sup> ont démontré par MALDI-TOF-MS, à partir de composés énantiomères marqués au deutérium, la formation de peptides homochiraux comme diastéréoisomères dominants à partir de mélanges racémiques d'acides aminés amphiphiles en solution aqueuse, effet observé dès que les oligopeptides devenaient assez longs pour s'organiser en structures secondaires. Ces auteurs utilisaient la polymérisation en solution en présence de CDI (cf Chapitre I.I.2.1.1). Il s'agit d'un scénario différent de nos conditions ; il ne semble pas que la question de l'énantiosélectivité ait été étudiée dans le scénario de polymérisation sur surfaces minérales.

### III. Description des systèmes étudiés

#### III.1. Choix des acides aminés

##### III.1.1. Le système de référence : acide glutamique + leucine

Dans ce travail, plusieurs systèmes d'acides aminés ont été étudiés dans le cadre du scénario de polymérisation sur surfaces de silice. Nos premières études se sont portées sur un système binaire, le système leucine (Leu, L) + acide glutamique (Glu, E) ; les structures de ces deux acides aminés sont surlignées sur la Figure I-9.

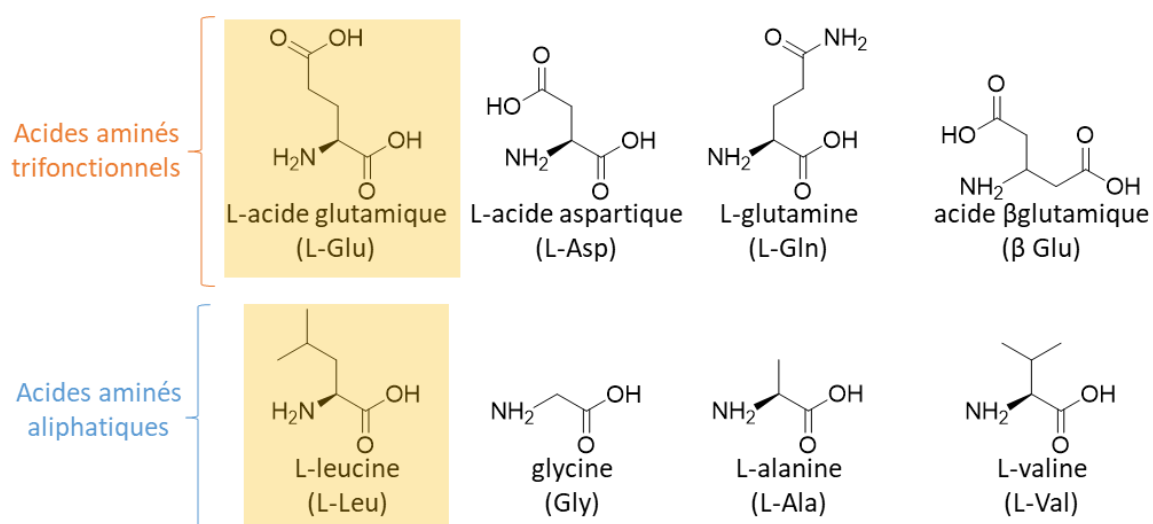


Figure I-9. Structures des acides aminés étudiés. Celles de l'acide glutamique et de la leucine sont surlignées.

Ce système est intéressant pour plusieurs raisons. Tout d'abord, Glu et Leu font partie des acides aminés primordiaux<sup>151</sup>, c'est-à-dire ceux dont la présence était la plus probable sur la Terre primordiale. En effet, ils sont formés dans les synthèses prébiotiques telles que l'expérience de Miller, et font aussi partie des premiers acides aminés à avoir été utilisés dans les protéines d'après les reconstitutions du protéome ancien<sup>152</sup>. De plus, ces deux acides aminés ont été identifiés parmi les acides aminés présents dans la météorite de Murchison<sup>153</sup>. Ils pourraient donc avoir été « livrés » sur Terre par le bombardement météoritique.

D'autre part, les polypeptides contenant Glu et Leu possèdent des propriétés structurales intéressantes. L'enchaînement des acides aminés constitue la structure primaire d'une protéine. Les protéines sont présentes dans les cellules sous forme repliée, chacune s'auto-organisant différemment en fonction de sa séquence, ce qui lui confère des propriétés propres. On peut distinguer des structures secondaires et tertiaires. Au niveau de la structure secondaire, les chaînes s'organisent surtout en hélices alpha ou en feuillets bêta. Bertrand<sup>154</sup> énonce comme règle générale que l'alternance des acides aminés hydrophiles (hi) tels que Glu (chaîne latérale acide), et hydrophobes (ho) tels que Leu (chaîne latérale aliphatique non ionisable), donc une séquence (-hi-ho-hi-ho-), induit une structure avec des feuillets bêta tandis qu'une polymérisation en blocs (-hi-hi-ho-ho-) entraîne une conformation en hélice alpha. Or les structures en feuillets bêta sont thermostables et résistantes à l'hydrolyse, contrairement aux hélices alpha, et pourraient donc avoir un rôle important dans l'origine de la vie<sup>19,155</sup>. On peut donc supposer qu'il y a eu une sélection pour les systèmes d'acides aminés capables de former des feuillets bêta.

Ce sont des considérations de ce type qui ont orienté les premières études de systèmes mixtes réalisées au LRS vers un nombre restreint de couples, incluant entre autres (Val + Asp) et (Leu + Glu)<sup>89</sup>. Comme nous l'avons signalé au Chapitre I.1.3.2, le système (Leu + Glu) s'est révélé particulièrement prometteur pour la formation d'oligopeptides linéaires. Il possède aussi un avantage analytique pratique que nous discuterons plus tard, la facilité d'extraction des peptides formés lors de l'étape d'activation thermique.

### III.1.2. Extension à d'autres systèmes d'acides aminés

Pour essayer de comprendre le rôle de chaque acide aminé dans la polymérisation, nous avons également étudié d'autres systèmes binaires comprenant un acide aminé trifonctionnel : leucine + acide aspartique, leucine + glutamine et leucine + acide  $\beta$ -glutamique. La chaîne latérale de l'acide aspartique (Asp, D) contient un groupe  $-CH_2-$  en moins, ce qui affecte la possibilité de cyclisation interne. Pour la glutamine (Gln, Q), le groupe acide carboxylique de la chaîne latérale est remplacé par un groupe amide. Enfin, dans le cas de l'acide  $\beta$ -glutamique, les groupements  $-COOH$  sont portés par les carbones  $\beta$  et  $\gamma$ , et non  $\alpha$  et  $\gamma$  comme dans l'acide glutamique précédemment présenté.

Par ailleurs, nous avons choisi d'étudier un système plus complexe, susceptible de fournir des informations sur la sélectivité en conditions plus réalistes prébiotiquement. Nous avons sélectionné le système GAVD (glycine (Gly, G), alanine (Ala, A), valine (Val, V), acide aspartique (Asp, D)) car ces quatre acides aminés sont considérés comme étant les acides aminés protéinogènes les plus abondants dans la météorite de Murchison et dans les expériences de synthèse prébiotique par décharge électrique<sup>156</sup>. De façon intéressante, une approche « top-down » à partir des protéomes des organismes actuels aboutit indépendamment à proposer ces quatre acides aminés comme les plus primordiaux<sup>157</sup>. Ces quatre acides aminés auraient permis de former des « protéines GAVD », utilisant un « alphabet » d'AA plus limité que celui des protéines actuelles<sup>158,159</sup>. Les structures de ces acides aminés sont montrées Figure I-9.

## III.2. Spécificités de l'acide glutamique

L'acide glutamique (ou l'ion glutamate en fonction de sa spéciation) joue un rôle clé dans le métabolisme de nombreux organismes, notamment dans l'assimilation de l'azote et la biosynthèse d'acides aminés. Il est aussi impliqué dans la production de métabolites secondaires comme les antibiotiques. Du point de vue de l'évolution, la causalité entre la présence importante de glutamate dans l'organisme et son implication dans de nombreux processus n'est pas clairement établie : la grande quantité de Glu dans l'organisme peut être une conséquence de son implication dans de nombreux processus biologiques, ou alors les processus se sont adaptés pour utiliser la grande quantité de Glu présente<sup>160</sup>.

### III.2.1. Cyclisation de Glu en pyroGlu

Une particularité de l'acide glutamique est sa capacité à subir une cyclisation interne pour former l'acide pyroglutamique (aussi appelé « Pyrrolidone Carboxylic Acid » (PCA) et noté pyroGlu ou pE), qui a une réactivité intéressante et complexe. La formation de pyroGlu est illustrée Figure I-10.

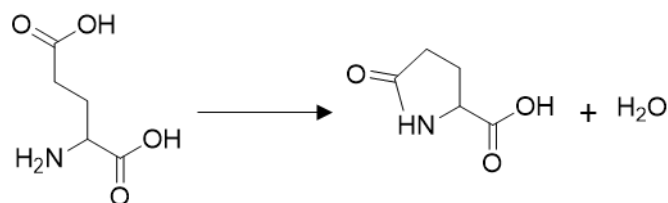


Figure I-10. Formation de l'acide pyroglutamique à partir de l'acide glutamique

### III.2.1.1. Réaction *in vivo*, dans les chaînes protéiques

#### III.2.1.1.1. pyroGlu en N-terminal dans les chaînes protéiques

De nombreux composés biologiques comportent un résidu acide pyroglutamique (pyroGlu ou pE) à leur extrémité N-terminale. La thyrotropin-releasing hormone (TRH), tripeptide L-pyroglutamyl-L-histidyl-L-prolinamide qui se trouve dans différentes cellules humaines, principalement du système nerveux central, en est un exemple. L'acide pyroglutamique peut être formé par cyclisation d'un résidu acide glutamique ou glutamine en N-terminal. La cyclisation de Gln en pyroGlu s'accompagne d'une élimination d'ammoniaque<sup>161,162</sup>, celle de Glu en pyroGlu d'une élimination d'eau. La glutamine peut être désamidée de façon enzymatique, une réaction catalysée dans ce cas par la glutamine cyclotransférase<sup>163</sup>, ou de façon non enzymatique<sup>164</sup>. La conversion de Glu en pGlu a lieu moins souvent et est en général plus lente que celle de Gln en pGlu<sup>165</sup>. Ainsi, Twardzik<sup>166</sup> a identifié le précurseur de pGlu en N-terminal des protéines de plasmocytome comme étant la glutamine.

Ainsi, la modification post-translationnelle la plus courante des anticorps monoclonaux est la conversion de la glutamine (Gln, Q) située en N-terminal des chaînes lourdes ou légères en acide pyroglutamique<sup>165</sup>. Ces chaînes, notamment celles des anticorps de l'immunoglobuline gamma, peuvent également contenir un Glu en N-terminal, qui se convertit aussi en pyroGlu<sup>167</sup>. De même, Chelius<sup>168</sup> montre la formation non enzymatique de pyroGlu à partir de Glu en N-terminal d'anticorps recombinants, après plusieurs semaines d'incubation dans plusieurs tampons à 37 °C et 45 °C. La formation de pyroGlu dans les anticorps thérapeutiques n'est pas un problème pharmacologique car pyroGlu est formé à partir des Glu et Gln N-terminaux *in vivo*. Une fois formé, le pyroGlu est très stable et a besoin de conditions dures (solutions alcalines ou fortement acides) pour s'hydrolyser en acide glutamique<sup>169,170</sup>.

## III.2.1.1.2. PyroGlu en résidu interne

Un résidu interne acide glutamique au sein d'une chaîne peptidique peut également se cycliser, soit sur l'amine du résidu précédent en formant un cycle à 5 atomes, soit sur l'amine du résidu suivant, donnant lieu à un cycle à 6 atomes (glutarimide), comme illustré sur la Figure I-11.

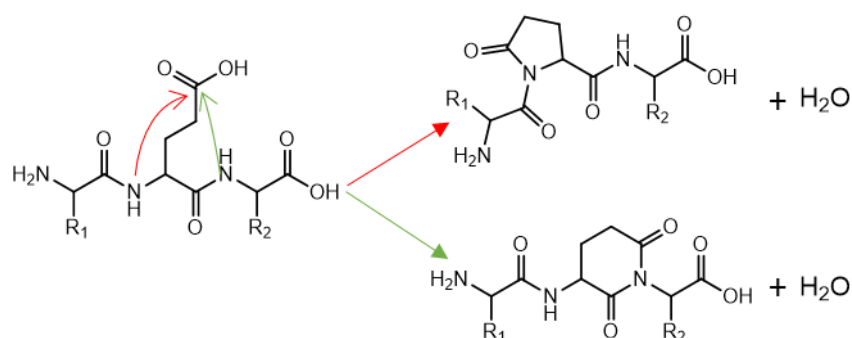


Figure I-11. Cyclisation du résidu Glu au sein d'une chaîne peptidique.

Clayton et Kenner ont montré qu'en présence de chlorure de thionyle et d'une amine tertiaire, le cycle à 6 est formé préférentiellement<sup>171</sup>. Il existe néanmoins des exemples de « pseudo acide pyroglutamique » au sein d'une chaîne peptidique. Ainsi, Howard<sup>172</sup> propose la présence d'un résidu acide pyroglutamique interne dans la protéine  $\alpha_2$ -Macroglobulin. Le pyroGlu peut alors subir une hydrolyse et entraîner la rupture de la liaison peptidique et ainsi former un nouveau résidu avec un pyrGlu en N-terminal. Khan<sup>173</sup> a repris cette idée et a synthétisé des polypeptides modèles ressemblant à la  $\alpha_2$ -Macroglobulin. Il suggère que certains composés ayant un pyroGlu en N-terminal peuvent provenir d'une hydrolyse régiosélective d'une fragmentation de chaîne d'un précurseur polypeptidique contenant un pyroGlu interne<sup>173</sup>. Les peptides modèles contenant un résidu pyroGlu à l'intérieur de la chaîne et le peptide modèle complètement cyclique sont interconvertibles dans des conditions physiologiques. Les résultats cinétiques suggèrent que la fragmentation de la chaîne protique s'opère par hydrolyse de la liaison peptidique du cycle lactame<sup>174</sup>.



### III.2.1.2. Formation de pyroGlu dans les études de chimie prébiotique

Plusieurs équipes travaillant dans le domaine de la chimie prébiotique ont étudié l'acide glutamique. Harada et Fox<sup>175 52</sup> ont étudié la polymérisation thermique d'acides aminés massiques (sans intervention d'une surface minérale). Ils ont remarqué que Glu avait un rôle particulier dans la copolymérisation, ce qui expliquerait sa forte proportion dans les protéines, même actuellement. Melius<sup>176</sup> a montré que les polypeptides formés de cette façon avaient un résidu pyroGlu en position N-terminale. Brack postule que si le Glu est chauffé avec d'autres acides aminés en l'absence de surface, l'acide glutamique a tendance à être localisé en position N-terminale car son groupe amine est bloqué par la cyclisation<sup>19</sup>.

Plusieurs équipes de recherches ont souligné le rôle de pyroGlu comme initiateur de la polymérisation, notamment Nakashima et al.<sup>55</sup> discutée au Chapitre I.I.2.3.

Il a également été montré que sur surface de silice, la cyclisation de Glu en pyroGlu a lieu à des températures inférieures à celles nécessaires à la condensation de deux acides aminés<sup>91</sup>

### III.2.2. La question de la régiosélectivité de la condensation : $\alpha$ - et $\gamma$ -peptides

En raison de sa chaîne latérale acide, l'acide glutamique peut former des  $\gamma$ -peptides (voir Figure I-12). C'est aussi le cas de la glutamine. Il faut noter que l'acide aspartique peut quant à lui former des  $\beta$ -peptides.

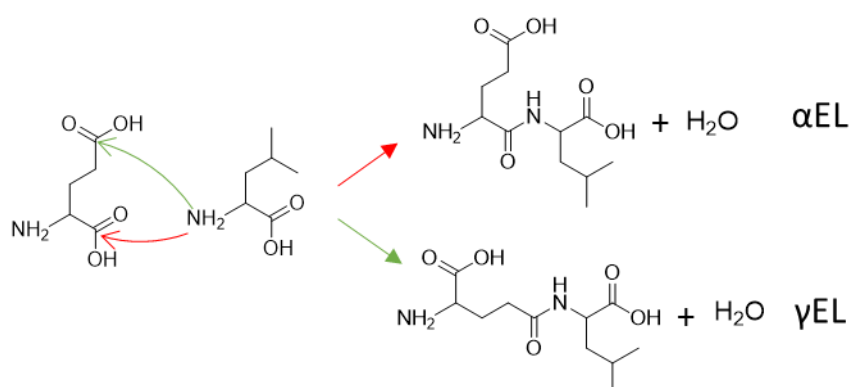


Figure I-12. Exemple de formation d'un  $\gamma$ -dipeptide : cas de EL.

Certains  $\gamma$ -glutamyl peptides, peptides contenant un acide glutamique  $\gamma$  à l'extrémité N-terminale de la molécule, existent naturellement, tels que le glutathion (L- $\gamma$ -glutamyl-L-cystéinyglycine ; GSH). Ce tripeptide, comportant une liaison peptidique entre le COOH en position  $\gamma$  de l'acide glutamique et le NH<sub>2</sub> de la cystéine, participe à la synthèse des protéines et des acides nucléiques et joue un rôle important dans le métabolisme des glucides et le transport des acides aminés<sup>177</sup>. Le type particulier de liaison le rend résistant à la majorité des peptidases cellulaires. De nombreux  $\gamma$ -glutamyl peptides sont des produits ou coproduits du cycle du glutathion dans les organismes<sup>178</sup>. Plusieurs dérivés de  $\gamma$ -glutamyl peptides ont été également isolés à partir de plantes<sup>179,180</sup>. Les  $\gamma$ -glutamyl di- et tripeptides sont également largement présents dans les aliments comme les légumineuses (notamment  $\gamma$ Glu-Glu,  $\gamma$ Glu-Val,  $\gamma$ Glu-Asp,  $\gamma$ Glu-Leu), les alliacées et dans les aliments fermentés comme le fromage, la sauce soja et le poisson fermenté<sup>178</sup>. Ils impactent le goût des aliments et sont notamment les molécules clés qui provoquent la saveur dite « kokumi » dans la nourriture fermentée<sup>181,182,183</sup>. Il a également été montré que les  $\gamma$ -glutamyl peptides présentent des effets antioxydants, anticancers et détoxifiants<sup>184,185</sup>. Ces peptides sont relativement stables dans le corps car ils ne sont hydrolysés que par une seule enzyme : la  $\gamma$ -glutamyl cyclotransférase<sup>186</sup>. Une autre particularité qui leur est attribuée est leur capacité à activer le récepteur de calcium extracellulaire (CaSR)<sup>187,188</sup>. L'activation de CaSR par  $\gamma$ EC (qui est le coproduit majoritaire du cycle du glutathion) et  $\gamma$ EV peut aider à réduire l'inflammation dans des conditions d'inflammations chroniques comme dans les maladies intestinales<sup>189,190</sup>. Par ailleurs, Kasai<sup>191</sup> a montré une différence de réactivité entre les  $\alpha$ - et  $\gamma$ - glutamyl peptides : le chauffage du dipeptide  $\alpha$ Glu-Leu mène à la formation du dipeptide pyroGlu-Leu tandis que celui de  $\gamma$ Glu-Leu aboutit uniquement à un monomère pyroGlu.

L'émergence des protéines nécessite des oligopeptides relativement longs, néanmoins certains dipeptides, peuvent déjà présenter des fonctions importantes. Les études bibliographiques que nous venons de citer indiquent que les dipeptides  $\gamma$ Glu-Xxx\* font partie de ceux-ci, ce qui constitue une raison supplémentaire pour s'intéresser à la possibilité de former ces systèmes.

---

\* Xxx correspond à un acide aminé quelconque



## Chapitre II. Matériels et méthodes

## I. Préparation des échantillons

La préparation des échantillons se compose de trois étapes : le dépôt sur surface minérale, l'activation thermique et l'étape de désorption avant leur analyse par HRMS.

### I.1. Dépôt sur surface minérale

La méthode utilisée pour l'adsorption des acides aminés est l'imprégnation à humidité naissante (IHN) (Incipient wetness impregnation, IWI). Cette méthode consiste à solubiliser les réactifs dans un volume minimal d'eau afin d'imprégner le support sans induire la formation d'une phase liquide séparée. En principe, la solution occupe le volume poreux (interparticulaire et, le cas échéant, intraparticulaire), et la valeur du volume d'imprégnation est déterminée préalablement en ajoutant à la phase solide de l'eau distillée jusqu'à observation d'une phase liquide séparée. Dans le cas des silices amorphes, les volumes d'imprégnation observés peuvent être très importants : les nanoparticules de SiO<sub>2</sub> peuvent causer la formation d'une phase de type sol, contenant une quantité d'eau jusqu'à 10 fois supérieure à celle de silice solide.

Dans notre expérience de base, les acides aminés ( $2 \cdot 10^{-4}$  mol) sont dissous dans 10 mL d'eau puis la solution obtenue est déposée sur 1 g de silice<sup>†</sup> afin d'obtenir une pâte homogène qui est ensuite séchée sous flux d'azote pendant une nuit jusqu'à l'obtention d'une poudre. Après avoir été broyé, l'échantillon est stocké dans un flacon en plastique et placé dans un dessiccateur.

A titre d'exemple, l'échantillon de notre système modèle (L+E)/SiO<sub>2</sub> a été préparé à partir de 13,1 mg de Leu et 14,7 mg de Glu, ce qui correspond à un rapport molaire 1 :1 (équimolaire) des deux acides aminés.

---

<sup>†</sup> La même procédure a été appliquée à deux autres solides : une montmorillonite et une alumine.

### I.1.1. Une estimation de la « monocouche »

Si on assimile les acides aminés à des rectangles de longueur et largeur données, il est possible de calculer l'aire occupée par les acides aminés sur une surface. Le Tableau II-1 donne ainsi les aires de l'acide glutamique et de la leucine. Cette estimation donne donc une idée générale de la quantité de molécules nécessaire pour recouvrir une surface, pour autant qu'ils soient adsorbés « à plat », avec une orientation maximisant leur projection sur la surface et disposés selon un réseau rectangulaire. On peut contester ces présuppositions ; néanmoins, elles fournissent au moins un ordre de grandeur pour ce qu'on appelle généralement une « monocouche physique ».

Tableau II-1. Estimations des dimensions des acides aminés Glu et Leu<sup>192</sup>.

	Longueur (Å)	Largeur (Å)	Aire (Å <sup>2</sup> = 10 <sup>-20</sup> m <sup>2</sup> )
Glu	8,0	3,5	28
Leu	8,5	5,0	42,5

Pour 1 g d'une silice de surface spécifique 380 m<sup>2</sup>/g, le nombre d'acides aminés correspondant à une monocouche physique est de  $\frac{380}{28 \cdot 10^{-20}} = 1,36 \cdot 10^{21}$  acides aminés pour

la leucine, ce qui correspond à un nombre de moles de  $\frac{1,36 \cdot 10^{21}}{6,02 \cdot 10^{23}} = 2,25 \cdot 10^{-3}$  mol/g.

Pour l'acide glutamique, on obtient  $\frac{380}{42,5 \cdot 10^{-20}} = 8,94 \cdot 10^{20}$  acides aminés soit  $1,48 \cdot 10^{-3}$

mol/g. Pour un échantillon contenant les deux acides aminés en proportions 1 :1, le calcul donne  $\frac{380}{(42,5+28) \cdot 10^{-20}} = 5,39 \cdot 10^{20}$  de chacun des acides aminés, soit  $8,95 \cdot 10^{-4}$  mol/g de chacun.

Comme nous le verrons dans les discussions ultérieures, cette notion de « monocouche physique » constitue un repère pour l'intuition, mais il n'y a pas de raison de penser qu'elle joue un rôle déterminant dans les phénomènes d'adsorption d'acides aminés, qui impliquent sans doute une adsorption sur site.

### I.1.2. Surfaces utilisées

La quasi-totalité des échantillons a été préparée sur surface de silice. La silice utilisée est une silice amorphe non poreuse de surface spécifique 380 m<sup>2</sup>/g commercialisée sous le nom d'Aerosil380. Une explication détaillée de la surface de silice est donnée Chapitre I.

Deux autres surfaces minérales ont été étudiées comme support de polymérisation, la montmorillonite et l'alumine. La montmorillonite est l'argile la plus commune du groupe des smectite sur la surface de la Terre et était présente sur la Terre Hadéenne<sup>59,60</sup>. Sa structure, montrée Figure II-1, est de type 2 :1 et consiste en un feuillet octaédrique intercalé entre deux feuillets tétraédriques.

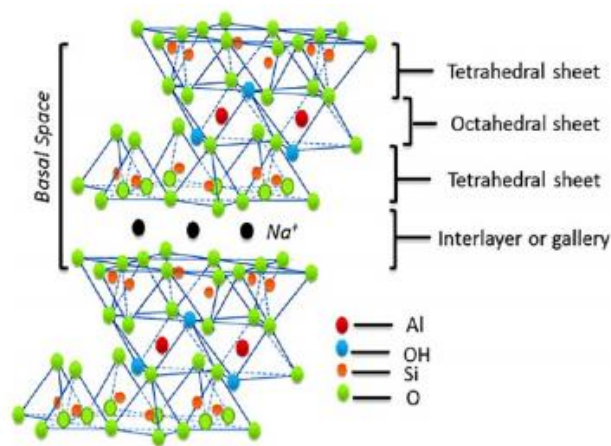


Figure II-1. Schéma d'une structure de montmorillonite de type 2 :1 <sup>193</sup>

Nous avons choisi d'utiliser une montmorillonite qui a déjà été caractérisée<sup>194</sup> : c'est une montmorillonite(M) Volclay SPV-200 (American Colloids), de rapport 2 :1 (Si/Al), avec une capacité d'échange de cations (CEC) de 76,4 cmol/kg (plus communément notée 76,4meq/100g) et une surface spécifique de 78 m<sup>2</sup>/g. La montmorillonite a été rendue homoionique en Na<sup>+</sup> en suivant le protocole de Banin et al<sup>195</sup>.

Par ailleurs, l'alumine a également déjà été utilisée comme support de polymérisation d'acides aminés. L'oxyde d'aluminium (alumine, Al<sub>2</sub>O<sub>3</sub>) existe sous plusieurs formes polymorphes. Nous avons choisi d'utiliser l'alumine  $\gamma$  pour plusieurs raisons. Malgré que l'alumine  $\alpha$  soit prébiotique et thermodynamiquement stable<sup>59,60</sup>, elle a une surface spécifique quasi nulle<sup>196</sup>, ce qui n'en fait pas une surface de choix pour la caractérisation des espèces supportées. Les acides aminés adsorbés seraient présents en trop faible quantité, ce

qui ne permettrait pas leur caractérisation. De plus, bien que l'alumine gamma soit métastable<sup>197</sup>, elle aurait pu être présente à l'ère prébiotique. Par ailleurs, les phases inorganiques métastables sont intéressantes pour activer des biomolécules. L'alumine  $\gamma$  est un matériau semi-cristallin, de surface spécifique élevée, largement utilisé en catalyse hétérogène. Elle a la particularité de ne pas adsorber l'eau de la même manière que les autres matériaux similaires. En effet, elle décompose les molécules d'eau et change partiellement sa composition et sa structure. L'alumine  $\gamma$  a une structure spinelle lacunaire : les atomes d'oxygènes sont disposés selon un empilement cubique face centrées et les atomes d'aluminium occupent une partie des sites octaédriques et tétraédriques. De nombreux travaux ont été consacrés à la nature de ses sites de surface, qui incluent au moins cinq types différents de groupes aluminols Al-OH (terminaux, pontants, coiffants, sur Al 4-coordonnés ou 6-coordonnés), et permettent aussi l'adsorption par formation de liaisons de coordination avec les ions  $\text{Al}^{3+}$ . La Figure II-2 présente plusieurs types de OH chimiquement différents suivant le modèle de Busca<sup>198</sup>. Certains détails sont contestés par d'autres auteurs, mais tous s'accordent à reconnaître l'existence d'un nombre significatif de groupes -OH différents.

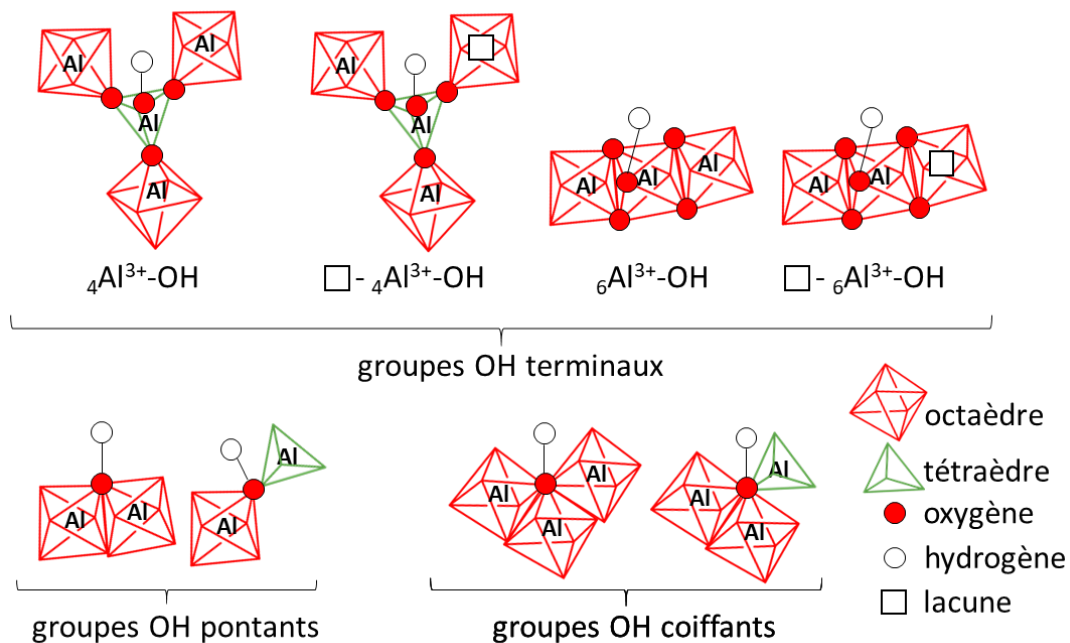


Figure II-2. Plusieurs types de groupements -OH de surface sur une alumine  $\gamma$ . La coordinence de l'atome d'aluminium (ou des atomes) qui les porte(nt) est notée par l'indice 4 ou 6.

Communication personnelle de J-F Lambert



L'alumine  $\gamma$  utilisée a été synthétisée à partir de boehmite Pural Scc de SASOL. La boehmite a été calcinée à 500 °C (rampe de 10 °C/min puis plateau à 500 °C pendant 12h) dans un four à moufle.

Les surfaces spécifiques de ces deux matériaux ont été mesurés par BET. Les surfaces obtenues sont 33 m<sup>2</sup>/g pour la montmorillonite et 203 m<sup>2</sup>/g pour l'alumine. Pour la montmorillonite la valeur attendue était de 70 m<sup>2</sup>/g mais le degré de délamination des feuillets peut changer selon les conditions de préparation.

Le calcul de la monocouche expliqué plus haut pour la silice peut être fait pour ces deux surfaces également. Dans le cas de la montmorillonite, le taux d'acides aminés déposé dépasse la monocouche physique.

Une fois déposés sur surface, les échantillons sont activés thermiquement dans le but de former des liaisons peptidiques (cf § I.3).

## I.2. Description de l'ensemble des échantillons testés

Dans l'échantillon de base de rapport molaire 1 :1,  $1.10^{-4}$  mol de chaque acide aminé ont été déposés, c'est-à-dire  $2.10^{-4}$  mol/g au total. Tous les échantillons préparés sur silice ont donc une charge en acides aminés nettement inférieure à la monocouche physique, y compris ceux avec des rapports molaires 5 :1 et 10 :1, qui contiennent  $6.10^{-4}$  mol/g (Chapitre IV). Si le rapport molaire n'est pas indiqué dans le nom de l'échantillon, il est de 1 :1 avec 1 correspondant à  $1.10^{-4}$  mol/g. Quand l'énantiométrie n'est pas précisée, c'est l'acide aminé L qui est utilisé.

Tous les échantillons testés sont répertoriés dans le Tableau II-2.

Tableau II-2. Ensemble des échantillons d'acides aminés supportés sur surfaces minérales discutés dans le manuscrit.

Nom de l'échantillon	Composition	Rapport molaire	T activation
(L+E)/SiO <sub>2</sub>	Leu + Glu	1 :1	160 °C + 100 °C, 115 °C, 130 °C et 145 °C
(L+Eabloqué)/SiO <sub>2</sub>	Leu + αGlu bloqué (COOH bloqué tertibutyle)	1 :1	160 °C
(L+Eγbloqué)/SiO <sub>2</sub>	Leu + γGlu bloqué (COOH bloqué tertibutyle)	1 :1	160 °C
(L <sub>L</sub> +(E <sub>L</sub> *+E <sub>D</sub> ))/SiO <sub>2</sub>	L-Leu + L-Glu <sup>13</sup> C +D-Glu	1 :0,5 :0,5	160 °C
(L <sub>L</sub> + (E <sub>L</sub> +E <sub>D</sub> *))/SiO <sub>2</sub>	L-Leu + L-Glu + D-Glu <sup>13</sup> C	1 :0,5 :0,5	160 °C
(E <sub>L</sub> +(L <sub>L</sub> *+L <sub>D</sub> ))/SiO <sub>2</sub>	L-Glu + L-Leu <sup>13</sup> C <sup>15</sup> N +D-Leu	1 :0,5 :0,5	160 °C
(E <sub>L</sub> +(L <sub>L</sub> +L <sub>D</sub> *))/SiO <sub>2</sub>	L-Glu + L-Leu + D-Leu <sup>13</sup> C	1 :0,5 :0,5	160 °C
(LE+L)/SiO <sub>2</sub>	Dipeptide Leu-Glu + Leu	1 :1	130 °C et 160 °C
(LE+E)/SiO <sub>2</sub>	Dipeptide Leu-Glu + Glu	1 :1	130 °C et 160 °C
(EE+L)/SiO <sub>2</sub>	Dipeptide Glu-Glu + Leu	1 :1	130 °C et 160 °C
(AA+E)/SiO <sub>2</sub>	Dipeptide Ala-Ala + Leu	1 :1	130 °C et 160 °C
(cycloEE+L)/SiO <sub>2</sub>	Dipeptide cycloGlu-Glu + Leu	1 :1	130 °C et 160 °C
(cycloAA+E)/SiO <sub>2</sub>	Dipeptide cycloAla-Ala + Glu	1 :1	130 °C et 160 °C
(pE+L)/SiO <sub>2</sub>	pyroGlu + Leu	1 :1	160 °C
(ELL+D)/SiO <sub>2</sub>	Tripeptide Glu-Leu-Leu + Asp	1 :1	160 °C
(L+E) <sub>(0,25 :0,25)</sub> /SiO <sub>2</sub>	Leu + Glu	0,25 :0,25	160 °C
(L+E) <sub>(3 :3)</sub> /SiO <sub>2</sub>	Leu + Glu	3 :3	160 °C
(L+E) <sub>(5 :1)</sub> /SiO <sub>2</sub>	Leu + Glu	5 :1	160 °C
(L+E) <sub>(2 :1)</sub> /SiO <sub>2</sub>	Leu + Glu	2 :1	160 °C
(L+E) <sub>(1 :2)</sub> /SiO <sub>2</sub>	Leu + Glu	1 :2	160 °C
(L+E) <sub>(1 :5)</sub> /SiO <sub>2</sub>	Leu + Glu	1 :5	160 °C
(L+E) <sub>(0,5 :5,5)</sub> /SiO <sub>2</sub>	Leu + Glu	0,54 :5,46	160 °C
(L+E)/SiO <sub>2</sub> 160+70	Leu + Glu	1 :1	160 °C puis 70 °C en conditions humides
(L+E)/SiO <sub>2</sub> 160+100	Leu + Glu	1 :1	160 °C puis 100 °C en conditions humides
(L+E)/SiO <sub>2</sub> dépôt séparé	Leu + Glu	1 :1	160 °C
(L+D)/SiO <sub>2</sub>	Leu + Asp	1 :1	160 °C
(L+Q)/SiO <sub>2</sub>	Leu + Gln	1 :1	160 °C
(L+βE)/SiO <sub>2</sub>	Leu + βGlu	1 :1	160 °C
(L+(E*+βE))/SiO <sub>2</sub>	Leu + Glu <sup>13</sup> C +βGlu	1 :0,5 :0,5	160 °C
(L+(E+D))/SiO <sub>2</sub>	Leu + Glu + Asp	1 :0,5 :0,5	160 °C
(L+(D+βE))/SiO <sub>2</sub>	Leu + Asp + βGlu	1 :0,5 :0,5	160 °C
(L+(E*+Q))/SiO <sub>2</sub>	Leu + Glu <sup>13</sup> C + Gln	1 :0,5 :0,5	160 °C
(GAVD)/SiO <sub>2</sub>	Gly + Ala + Val + Asp	1 :1 :1 :1	130 °C et 160 °C
(GAVE)/SiO <sub>2</sub>	Gly + Ala + Val + Glu	1 :1 :1 :1	130 °C et 160 °C
(G <sub>L</sub> +A <sub>L</sub> +V <sub>L</sub> +E <sub>L</sub> +E <sub>D</sub> *)/SiO <sub>2</sub>	Gly + Ala + Val + L-Glu + D-Glu <sup>13</sup> C	0,5 :0,5 :0,5 :0,25 :0,25	160 °C
(G <sub>L</sub> +A <sub>L</sub> +V <sub>L</sub> +E <sub>L</sub> *+E <sub>D</sub> )/SiO <sub>2</sub>	Gly + Ala + Val + L-Glu <sup>13</sup> C + D-Glu	0,5 :0,5 :0,5 :0,25 :0,25	160 °C
L+E-CDI	Leu + Glu	1 :1	Préparé avec CDI
(L+E)/Mt	Leu + Glu	1 :1	160 °C
(L+E)/Al <sub>2</sub> O <sub>3</sub>	Leu + Glu	1 :1	160 °C
(ATP+GTP)/SiO <sub>2</sub>	Adénosine triphosphate + guanosine triphosphate	1 :1	160 °C
(ATP+UTP)/SiO <sub>2</sub>	Adénosine triphosphate + uridine triphosphate	1 :1	160 °C et 200 °C

Outre les expériences menées sur surface, nous avons effectué une réaction de polymérisation en solution aqueuse, en utilisant un activateur sacrificiel, le N,N-carbonyldiimidazole (CDI). Le protocole utilisé est le suivant<sup>199</sup> : 1,5 mM L-Leu, 1,5 mM L-Glu et 10 mM KCl sont dissous dans 10 mL d'eau. 6 mM de carbonyldiimidazole sont ajoutés et la solution est laissée à température ambiante pendant 24 h.

### I.3. Protocole d'activation thermique

Pour chaque échantillon AAs/SiO<sub>2</sub>, environ 100 mg sont activés thermiquement dans un tube en U traversé par un flux d'argon. Le tube est placé dans un four cylindrique contrôlé par un programmateur de température. Dans l'échantillon de base (L+E)/SiO<sub>2</sub>, la température est augmentée de l'ambiante jusqu'à 160 °C avec une rampe de 1 °C/min suivie d'un palier de 30 minutes à la température cible.

L'échantillon est ensuite placé dans un tube eppendorf puis peut être caractérisé par des techniques in situ ou bien subir une étape d'extraction dans le but d'être analysé ex-situ. En effet, différentes méthodes d'analyse ont été effectuées selon l'état physique de l'échantillon. Une analyse directe des produits présents à la surface de la silice peut être opérée par des méthodes de caractérisation de matériaux comme la spectroscopie infrarouge, la diffraction des rayons X et l'analyse thermogravimétrique. Pour des analyses plus fines et sélectives des produits de polymérisation, la solution obtenue après l'étape d'extraction est analysée par spectrométrie de masse.

### I.4. Etape d'extraction en solution

Une étape d'extraction est nécessaire afin de désorber les espèces produites par activation thermique en solution aqueuse avant leur caractérisation par spectrométrie de masse. 10 mg de solide sont suspendus dans 1 mL d'eau MilliQ et le mélange est agité avec un vortex. La phase solide est séparée du surnageant par centrifugation (15000 tours/min pendant 10 minutes). 900 µL de solution sont récupérés.

L'efficacité de cette étape est vérifiée par spectroscopie infrarouge, mais le rendement exact d'extraction est difficile à quantifier. Trois étapes d'extraction successives ont été testées et il a été montré que les produits de condensation étaient majoritairement présents dans la

première solution de désorption et non dans les deux suivantes. De plus il ne semble pas y avoir de sélectivité de désorption : les espèces présentes dans les trois extraits successifs sont les mêmes. Seule la première solution de désorption est donc analysée pour l'ensemble des échantillons.

## II. Techniques de caractérisation *in-situ* des systèmes AAs/SiO<sub>2</sub>

### II.1. Analyse thermogravimétrique

L'analyse thermogravimétrique (ATG) est réalisée sur un TA Instruments Waters avec un détecteur SDT Q600. Les échantillons sont chauffés avec un gradient de 1 °C/min jusqu'à 200 °C sous flux d'air avec un débit de 100 mL/min. La mesure de la variation de masse en fonction de la température permet de mettre en évidence la formation de la liaison peptidique, concomitante à la libération d'eau en phase gazeuse.

En outre, la thermogravimétrie a également été utilisée pour déterminer s'il subsiste une quantité de matière organique résiduelle après l'étape d'extraction ; dans ce cas sa décomposition thermique induirait une perte de masse visible. A cet effet, les échantillons ont été chauffés avec un gradient de 5 °C/min jusqu'à 800 °C sous flux d'air avec un débit de 100 mL/min.

### II.2. Diffraction des rayons X

L'analyse par diffraction des rayons permet de vérifier que les acides aminés sont adsorbés à la surface de la silice (Chapitre I) et non précipités indépendamment de celle-ci. Les diffractogrammes ont été enregistrés sur un diffractomètre Bruker D8 utilisant la raie K $\alpha$  du Cu ( $\lambda=1.5404$  Å). Le détecteur est monté sur un goniomètre qui tourne autour de l'échantillon et collecte les rayons X diffusés. L'échantillon sous forme de poudre est placé dans un support en plastique. En général, nous avons enregistré les diffractogrammes sur une gamme de  $2\theta$  allant de 10 à 45° avec un nombre de pas de 3810 et un temps d'accumulation de 1 sec par point.

### II.3. Spectroscopie infrarouge

L'appareil de spectroscopie infrarouge (IR) utilisé dans le présent travail est un spectromètre Infrarouge VERTEX 80 à transformée de Fourier comprenant un détecteur de type MCT ayant une résolution de  $4\text{ cm}^{-1}$ . En fonction du support utilisé et de l'état de l'échantillon, les analyses IR sont faites en transmission (dans le cas des échantillons supportés sur silice et alumine notamment) ou en réflectance totale atténuée (ATR, Attenuated Total Reflectance). Quand cela est possible, le mode ATR est préféré car il donne la possibilité d'analyser des échantillons solides et liquides de manière rapide et répétable sans préparation préalable. Malheureusement, les échantillons la silice et la montmorillonite ne sont pas visibles par cette méthode. Pour le mode transmission, une quantité de poudre est placée entre deux fenêtres  $\text{CaF}_2$  puis l'ensemble est placé sur le porte échantillon dans le compartiment du spectromètre sous atmosphère d'azote. Quelle que soit la technique utilisée, 250 scans sont accumulés pour un domaine spectral de  $7800$  à  $550\text{ cm}^{-1}$ .

Les échantillons ont été analysés par IR avant et après activation thermique afin de vérifier la formation de liaisons peptidiques durant l'activation. Après l'étape d'extraction, le culot est séché sous flux d'azote puis est également analysé par IR afin de vérifier l'efficacité de cette désorption.

Pour chaque acquisition, une correction de la ligne de base est effectuée automatiquement, puis une soustraction du spectre du support minéral est effectuée. En outre, en mode ATR, une compensation atmosphérique est effectuée.

Ces méthodes de caractérisation donnent des informations sur la formation de liaisons peptidiques, mais des techniques sensibles et spécifiques doivent être mises en œuvre afin de mieux caractériser la composition des échantillons. Dans nos travaux de recherches, les méthodes choisies sont les méthodes séparatives couplées avec la spectrométrie de masse, approches analytiques classiques du domaine de la protéomique et de l'analyse de mélanges complexes. Nous avons notamment privilégié l'utilisation de l'injection directe dans une instrumentation de type FTICR. Nous avons également eu accès ponctuellement à un spectromètre de masse Q-TOF pouvant être utilisé en couplage avec la chromatographie liquide ou la mobilité ionique. En raison de la complexité des échantillons et des quantités disponibles, la RMN n'a pas été choisie, elle nécessite d'avoir des concentrations élevées.

## III. Analyses *ex situ* de la solution de désorption

### III.1. HPLC-UV

Des essais préliminaires par chromatographie en phase inverse ont été faits sur un appareil HPLC Elite LaChrome (VWR-Hitachi) disposant d'un détecteur à barrette de diodes dans l'objectif de séparer les produits obtenus par condensation d'acides aminés sur une surface de silice. Le mode phase inverse a été choisi car il est couramment utilisé pour séparer des peptides et protéines. Les colonnes choisies, de marque Phenomenex, sont des colonnes C18, modifiées afin d'ajouter des interactions ioniques et d'améliorer la rétention des peptides les plus polaires. Sur l'une d'entre elles (PS-C18), des charges positives sont présentes en surface, sur l'autre (XB-C18), les silanols résiduels sont recouverts par des groupements isobutyles.

Malgré l'utilisation de colonnes adaptées à notre système, les chromatogrammes obtenus ne se sont pas interprétables, en raison du très grand nombre de composés de structures et d'hydrophobicité très proches dans l'échantillon. De plus, nous ne disposons pas de références de tous les peptides possibles, permettant l'attribution sans ambiguïté des pics chromatographiques.

Cette technique a donc été abandonnée au profit de la LC/MS.

### III.2. Spectromètre de masse à haute résolution FTICR

#### III.2.1. Description générale de l'appareil

Nous disposons d'un appareil hybride QhFTICR Solarix 7 T (Solarix<sup>XR</sup>, BrukerDaltonics, Germany) dont un schéma est montré Figure II-3.

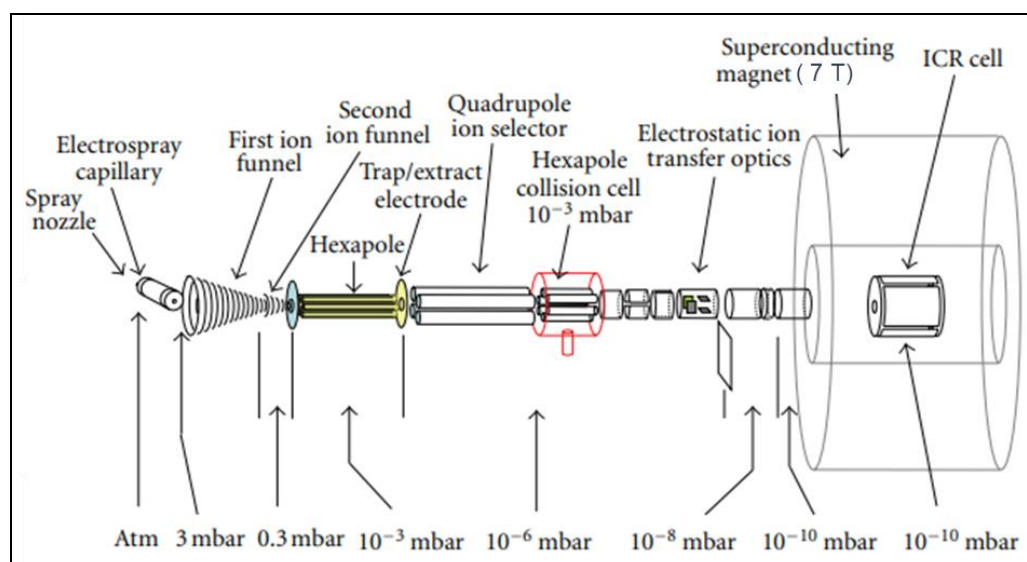


Figure II-3. Schéma d'un appareil FTICR<sup>200</sup>, modifié.

Cet instrument hybride est composé d'une source ESI permettant l'ionisation des molécules, d'un système d'optique, ainsi que de plusieurs analyseurs en masse dont la cellule ICR qui est sous vide ultra poussé.

Comme expliqué au Chapitre I, la cellule ICR permet de piéger les ions en appliquant un champ magnétique et électrostatique<sup>201</sup>. Elle a été mise en œuvre pour la première fois en 1936 par Penning. Depuis, différentes géométries de cellules ont été développées<sup>202-205</sup> dans le but d'augmenter la résolution sans changer la valeur du champ. Le FTICR utilisé est équipé d'une cellule de dernière génération dite « dynamically harmonized cell ». Le but de cette cellule, inventée par Nikolaev<sup>206</sup>, est d'obtenir un meilleur pouvoir résolutif. Pour cela, il faut garder la fréquence cyclotronique pour tous les ions d'un même rapport  $m/z$  durant le temps de détection du signal, ce qui nécessite d'abord un volume de cellule de piégeage plus grand. L'objectif de cette nouvelle géométrie est aussi de réduire l'influence du champ électrostatique axial et donc de minimiser le déphasage du paquet d'ions pendant l'étape de détection et en conséquence d'améliorer la résolution accessible. En effet, même sous un champ magnétique idéal homogène et sans interactions ion-neutre ou ion-ion, des paquets d'ions avec les mêmes valeurs  $m/z$  mais différentes amplitudes d'oscillation entraînent des fréquences cyclotroniques différentes, ce qui entraîne un décalage en fréquence et donc diminue le pouvoir résolutif.

Pour ce faire, Nikolaev a modifié la forme de la cellule. La surface interne des électrodes est sphérique, ce qui permet d'avoir des champs homogènes et de garder un paquet d'ions

dense. Le potentiel de piégeage est donc plus lisse et la distribution sera plus étroite. Cette nouvelle cellule, nommée cellule de Nikolaev, montre la meilleure performance en termes de pouvoir résolutif et de gamme d'excitation atteint pour un champ magnétique de 7 T. Cet appareil hybride donne la possibilité d'effectuer des expériences MS/MS en amont de la cellule ICR, dans un quadropôle en mode CID (Collision-Induced Dissociation). C'est le mode que nous avons privilégié.

### III.2.2. Conditions expérimentales

Les échantillons répertoriés dans le Tableau II-2 ont tous été analysés à l'aide du FTICR. Les solutions récupérées après étape de d'extraction sont diluées avec un mélange MeOH/H<sub>2</sub>O acidifié à 0,1 % en volume avec de l'acide formique afin de favoriser la formation d'espèces protonées. Le facteur de dilution est de 10, 25 ou 50 en fonction de la température d'activation et de la sensibilité de l'appareil. La dilution est ainsi optimisée afin d'obtenir des intensités de pics suffisantes tout en maintenant un nombre d'ions équivalent pour éviter les phénomènes d'espace de charge dans le système de piégeage pour chaque échantillon (voir Chapitre I).

Dans tous les cas, excepté les échantillons de nucléotides, l'analyse est opérée en ESI mode positif avec une gamme de détection allant de  $m/z$  100 à  $m/z$  2000. En fonction de l'étude, les échantillons ont été analysés par injection directe ou en utilisant un passeur automatique d'échantillon (mode FIA, *Flow Injection Analysis*). L'intérêt du mode FIA est de procéder à des analyses automatisées rapides et reproductibles, il a ainsi été utilisé notamment pour l'étude cinétique (voir Chapitre IV). Dans les deux cas, le temps d'accumulation des ions dans la cellule ICR choisi est de 0,2 s et le temps de vol est de 1,2 ms. Les paramètres opératoires les plus importants de la méthode sont répertoriés dans le Tableau II-3. Pour une injection directe, l'échantillon est injecté en mode infusion dans la source ESI avec un débit de 180  $\mu\text{L}/\text{h}$ . Dans le cas du mode FIA, un appareil de chromatographie liquide, commandé par le logiciel Chromeleon, est utilisé sans colonne. La méthode de chromatographie associée a une durée d'acquisition de 5 min, le débit est de 2  $\mu\text{L}/\text{min}$  et le volume injecté de 2  $\mu\text{L}$ . Le logiciel Compass Hystar est utilisé en mode FIA-FTICR.



Tableau II-3. Paramètres importants utilisés pour les analyses sur FTICR. Par soucis de clarté, les termes propres à l'instrument sont laissés en anglais.

Paramètres de la source		Paramètres de optiques de transfert	
Capillary	4500 V	Capillary exit	200 V
End plate offset	-500 V	Deflector plate	180 V
Nebuliseur	1,5 bar	Funnel 1	120 V
Dry gas	2,0 L/min	Skimmer 1	30 V
Dry temperature	200 °C	Funnel RF Amplitude	150 Vpp

Les expériences MS/MS ont été menées en utilisant le mode CID dans la cellule de collision hexapole pour étudier les fragmentations des produits de condensation. La fenêtre d'isolation de l'ion précurseur varie de 1 m/z à 5 m/z en fonction de l'intensité de l'ion précurseur et/ou de la présence d'ions proches en m/z. L'énergie de collision appliquée dépend de l'intensité relative de l'ion précurseur (Elab de 4-20 eV).

### III.3. Spectromètre de masse TIMS-TOF

#### III.3.1. Description générale de l'appareil

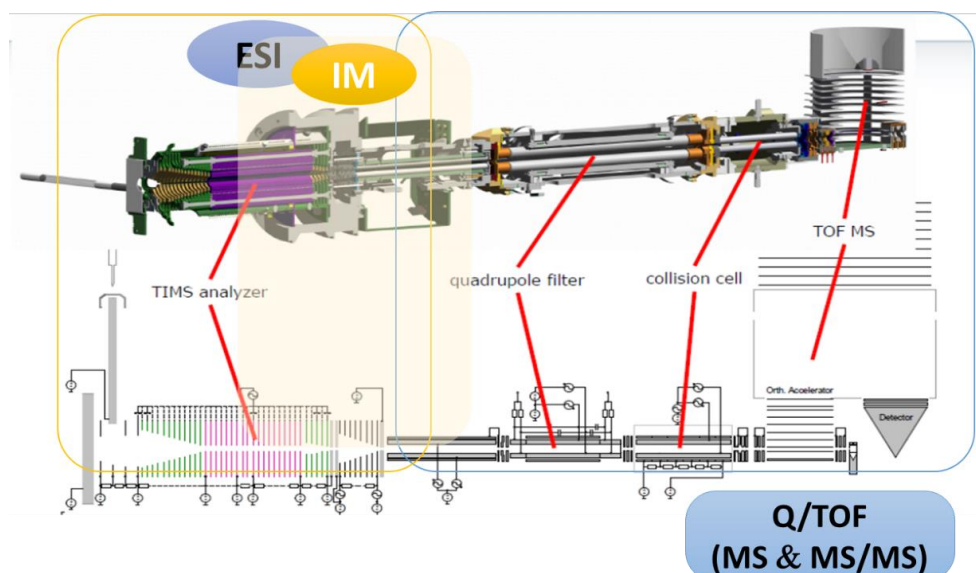


Figure II-4. Schéma de l'appareil TIMS-TOF. Communication personnelle S. Alves.

La Figure II-4 montre le schéma de l'appareil utilisé, un instrument hybride composé d'un quadripôle et d'un analyseur à temps de vol. Il peut être utilisé en couplage avec une LC

et/ou de la mobilité ionique (cellule TIMS). En effet, il comporte une cellule de mobilité ionique dite TIMS, récemment développé par Bruker, qui peut être utilisée pour faire du couplage IM-MS, ou pas (mode MS).

Qu'il soit utilisé avec ou sans séparation préalable, le quadripôle a deux modes de fonctionnement. Pour une analyse MS, le quadripôle est utilisé comme un guide d'ions additionnel (mode RF only), il ne filtre pas les ions en tant qu'analyseur en masse mais transmet le paquet d'ions de la source vers l'analyseur à temps de vol, une valeur d'énergie de collision faible est appliquée pour assurer la transmission tout en évitant la fragmentation. Pour les expériences MS/MS, le quadripôle est utilisé comme un filtre pour isoler un ion (précurseur) ou une gamme de masse (isolation ajustable entre 0,01 et 300 Daltons). Les ions précurseurs sélectionnés dans le quadripôle sont fragmentés dans la cellule de collision (mode CID) grâce à la présence d'un gaz de collision neutre (N<sub>2</sub>) et par accélération des ions (Elab).

L'analyseur à temps de vol (Time Of Flight, TOF) permet la mesure de masses précises des ions formés dans la source ESI (mode MS) ou des ions fragments (mode MS/MS). Le détecteur utilise la technologie MCP (microchannel plate). Il permet de convertir le signal des ions en signal électrique et de l'amplifier. L'ion entre dans un microcanal incurvé recouvert d'un dépôt métallique et soumis à un champ électrique. L'impact de l'ion entrant sur la paroi du microcanal provoque l'émission en cascade d'électrons, accélérés par un champ électrique.

### III.3.2. Utilisation de la mobilité ionique en couplage avec la spectrométrie de masse (approche IM-MS)

Le principe général de la mobilité ionique est expliqué dans le Chapitre I. Le spectromètre utilisé dans ce travail dispose d'une cellule TIMS (Trapped Ion Mobility Spectrometry) et repose sur le piégeage des ions et leur libération successive, permettant leurs séparations par mobilité<sup>207</sup>.

### III.3.2.1. Description de la cellule de mobilité ionique

La Figure II-5 montre un schéma plus détaillé d'une cellule TIMS et notamment des trois régions qui la composent (funnel d'entrée, tunnel TIMS et funnel de sortie). L'utilisation de cette cellule ajoute une dimension séparative en phase gazeuse basée sur le rapport  $\Omega/z$  de l'ion, avec  $\Omega$  la section efficace de collision et  $z$  sa charge (voir Chapitre I). Son principe de séparation des ions en phase gazeuse diffère des cellules classiques DTIMS. Au lieu de conduire les ions à travers un gaz stationnaire comme dans une cellule de dérive classique, le TIMS agit comme un système de piégeage et procède par des étapes temporelles<sup>119</sup>. D'abord le flux de gaz entraîne les ions le long de l'appareil tandis que le champ électrique les ralentit<sup>116</sup>. Chaque ion migre dans le tube jusqu'à la région où le champ électrique est suffisamment fort pour compenser les forces de dérive générées par le flux de gaz. L'ion s'arrête donc à une position spécifique le long de la cellule TIMS et est ainsi piégé par combinaison d'un gradient de champ électrique axial et d'un potentiel radiofréquence dans la dimension radiale. Le temps de piégeage est souvent de l'ordre de quelques millisecondes.

Les ions sont ensuite « élus » : la rampe de potentiels (profil EFG, Electric Field Gradient) est baissée graduellement de sa valeur initiale  $E_0$  jusqu'à une valeur définie pour libérer les ions. Les ions sont élus en fonction de leur mobilité : les premiers ions détectés sont ceux ayant la mobilité la plus faible. Les ions avec  $\Omega$  grand auront une plus grande énergie cinétique ; ils pourront aller plus loin sur la rampe de potentiels donc un potentiel plus élevé est nécessaire pour les piéger. La distance que les ions parcourent dépend de la vitesse du gaz et de la vitesse à laquelle le profil EFG décroît<sup>207</sup>. Les ions sont donc libérés à une valeur de champ électrique donnée, qui permet de calculer leurs mobilités avec l'équation (1) du Chapitre I (§ II.2.2.1). Dans une cellule TIMS, le calcul de CCS à partir de  $K$  n'est pas direct ; une étape d'étalonnage avec des standards de valeurs CCS connues est obligatoire.

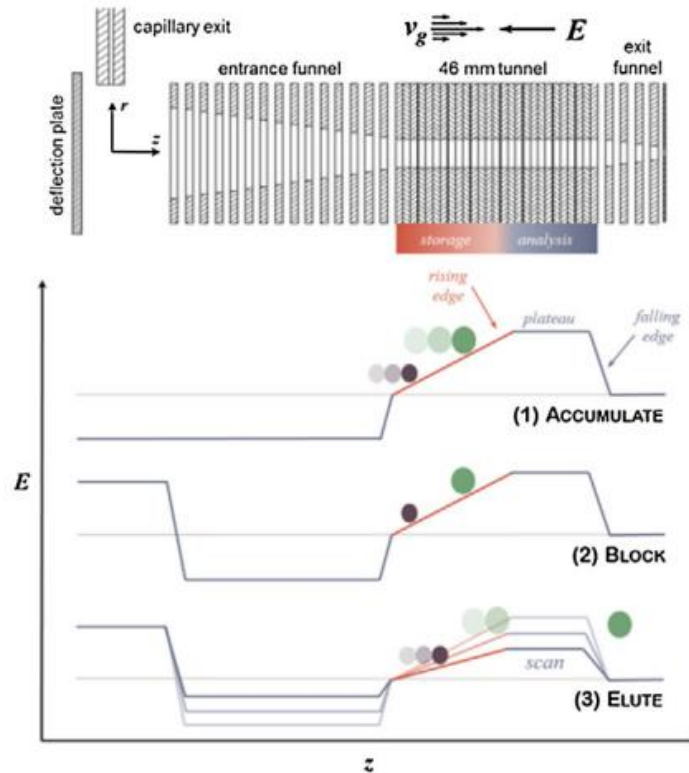


Figure II-5. Représentation schématique d'une cellule TIMS et de son système de fonctionnement<sup>208</sup>.

L'intérêt de cette cellule innovante réside dans sa versatilité. En effet, la résolution de la mobilité ionique, qui dépend de la vitesse à laquelle la rampe de potentiel est baissée, peut être changée au besoin : plus elle est baissée lentement, meilleure est la résolution. La Figure II-6 montre ainsi les trois modes de détection possibles : survey, detect ou ultra. Cependant le gain en résolution se fait au détriment de la gamme en mobilités réduites détectées. Dans le mode ultra, on se focalise sur une faible gamme de  $1/K_0$ . Dans le cas de notre instrument TIMS, on mesure la mobilité réduite  $1/K_0$  car on travaille dans des conditions standards de pression et de température.

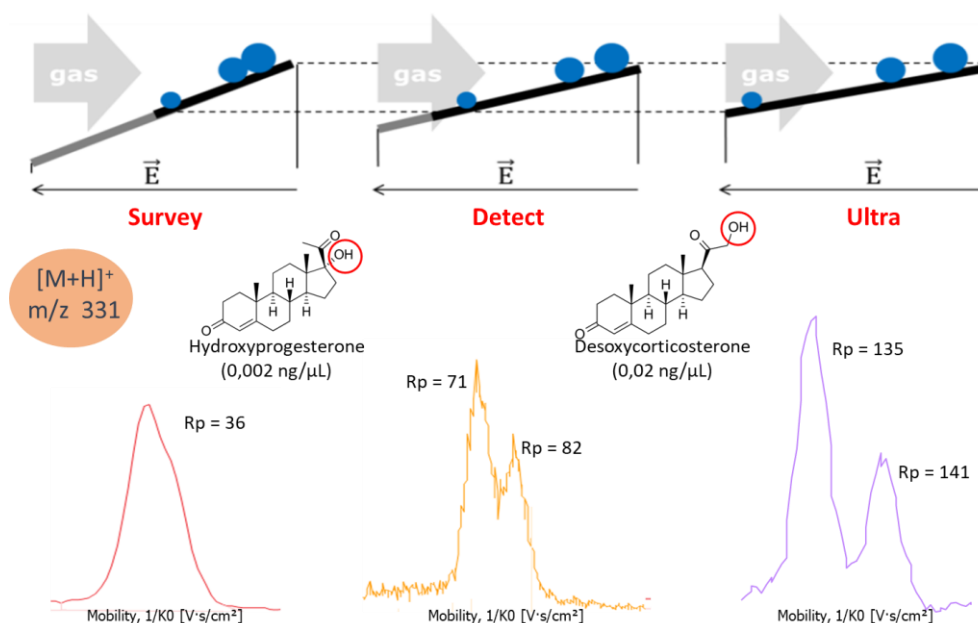


Figure II-6. Comparaison des mobilogrammes obtenus en fonction du mode d'acquisition choisi pour la distinction de deux isomères. Le mode ultra permet d'accéder à de meilleures résolutions mais la gamme de mobilité étudiée est plus faible. Communication personnelle S. Alves.

### III.3.2.2. Conditions expérimentales des expériences IM-MS

Les expériences de mobilité ionique ont été effectuées en infusion directe sur une gamme de masse de  $m/z$  100 à  $m/z$  1350. Le mode utilisé par défaut est le mode detect, qui permet d'avoir une large gamme de détection de mobilités réduites  $1/K_0$  0,55-1,9  $V \cdot s \cdot cm^{-2}$ , correspondant à une gamme large en  $m/z$ . Le temps d'abaissement de la rampe de potentiel (« ramp time ») est de 100 ms. Le mode ultra permet d'améliorer la résolution des mobilogrammes en se focalisant sur une gamme de  $1/K_0$ . Le pouvoir résolutif (ou résolution) d'un pic de mobilogramme correspond au rapport  $K_0/\Delta K_0$ . Pour le mode ultra, le pouvoir résolutif est d'environ 120 tandis qu'il est de 50-70 pour le mode detect. Dans les cas où le mode ultra est utilisé, la gamme de mobilités sera précisée dans les résultats. Un contrôle du nombre d'ions ICC (ion current control) est de 2 Moi. Les paramètres importants de la méthode sont regroupés dans le Tableau II-4.

Tableau II-4. Paramètres importants utilisés pour les analyses IM-MS. Par soucis de clarté, les termes propres à l'instrument sont laissés en anglais.

Paramètres source			
End plate offset	500 V	Dry gas	5 L/min
Capillary	4200 V	Dry temp	200 °C
Nebulizer	1 bar	Flow	10 µL/min
Paramètres TIMS			
Δ1 (Deflection transfer → capillary exit)	-20 V	Δ5 (Accumulation exit → accumulation transfer)	0 V
Δ2 (Deflection transfer → deflection discard)	-120 V	Δ6 (Ramp start → Accumulation exit)	100 V
Δ3 (Funnel 1 In → Deflection transfer)	80 V	Funnel 1 RF	250 Vpp
Δ4 (Accumulation Trap → Funnel 1 In)	100 V	Collision Cell In	220 V
Paramètres MS			
Transfer		Collision cell	
Funnel 1 RF	250 Vpp	Collision energy	10 eV
IsCID Energy	0 Ev	Transfer time	54 µs
Deflection delta	80 V	Collision RF	1100 Vpp
Funnel 2 RF	250 Vpp	Pre pulse storage	5 µs
Multipole RF	200 Vpp		

### III.3.3. Approche LC/MS

Indépendamment de la cellule de mobilité, le spectromètre QToF peut être couplé à un système de chromatographie liquide. Les échantillons sont prélevés avec un passeur automatique puis traversent une colonne de chromatographie permettant la séparation des composés avant d'être injectés dans la source ESI. Pour chaque acquisition, une solution de cluster est injectée avant l'échantillon à analyser pour un étalonnage interne.

La gamme de masse étudiée est  $m/z$  100-2000. Un gradient entre les phases aqueuse et organique est opéré. La phase aqueuse A est composée de 95 % de H<sub>2</sub>O et 5 % MeOH et est acidifié avec 0,1 % volumique d'acide formique. La phase organique B est composée uniquement de MeOH acidifié avec 0,1 % volumique d'acide formique. Parmi les différents gradients testés, celui qui a été utilisé est détaillé au Tableau II-5.

Différentes colonnes ont été testées : une colonne C18 classique Intensity Solo 2 C18 (100 x 2,1 mm) distribuée par Bruker et une colonne C18 modifiée présentant des charges positives en surface (PS-C18 de Phenomenex). Les séparations ne montrant pas de différences significatives, la colonne choisie est la C18 classique.

Tableau II-5. Gradient utilisé pour les séparations de chromatographie liquide.

Temps (en min)	% phase A	% phase B
0	100	0
5	100	0
85	25	75
100	0	100
105	0	100
106	100	0
120	100	0

Les paramètres importants de la méthode sont regroupés dans le Tableau II-6.

Tableau II-6. Paramètres importants utilisés pour les analyses LC/MS. Par soucis de clarté, les termes propres à l'instrument sont laissés en anglais.

Paramètres source			
End plate offset	500 V	Dry gas	9 L/min
Capillary	3500 V	Dry temp	200 °C
Nebulizer	2.8 bar	Débit	200 µL/min
Paramètres MS			
Transfer		Collision cell	
Funnel 1 RF	80 Vpp	Collision energy	5 eV
IsCID Energy	0 Ev	Transfer time	20 - 90 µs
Deflection delta	70 V	Collision RF	200-1000 Vpp
Funnel 2 RF	100 Vpp	Pre pulse storage	6 µs
Multipole RF	100 Vpp		

## IV. Considérations théoriques sur la polymérisation des acides aminés

Une des questions soulevées dans le présent travail est celle de la sélectivité de polymérisation des acides aminés. Par rapport à cette question, « l'hypothèse-zéro » est celle d'une polymérisation aléatoire où, pour chaque position dans un oligopeptide de longueur  $n$ , chaque acide aminé apparaîtrait avec la même probabilité que sa probabilité statistique dans le mélange initialement introduit.

Considérons un système binaire, par exemple le système de base Leu + Glu que nous avons le plus étudié. Soit  $x$  la fraction molaire de Leu ; la fraction molaire de Glu est donc  $(1-x)$ . Dans notre échantillon de référence, contenant des quantités équimolaires de Leu et Glu,

$x = (1-x) = 1/2$ . Les oligopeptides de longueur  $n$  contiennent  $2^n$  séquences possibles, par exemple pour des tripeptides, les séquences LLL, LLE, LEL, LEE, ELL, ELE, EEL, EEE. Si la polymérisation est aléatoire, chacune de ces séquences est équiprobable, avec une probabilité de  $2^{-n}$  ( $= 1/8$ ).

Toutefois, la séquence ne peut être déterminée directement par spectrométrie de masse avec laquelle on peut uniquement avoir accès à la stœchiométrie, de type  $E_m L_{n-m}$ . Le nombre de séquences de cette stœchiométrie se calcule en résolvant le coefficient binomial de l'équation (1).

$$\binom{n}{m} = \frac{n!}{m!(n-m)!} \quad (1)$$

Les probabilités des différentes stœchiométries dans le modèle aléatoire pour les tripeptides sont données dans le Tableau II-7, à titre d'exemple.

Tableau II-7. Probabilités des différentes stœchiométries dans le modèle aléatoire.

Stœchiométrie	E <sub>3</sub>	EL <sub>2</sub>	E <sub>2</sub> L	L <sub>3</sub>
m	3	1	2	0
$\binom{n}{m}$	1	3	3	1
Probabilité	0,125	0,375	0,375	0,125

Au cours de ce travail, des échantillons contenant Leu et Glu en proportions non équimolaires seront étudiés. Dans le cas d'une polymérisation aléatoire d'un mélange contenant une fraction molaire  $x$  quelconque de Glu, la probabilité qu'un oligomère de longueur  $n$  contienne exactement  $m$  résidus Glu est donnée par l'équation (2), appelée « fonction de masse ».

$$f(m, n, x) = \binom{n}{m} \cdot x^m \cdot (1-x)^{n-m} = \frac{n!}{m!(n-m)!} \cdot x^m \cdot (1-x)^{n-m} \quad (2)$$

Cette équation est utilisable pour des mélanges contenant plus de deux acides aminés, en utilisant les fractions molaires propres de chacun des acides aminés. A titre d'exemple, pour un mélange équimolaire de quatre acides aminés A, B, C, D, les fractions molaires de chacun des acides aminés valent  $x_A = x_B = x_C = x_D = 0,25$ . L'équation (2) appliquée aux dipeptides s'écrit  $f(m, 2, 0,25) = \frac{2!}{m!(2-m)!} \cdot (0,25)^m \cdot (0,75)^{2-m}$  et les probabilités de contenir zéro, un, ou deux résidus A sont :



$$P(0A) = f(0, 2, 0,25) = \frac{2!}{0!2!} \cdot (0,25)^0 \cdot (0,75)^2 = 0,5625.$$

$$P(1A) = f(1, 2, 0,25) = \frac{2!}{1!1!} \cdot (0,25)^1 \cdot (0,75)^1 = 0,375.$$

$$P(2A) = f(2, 2, 0,25) = \frac{2!}{2!0!} \cdot (0,25)^2 \cdot (0,75)^0 = 0,0625.$$

(La somme des trois probabilités vaut bien 1.)

Ces calculs de distributions seront utilisés au Chapitre III pour comparer les résultats expérimentaux à « l'hypothèse zéro » d'une polymérisation aléatoire. Tout écart des résultats expérimentaux par rapport à ces valeurs est l'indice d'une sélectivité dans la formation des oligomères dans nos synthèses prébiotiques.

Il n'est pas possible d'obtenir n'importe quelle distribution de séquences ni de stœchiométries des oligopeptides à partir d'un mélange de composition donnée. Soit  $n_L^o$  le nombre de moles de Leu introduit dans le système,  $n_E^o$  le nombre de moles de Glu introduit, et  $n_{ij}$  le nombre de moles de chaque polypeptide de stœchiométrie  $E_iL_j$  formé après activation. Si on suppose que les acides aminés ne peuvent être ni détruits (par dégradation), ni éliminés (par désorption en phase gazeuse), la conservation de la matière fournit :  $\sum_{i,j=0}^{\infty} j \cdot n_{ij} = n_L^o$  pour la leucine, et  $\sum_{i,j=0}^{\infty} i \cdot n_{ij} = n_E^o$  pour l'acide glutamique.

Alternativement, on peut exprimer la contrainte de conservation de la masse en termes de fractions massiques. Soient  $w_{ij}$  les fractions massiques des oligomères individuels après condensation, et  $w_E^o$ ,  $w_L^o$ , les fractions massiques des deux acides aminés initialement introduits. On doit avoir :

$$\frac{\sum_{i,j=0}^{\infty} i \cdot w_{ij}}{\sum_{i,j=0}^{\infty} (i + j) \cdot w_{ij}} = w_E^o \quad (3)$$

Et une formule semblable pour la leucine.

On pourrait donc vérifier la cohérence des analyses d'après ces deux relations, si on avait accès avec précision aux quantités  $n_{ij}$ , ou aux  $w_{ij}$ , ou, ce qui revient au même, aux concentrations  $c_{ij}$  dans la solution de désorption. Néanmoins nous verrons au Chapitre III, qu'il est très difficile d'évaluer avec précision les concentrations en oligopeptides dans nos systèmes.

Chapitre III. Sélectivité de  
polymérisation dans un système binaire  
(L+E)/SiO<sub>2</sub>

La plupart des résultats expérimentaux de ce chapitre ont été publiés dans *ACS Earth Space Chem.* 2020, 4, 1802-1812, reproduit en Annexe 1. Rappelons (cf. Chapitre II) que notre « système de base » est un échantillon (L+E)/SiO<sub>2</sub> préparé par imprégnation à humidité naissante, contenant 1.10<sup>-4</sup> mole d'acide glutamique et 1.10<sup>-4</sup> mole de leucine par gramme de silice, soit une charge massique de 1,31 %<sub>m</sub> Leu et 1,47 %<sub>m</sub> Glu (toujours par rapport au support silice). Dans ce chapitre, lorsque l'on parle de (L+E)/SiO<sub>2</sub> sans autre précision, c'est de cet échantillon qu'il s'agit. De même, lorsque l'on parle « d'activation », il s'agit d'une activation thermique à 160 °C dans une cellule traversée par un flux d'argon, avec une rampe de 1 °C par minute et un plateau final de 30 minutes.

## I. Etude préliminaire du système Leu + Glu

Plusieurs techniques sont utilisées pour étudier *in-situ* l'échantillon (L+E)/SiO<sub>2</sub> avant et après activation thermique (voir le Chapitre II pour les protocoles).

Le diffractogramme de l'échantillon (L+E)/SiO<sub>2</sub>, reporté en Figure III-1a, ne montre aucun pic de Bragg fin, seulement un signal très large dû à la diffusion des RX par la silice amorphe. Tous les acides aminés sont donc adsorbés sur la surface de silice puisqu'ils ne forment pas de cristaux massiques. Cette conclusion est compatible avec les « recouvrements de saturation » déterminés pour le même système par Sakhno et al<sup>89</sup>.

L'analyse thermogravimétrique (ATG) de l'échantillon modèle (L+E)/SiO<sub>2</sub> est présentée en Figure III-1b. Le thermogramme montre trois évènements. Le premier, avec un maximum à 38 °C, correspond à l'évaporation de l'eau physisorbée sur la silice. Deux évènements sont observés à 115 °C et 148 °C, dans une région où le support silice lui-même ne présente pas de perte de poids significative. Comme cela a été rapporté dans des études précédentes<sup>76,80,209</sup>, des pics dans cette gamme de température peuvent être attribués à une réaction de condensation entre des groupements amine et acide carboxylique, entraînant la formation d'une liaison amide et la libération d'une molécule d'eau. Pour la plupart des acides aminés, un seul évènement de condensation est observé dans cette région. La présence de deux évènements séparés est ici typique de systèmes contenant des acides aminés trifonctionnels comme l'acide glutamique, pour lesquels une cyclisation interne de ce résidu se produit. Ces analyses justifient le choix des conditions d'activation thermique, à

savoir une température finale de 160 °C, puisque les évènements attribuables à la condensation peptidique se produisent avant cette valeur limite de 160 °C. Une rampe de température de 1 °C/min est également choisie. Au-delà de 200 °C (données non présentées), d'autres évènements de perte de masse correspondant à des dégradations des molécules organiques sont observés. Nous n'avons pas exploré ce régime de températures élevées, car les biomolécules sont alors dégradées.

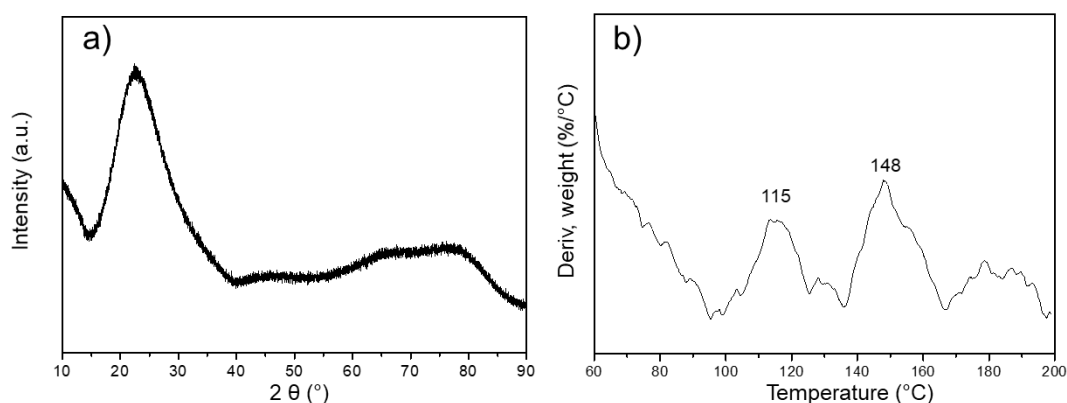


Figure III-1. a) Diffractogramme RX et b) profil thermogravimétrique de l'échantillon (L+E)/SiO<sub>2</sub>.

La Figure III-2 compare les spectres IR obtenus à partir de l'échantillon (L+E)/SiO<sub>2</sub>, avant et après activation thermique, et après l'étape de désorption. Seule la région 1800-1300 cm<sup>-1</sup>, caractéristique des bandes organiques, est représentée. En-dessous de 1300 cm<sup>-1</sup>, le signal est saturé à cause de l'absorption par la silice. La région 3200-3800 cm<sup>-1</sup> montre des bandes dues aux vibreurs OH (silanols de surface, eau) et NH (fonctions amine/ammonium), mais ces signaux sont larges et d'interprétation difficile.

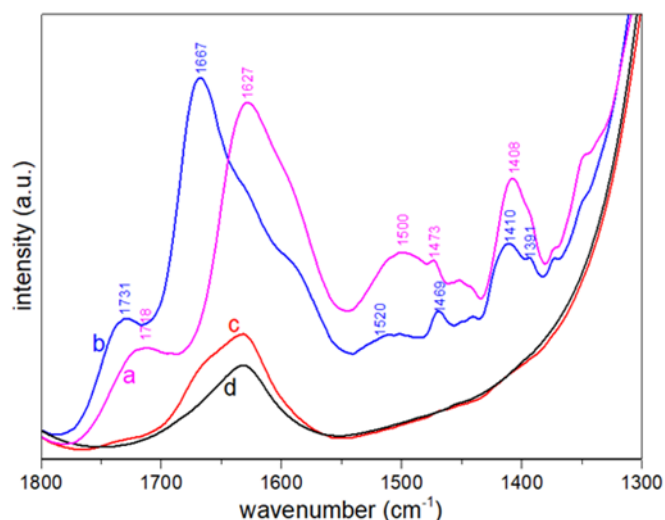


Figure III-2. Spectres IR normalisés sur la gamme 1800-1300 cm<sup>-1</sup> de l'échantillon (L+E)/SiO<sub>2</sub> (a) avant et (b) après activation thermique, (c) du solide après désorption dans de l'eau et (d) du support SiO<sub>2</sub>.

Le spectre IR permet de connaître la spéciation des acides aminés adsorbés. La bande à 1627 cm<sup>-1</sup> correspond à la déformation asymétrique du groupement NH<sub>3</sub><sup>+</sup> ( $\delta_{as\ NH_3^+}$ ), la bande à 1500 cm<sup>-1</sup> correspond à sa déformation symétrique. L'épaule à 1592 cm<sup>-1</sup> correspond à la vibration d'élongation asymétrique  $\nu_{as}(\text{COO}^-)$  du CO<sub>2</sub><sup>-</sup> de la fonction acide aminé, et la bande à 1408 cm<sup>-1</sup> à sa vibration d'élongation symétrique. On a donc des fonctions carboxylate déprotonées, et des fonctions ammonium protonées : une fois adsorbés, les acides aminés sont donc présents sous forme zwitterionique comme dans la solution, conformément à ce qui avait été observé par Sakhno et al<sup>89</sup> et par Abadian<sup>91</sup>. On voit aussi une bande à 1718 cm<sup>-1</sup> attribuable à l'élongation C=O dans la fonction acide carboxylique (non déprotonée, elle) dans la chaîne latérale de l'acide glutamique.

La plupart de ces bandes d'absorption disparaissent après activation thermique. Deux nouvelles bandes apparaissent : à 1667 cm<sup>-1</sup>, assignée à la vibration amide I (principalement due à l'élongation  $\nu(\text{C=O})$ ), et à 1520 cm<sup>-1</sup>, assignée à la vibration amide II, révélant la formation d'une liaison amide durant l'activation thermique. Il ne reste pratiquement plus d'acides aminés sous formes de monomères, et des peptides ont été formés. De plus, la présence de la bande amide II implique la présence de polypeptides linéaires<sup>89</sup>, car cette bande n'est pas présente dans les dipeptides cycliques de type DKP.

Par ailleurs, la spectroscopie IR a été utilisée pour estimer l'efficacité de l'étape de désorption<sup>28</sup>. En effet, l'intensité des bandes des molécules organiques en surface peut être

estimée en soustrayant le spectre du support de silice à celui de l'échantillon, après normalisation pour les bandes du support (les bandes à 1865 et 1980 cm<sup>-1</sup> sont généralement utilisées à cet effet). En appliquant cette procédure, il apparaît que l'échantillon désorbé (spectre c de la Figure III-2) conserve moins de 5 % de l'intensité des bandes représentant les molécules organiques, comparé au même échantillon avant désorption (spectre b).

Les produits de réaction de l'activation thermique sont donc presque quantitativement désorbés en une seule étape de lavage. Nous ne développons pas cette conclusion car elle confirme les résultats obtenus préalablement sur nos systèmes. Mais elle est très importante, car elle signifie que l'étape d'extraction nécessaire à une analyse *ex situ* n'introduit pas *à priori* de ségrégation des molécules présentes selon leur affinité avec la surface.

Il est donc possible d'analyser la solution de désorption, notamment par spectrométrie de masse à haute résolution (HRMS High Resolution Mass Spectrometry), et d'obtenir des résultats qui reflètent effectivement la composition du mélange formé sur la surface de silice. L'efficacité de la désorption a été vérifiée pour chaque nouveau système étudié.

## II. Sélectivité de composition des produits de polymérisation

La solution de désorption de l'échantillon (L+E)/SiO<sub>2</sub> activé à 160 °C a été ensuite analysée à l'aide d'un instrument FTICR avec la procédure décrite dans le Chapitre II. Le spectre de masse ESI en mode positif est reporté dans la Figure III-3.

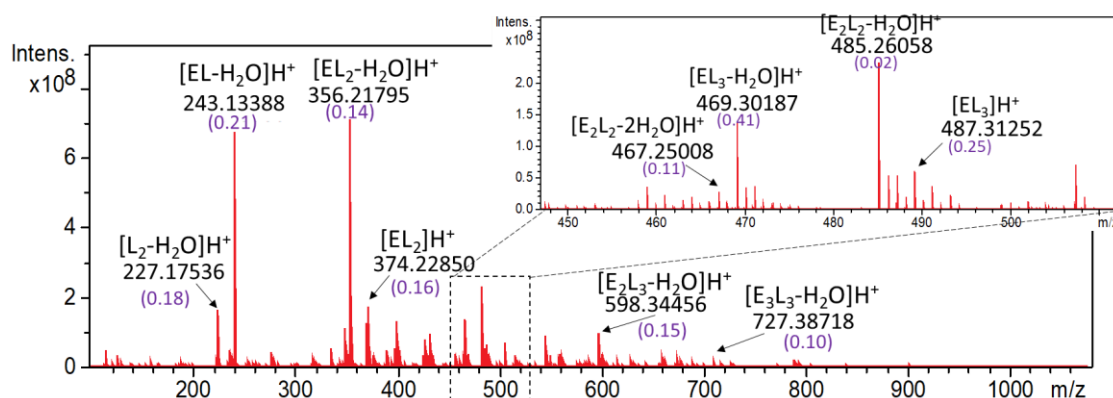


Figure III-3. Spectre de masse ESI en mode positif des produits d'activation de l'échantillon (L+E)/SiO<sub>2</sub> (instrument FTICR). L'erreur en ppm sur les mesures de m/z est reportée entre parenthèses.

La majorité des pics peut être attribuée à des produits de polymérisation peptidique. Des ions protonés et cationisés d'homo- et d'hétéro- oligopeptides sont détectés, prouvant la formation d'oligopeptides par activation thermique. Des oligopeptides allant jusqu'aux octapeptides (huit résidus d'acides aminés) sont détectés sous formes d'ions moléculaires avec une précision de mesure inférieure à 1 ppm, avec un degré de certitude élevé de l'annotation. Les quatre pics majoritaires sont attribués respectivement aux stœchiométries [EL-H<sub>2</sub>O]H<sup>+</sup>, [EL<sub>2</sub>-H<sub>2</sub>O]H<sup>+</sup>, [EL<sub>2</sub>]H<sup>+</sup> et [E<sub>2</sub>L<sub>2</sub>-H<sub>2</sub>O]H<sup>+</sup>.

Les espèces les plus abondantes correspondent à ces formes oligomériques de type « déshydraté ». En effet, après les quatre pics majeurs déjà cités, les espèces suivantes par ordre d'intensité des pics sont [L<sub>2</sub>-H<sub>2</sub>O]H<sup>+</sup>, [E<sub>2</sub>L-H<sub>2</sub>O]H<sup>+</sup> et [E<sub>2</sub>L<sub>3</sub>-H<sub>2</sub>O]H<sup>+</sup>.

Afin d'étudier la reproductibilité et de vérifier les possibles effets de vieillissement des échantillons, plusieurs analyses HRMS ont été réalisées sur des solutions désorbées de l'échantillon (L+E)/SiO<sub>2</sub> après différents temps de stockage du solide, immédiatement après désorption ou après différents cycles de congélation/décongélation. Les spectres de masse de ces différents échantillons sont similaires, indiquant que les conditions de stockage ne modifient pas significativement la composition du mélange d'oligopeptides.

## II.1. Etude de la stœchiométrie des oligopeptides formés

### II.1.1. Distributions expérimentales *versus* polymérisation aléatoire

L'objectif initial de ce travail était d'étudier la sélectivité et la spécificité de polymérisation des acides aminés dans les scénarios d'activation sur surface.

Le premier aspect de la question de la sélectivité est de savoir si la production de stœchiométries privilégiées pour une même longueur de peptides est obtenue. Si on prend l'exemple des tripeptides, les stœchiométries possibles sont L<sub>3</sub>, EL<sub>2</sub>, E<sub>2</sub>L et E<sub>3</sub>. En considérant une polymérisation aléatoire pour deux acides aminés présents en quantité équimolaire, la probabilité que chaque position dans la chaîne soit occupée par Leu ou Glu est de 0,5 ( $p_E = p_L = 0,5$ ). D'après les considérations développées au Chapitre II, les rapports attendus entre les tripeptides de chaque stœchiométrie sont donc 1 : 3 : 3 : 1, ou 1/3 : 1 : 1 : 1/3 en normalisant à la quantité la plus importante. Il peut sembler assez simple de comparer ces quantités relatives calculées selon « l'hypothèse-zéro » (polymérisation aléatoire) avec les quantités relatives expérimentalement observées.

Toutefois, dans les expériences de spectrométrie de masse, on ne mesure pas directement les concentrations dans la solution analysée, mais l'intensité des pics (à partir des ions produits en ESI). Les observations expérimentales sont donc *a priori* biaisées par des phénomènes de discrimination<sup>129</sup> (voir la discussion dans le Chapitre I). En pratique, chaque molécule sera caractérisée par une « efficacité d'ionisation » différente, difficilement prédictible. Des études vraiment quantitatives par spectrométrie de masse nécessitent l'utilisation de composés de référence purs, généralement des analogues marqués par des isotopes stables, et les méthodes sont développées pour un (ou une famille de) composé(s) ciblé(s)<sup>210</sup>.

### II.1.2. Evaluation des efficacités d'ionisation des oligopeptides

Dans le cas des tripeptides, nous avons pu caractériser les peptides de référence purs, afin d'évaluer les différences d'efficacité d'ionisation entre les différentes séquences. Nous avons analysé selon notre procédure standard des mélanges des huit tripeptides de référence, à savoir les séquences : LLL, EEE, LLE, ELL, LEL, LEE, EEL et ELE. Chaque mélange était composé



en quantités équimolaires de trois tripeptides de référence (LLL, EEE et un des six autres hétéro-tripeptides isomères). Les spectres de masse ESI résultants sont reportés dans la Figure III-4, ci-dessous.

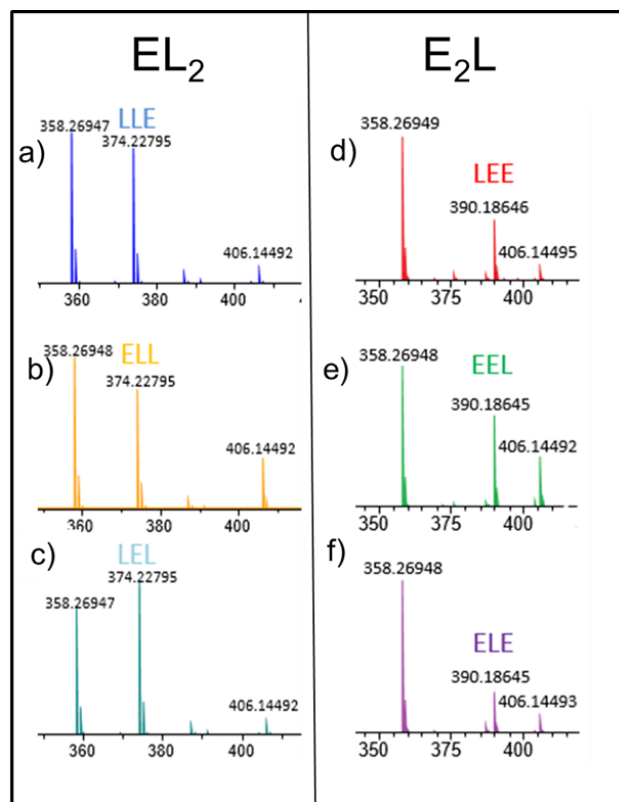


Figure III-4. Spectre de masse ESI en mode positif de mélanges de tripeptides de référence (instrument FTICR). Chaque solution contient en quantités égales les tripeptides LLL ( $m/z$  358,26945), EEE ( $m/z$  406,14492) et un hétéro-tripeptide de séquence a) LLE, b) ELL, c) LEL, d) LEE, e) EEL et f) ELE.

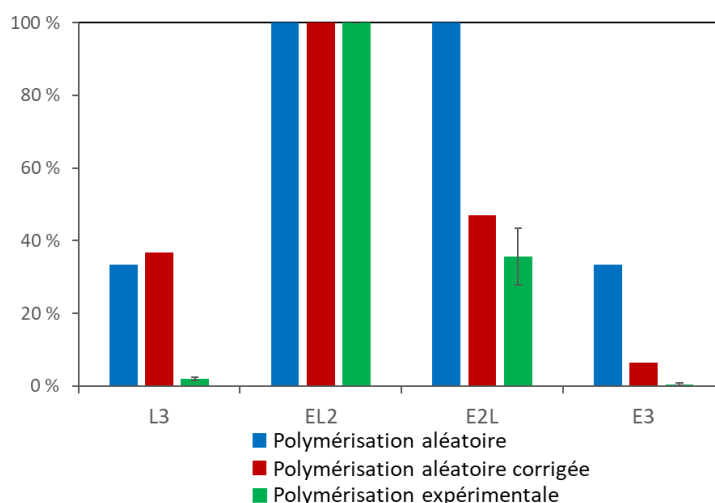
Comme prévu, les tripeptides montrent des efficacités d'ionisation différentes. De façon générale, plus les peptides sont riches en leucine, plus leur efficacité d'ionisation est importante. Cela concorde avec les prévisions théoriques. En effet, dans l'ionisation par électrospray, les analytes ayant des parties hydrophobes ont une réponse ESI plus importante que ceux ayant des parties plus polaires en raison de la grande affinité pour la surface des gouttelettes électrospray<sup>129</sup> (voir Chapitre I). Selon nos résultats expérimentaux, un ordre relatif des efficacités d'ionisation peut être établi :  $L_3 > EL_2 > E_2L > E_3$ , la différence entre les efficacités d'ionisation de  $L_3$  et  $E_3$  atteignant un ordre de grandeur. Par ailleurs, les tripeptides de même stœchiométrie mais de séquences différentes ont des réponses MS proches.

Il était impossible de mener cette démarche jusqu'à son terme parce que les espèces peptidiques observées expérimentalement incluent, outre des formes cationisées minoritaires, les formes déshydratées déjà mentionnées plus haut, dont la nature chimique est complexe (cf. Chapitre VI) et qui ne sont en tout état de cause généralement pas disponibles à l'état pur.

Nous avons donc fait l'hypothèse simplificatrice que l'ordre d'efficacité d'ionisation observé pour les tripeptides intacts, et déterminé plus haut, restait valable pour les tripeptides déshydratés, et avons recalculé les intensités relatives attendues pour une polymérisation aléatoire en tenant compte de la réponse analytique. D'autre part, nous avons sommé pour chaque stœchiométrie les intensités expérimentales de tous les ions d'oligopeptides de cette stœchiométrie, qu'ils soient protonés ou cationisés, intacts ou déshydratés.

### II.1.3. L'exemple des tripeptides : preuve d'une sélectivité de composition

La Figure III-5 compare ainsi les intensités attendues pour une polymérisation aléatoire (barres rouges ; abondances corrigées avec les facteurs de réponse des peptides de référence) à celles expérimentalement observées (en vert). Les valeurs expérimentales correspondent à des valeurs moyennes sur sept répétitions ; les barres d'erreurs correspondant aux écarts-types sont reportées sur la distribution expérimentale (en vert).



*Figure III-5. Comparaison de la distribution des intensités des tripeptides pour une polymérisation expérimentale par rapport à la polymérisation aléatoire corrigée (voir discussion ci-dessus). Toutes les distributions sont normalisées par rapport à EL<sub>2</sub>, pris comme référence à 100 %.*

Il apparaît que la correction tenant compte des efficacités d'ionisation est indispensable pour évaluer plus finement les répartitions des stœchiométries. Ainsi, les données non corrigées laisseraient penser qu'il y a une prédominance importante des tripeptides de stœchiométrie EL<sub>2</sub> par rapport aux E<sub>2</sub>L. Après correction, la différence est moins nette, encore qu'elle reste statistiquement significative.

Au contraire, l'absence presque totale de peptides L<sub>3</sub> est encore plus remarquable après correction car cette espèce a un facteur de réponse important, et donc devrait être particulièrement visible. On peut également noter l'absence d'espèces de stœchiométrie E<sub>3</sub>. La procédure que nous avons utilisée pour évaluer de façon semi-quantitative la sélectivité dans le cas des tripeptides est exigeante et nécessite de disposer de tous les peptides différents sous forme de standards purs. Le nombre de séquences possibles étant égal à 2<sup>n</sup>, deux fois plus d'expériences sont nécessaires chaque fois qu'on augmente la longueur du polypeptide d'une unité. Nous n'avons donc pas appliqué systématiquement cette approche. Néanmoins, la Figure A1 en Annexe 2 compare les intensités expérimentales obtenues aux intensités théoriques non corrigées pour une polymérisation aléatoire pour toutes les longueurs jusqu'aux octapeptides. De façon générale, pour toutes les tailles d'oligopeptides, une abondance bien supérieure est observée pour les stœchiométries contenant les deux acides aminés (hétéro-oligopeptides), par rapport aux valeurs attendues pour une distribution aléatoire.

## II.2. Origine de la sélectivité de stœchiométrie

La sélectivité de la réaction de polymérisation peut avoir deux origines. Elle peut dépendre d'un contrôle thermodynamique, c'est-à-dire que certains oligopeptides seraient formés préférentiellement car plus stables intrinsèquement, ou dépendre d'un contrôle cinétique, c'est-à-dire que la surface de silice favorise une voie mécanistique aboutissant à certains peptides particuliers en abaissant son énergie d'activation (catalyse hétérogène). Afin de comprendre l'origine de la sélectivité de polymérisation obtenue sur surface de silice, nous avons effectué une réaction de polymérisation dans d'autres conditions : en solution aqueuse, en utilisant un activateur sacrificiel, le N,N-carbonyldiimidazole (CDI). La Figure III-6 montre le mécanisme de condensation peptidique induit par le CDI. A l'issue de la séquence

de réactions, une liaison peptidique est formée, et le CDI est décomposé en imidazole et CO<sub>2</sub>, la réaction globale étant thermodynamiquement favorable.

Bien que la voie d'activation par CDI ait été à plusieurs reprises utilisée dans le cadre d'études de chimie prébiotique<sup>211,212,213</sup>, il est assez improbable que le CDI ait pu constituer un agent activant prébiotique<sup>37</sup>. Nous l'avons utilisé essentiellement pour déterminer si le résultat final de la polymérisation était le même que dans le scénario « surfaces de silice ».

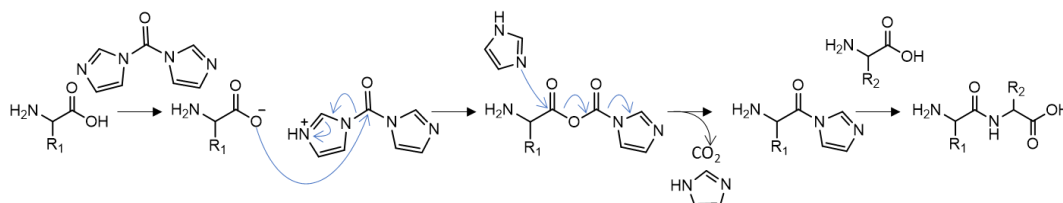


Figure III-6. Mécanisme réactionnel de la condensation peptidique avec CDI proposé par Ehler<sup>211</sup>.

Deux répétitions de cette condensation ont été opérées. Les produits ont ensuite été analysés par spectrométrie de masse à haute résolution sans purification. Le spectre de masse ESI des produits de condensation est reporté sur la Figure III-7, ci-dessous.

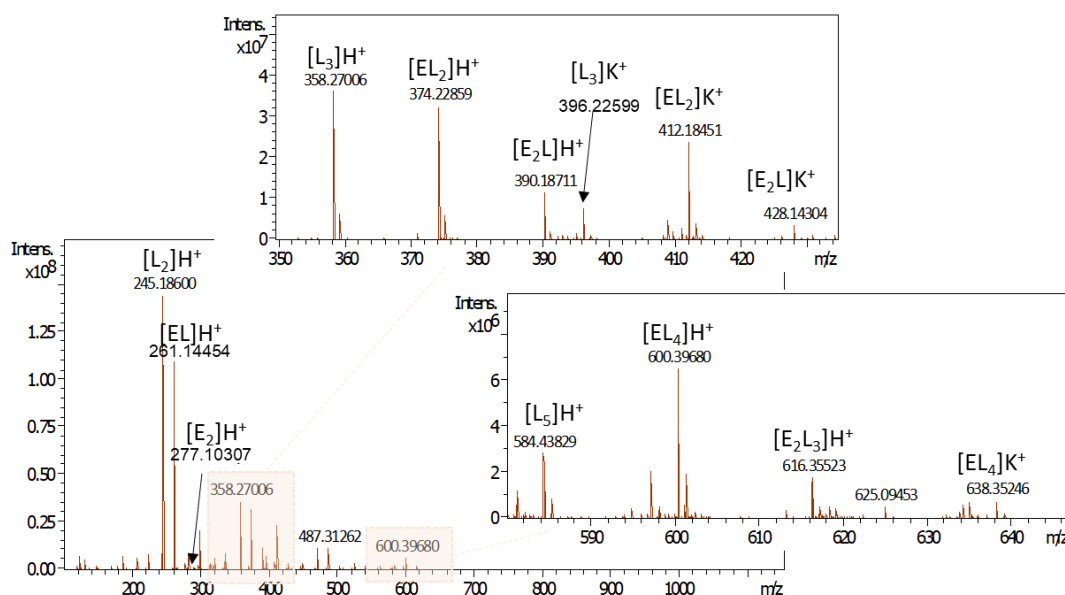


Figure III-7. Spectre de masse ESI en mode positif du produit de condensation obtenu en solution en utilisant le *N,N*-carbonyldiimidazole comme activateur sacrificiel avec des zoom sur les zones des tripeptides et pentapeptides.

Des oligopeptides allant jusqu'aux pentapeptides sont détectés, mais les résultats sont différents de ceux obtenus avec une polymérisation sur surface selon deux aspects. Premièrement, les espèces déshydratées sont minoritaires. Deuxièmement, on n'observe pas les mêmes tendances de sélectivité dans la réaction de polymérisation en solution par rapport à la polymérisation de surface. La Figure III-8 compare les répartitions en fonction de la stœchiométrie pour le cas des tripeptides, le plus précisément étudié. La différence la plus remarquable se situe au niveau des tripeptides L<sub>3</sub> : quasiment absents dans le scénario « surfaces de silice », ils sont au contraire plus abondants que prévu dans le scénario « polymérisation en solution ».

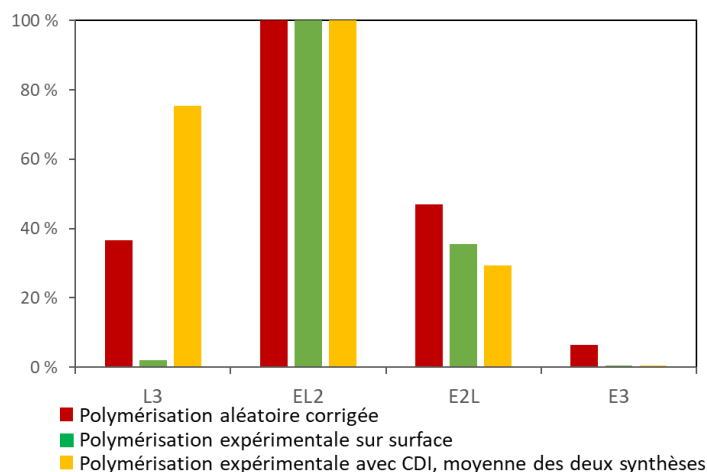


Figure III-8. Comparaison de la distribution des abondances expérimentales des tripeptides obtenues pour la polymérisation induite par carbonyldiimidazole (en jaune) avec celle obtenue pour surface de silice (en vert). La distribution de la polymérisation aléatoire corrigée est rappelée (en rouge). Toutes les distributions sont normalisées par rapport à EL<sub>2</sub>, pris comme référence à 100 %.

On peut donc dire que les scénarios de polymérisation n'ont pas convergé vers les mêmes distributions. Les sélectivités sont différentes, mais la polymérisation est-elle « plus sélective » dans le scénario de surface, ou dans la polymérisation en solution ? Pour donner un sens à cette question, il faudrait disposer d'une mesure unique de la sélectivité d'une distribution. La définition d'une telle mesure est l'objet du paragraphe suivant.

### II.3. L'entropie de Shannon pour évaluer la sélectivité de polymérisation

La définition de la sélectivité d'une réaction généralement utilisée en catalyse hétérogène/industrielle (quantité de molécules de départ transformées en produit d'intérêt) est assez restrictive et insuffisante pour évaluer l'intérêt prébiotique. Dans le cas d'un biopolymère fonctionnel dans une cellule (acide nucléique, mais aussi protéine), on peut dire que le système « contient de l'information », puisque la molécule consiste en une suite bien définie de monomères, assimilable à une suite de « lettres » bien définies. La construction biochimique d'une protéine est un processus parfaitement sélectif. La situation opposée est la répartition aléatoire des oligopeptides, « l'hypothèse zéro » que nous avons définie au Chapitre II.

Le scénario sur surface minérale est sélectif, dans le sens où certains produits de réaction sont favorisés et d'autres quasi absents. L'objectif, ici, est de quantifier cette sélectivité. Pour ce faire, nous avons besoin d'une mesure quantitative du degré auquel une distribution expérimentalement observée d'oligopeptides s'écarte de la répartition aléatoire.

Nous pouvons établir un parallèle entre cette question et un problème souvent traité en biochimie théorique qui est l'évaluation de l'hétérogénéité génétique parmi une population d'organismes. Par exemple, Adami<sup>214</sup> a quantifié la diversité génétique dans une population contenant des génomes différents. Mathématiquement, il s'agit du même problème que le nôtre si on remplace les nucléotides par des acides aminés, et que la longueur du polymère passe de quelques millions à quelques unités.

Dans le traitement d'Adami, X est un ensemble de séquences connues, dans son cas les chaînes d'ADN de longueurs L (la longueur de la séquence peut atteindre deux millions de résidus pour un simple gène). Le nombre de séquences possibles N dépend de la longueur des séquences et du nombre de monomères D (la taille de « l'alphabet »),  $N = D^L$ . Dans le cas des séquences d'ADN,  $D = 4$  (il y a 4 nucléobases).

Dans notre cas, X est un ensemble d'oligopeptides. Si on se limite à la population des tripeptides, la longueur de la séquence est donc  $L = 3$ , et la « taille de l'alphabet » est  $D = 2$  (les monomères peuvent être Leu ou Glu). On a donc  $N = 2^3 = 8$  séquences différentes. Dans le modèle aléatoire, elles sont toutes équiprobables,  $p_i = 1/8 = 0,125$ .

« L'entropie de Shannon »  $H^\ddagger$  est calculée avec l'équation (1) dans laquelle la somme est sur toutes les séquences possibles dans l'ensemble  $X$  et  $p_i$  correspond aux probabilités respectives.

$$H(X) = -\sum_{i=1}^N p_i \cdot \ln p_i \quad (1)$$

Dans l'hypothèse aléatoire, toutes les séquences sont équiprobables, et la probabilité de chacune est  $1/N$ , donc la relation (1) s'écrit :

$$H(X_{\text{aléatoire}}) = H_{\text{max}} = -\sum_{i=1}^N \frac{1}{N} \cdot \ln \left( \frac{1}{N} \right) = \ln N \quad (2)$$

C'est la valeur la plus élevée qu'on puisse atteindre pour une distribution d'oligopeptides. A l'opposé, si tous les oligopeptides sont identiques, il n'existe qu'une seule séquence de probabilité  $p = 1$ . Cette répartition correspond à une entropie de Shannon  $H_{\text{min}} = 0$ .

Il est important de noter que l'entropie de Shannon est un nombre sans dimension qui est différente de l'entropie molaire  $S^\circ$ , laquelle a les unités de  $\text{J} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$ .<sup>§</sup>

L'entropie de Shannon est la même que l'on prenne Leu, Glu, ou n'importe quel acide aminé. Il s'agit de quantifier l'information contenue dans une distribution de différentes séquences, ce qui ne dépend pas de la nature de l'acide aminé, mais seulement de la répartition.

Les génomes dans un environnement réel  $E$  sont soumis à des pressions de sélection où les probabilités de séquences  $p_i$  sont très différentes de l'hypothèse aléatoire. L'entropie conditionnelle d'un ensemble de séquences  $X$  dans un environnement  $E$  est très inférieure à  $H_{\text{max}}$ , et notée  $H(X|E)$ . La différence  $H_{\text{max}} - H(X|E)$  est notée  $I(X:E)$  et peut être considérée comme un contenu en information. Pour Adami, cette grandeur correspond à « l'information que  $X$  a sur  $E$  », donc l'information que possède une population de génomes sur son environnement.

On peut étendre ces concepts à notre système. Dans notre cas, on définit les  $p_i$  comme les abondances relatives (probabilités statistiques) des oligopeptides dans le mélange de

<sup>‡</sup> L'utilisation du symbole  $H$  pour désigner l'entropie de Shannon<sup>215</sup> est maintenant passée dans l'usage, bien qu'elle puisse être source de confusion puisqu'en thermodynamique  $H$  désigne habituellement l'enthalpie. A l'origine, l'entropie de Shannon a été développée dans le cadre d'une théorie de la transmission d'information.

<sup>§</sup> La relation entre les deux quantités est du même type que celle exprimée dans l'équation de Boltzmann  $S = k \ln \Omega$ .

désorption. Ainsi,  $I(X:SiO_2)$  est une évaluation quantitative du caractère non-aléatoire de la distribution d'oligopeptides observée dans le scénario de polymérisation sur silice, et  $I(X:CDI)$  la quantité correspondante pour la distribution observée dans le scénario en solution, en présence de CDI.

Dans le cas des tripeptides, les abondances relatives sont corrigées par les efficacités d'ionisation estimées expérimentalement comme expliqué plus haut. Pour cette longueur, les étapes de calcul de  $I(X:SiO_2)$  sont détaillées dans le Tableau III-1 et la quantité trouvée est de  $1,256 - 0,788 = 0,467$  unités d'entropie. Par ailleurs, pour la polymérisation utilisant l'activateur sacrificiel CDI,  $I(X:CDI)$  est de  $1,256 - 1,034 = 0,222$  unités d'entropie. Ce calcul est en accord avec la discussion du § II.2 : la distribution obtenue pour la polymérisation avec CDI se rapproche plus de la distribution aléatoire. La sélectivité est donc plus importante pour la polymérisation sur surface de silice.

*Tableau III-1. Calculs de l'entropie de Shannon pour les distributions de tripeptides expérimentales obtenues par condensation sur surface et par condensation par CDI, comparées à une distribution aléatoire. Le logarithme naturel a été utilisé, une conversion en base 2 peut être obtenue en multipliant par 1,4427.*

	Aléatoire			SiO <sub>2</sub>			CDI		
	$p_i$	$p_i \cdot \ln(p_i)$	$H_{max}$	$p_i$	$p_i \cdot \ln(p_i)$	$H$	$p_i$	$p_i \cdot \ln(p_i)$	$H$
<b>L<sub>3</sub></b>	0,125	-0,260	1,256	0,010	-0,045	0,788	0,300	-0,361	1,034
<b>EL<sub>2</sub></b>	0,375	-0,368		0,570	-0,321		0,495	-0,348	
<b>E<sub>2</sub>L</b>	0,375	-0,368		0,407	-0,366		0,205	-0,325	
<b>E<sub>3</sub></b>	0,125	-0,260		0,013	-0,057		0,000	-0,000	

Pour les oligopeptides de longueurs supérieures, nous avons effectué ces calculs sans facteur correctif. Les résultats sont repris au Tableau III-2. Pour le scénario de polymérisation avec CDI, on peut calculer l'entropie de Shannon jusqu'aux pentapeptides.

*Tableau III-2. Entropie de Shannon pour les distributions d'oligopeptides expérimentales, obtenues par condensation sur surface et par condensation par CDI, comparées avec l'entropie des distributions aléatoires correspondantes.*

Polypeptides	$H_{max}$	$H(X SiO_2)$	$\frac{H_{max} - H(X SiO_2)}{H_{max}}$	$H(X CDI)$	$\frac{H_{max} - H(X CDI)}{H_{max}}$
térapeptides	1,408	0,880	37 %	0,993	29 %
pentapeptides	1,524	0,535	65 %	0,584	62 %
hexapeptides	1,617	0,952	41 %		
heptapeptides	1,696	0,790	53 %		
octapeptides	1,763	0,612	65 %		



Pour toutes les longueurs de peptides, les valeurs d'enthalpie de Shannon de la distribution expérimentale sont différentes des valeurs aléatoires. Par ailleurs, les valeurs obtenues pour une polymérisation avec CDI sont plus proches de l'aléatoire. Cependant, le fait de ne pas prendre en compte les différences d'efficacité d'ionisation introduit une erreur dans le calcul.

On n'observe pas de tendance claire de l'évolution de l'enthalpie en fonction de la longueur des oligopeptides. On peut néanmoins affirmer que les oligopeptides les plus longs ne se rapprochent pas de la distribution aléatoire.

Le calcul des entropies selon le modèle probabilistique de Shannon nous permet donc de mettre sur une base quantitative l'estimation de la sélectivité de polymérisation dans le scénario sur surface de silice. Nous pouvons, par exemple, affirmer que la sélectivité est plus importante que celle du scénario avec activateur sacrificiel, traduisant l'implication de la surface et/ou de la température dans la réactivité des acides aminés. Ces paramètres ont été également étudiés en modifiant la température d'activation (voir Chapitre IV) et la surface utilisée (voir Chapitre V).

La présence d'une sélectivité de polymérisation au niveau de la stœchiométrie des oligopeptides a été mise en évidence. En outre, différentes séquences sont envisageables pour des peptides de même stœchiométrie ; un tripeptide de composition EL<sub>2</sub> peut ainsi avoir une séquence ELL, LEL ou LLE. D'un point de vue prébiotique, l'existence ou non d'une sélectivité dite de séquence est une question importante, car la fonction biologique d'un peptide dépend entre autres de sa structure, c'est-à-dire de l'agencement des résidus d'acides aminés. Les peptides de même stœchiométrie mais de séquences différentes n'étant pas distinguables en spectrométrie de masse dite à une dimension, l'étude de la sélectivité de séquence est reportée au Chapitre VI car elle nécessite l'utilisation d'autres techniques analytiques comme la spectrométrie de masse en mode tandem (MS/MS).

### III. Etude de la régiosélectivité de la polymérisation de (L+E)/SiO<sub>2</sub>

L'importance de l'acide aminé trifonctionnel dans les réactions de polymérisation a été démontrée dans nos travaux et au cours d'études précédentes<sup>51-53,55</sup> (voir Chapitre I). Mais l'existence d'une troisième fonction réactive dans la chaîne latérale pose la question de la régiosélectivité de la polymérisation. L'acide glutamique possède deux groupements acide carboxylique, qui peuvent tous deux se condenser et conduire à la formation d'une liaison amide. D'après le logiciel Molcalc, les charges partielles des groupements carboxyle en  $\alpha$  et en  $\gamma$  sont respectivement de  $\delta = +0,40$  et  $\delta = +0,41$ . De ce point de vue, ils auraient donc, à priori, la même probabilité d'être impliqués dans la formation d'une liaison amide. En revanche le carboxyle en  $\gamma$  est moins encombré. Une réactivité similaire des deux carboxyl aurait pour conséquence une augmentation considérable du nombre d'isomères possibles pour une séquence donnée (voir l'exemple illustré Figure III-9).

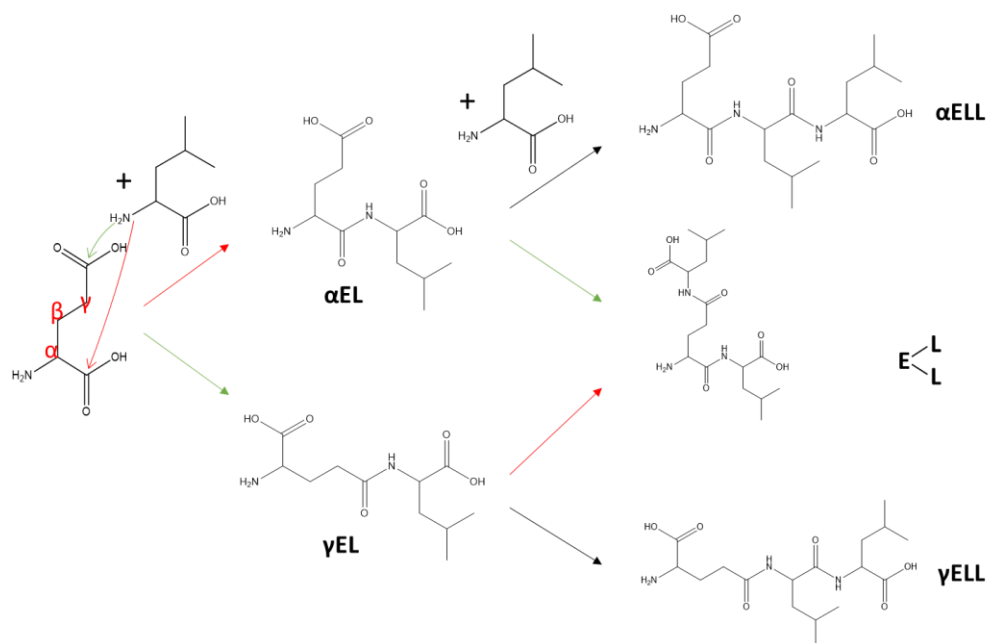


Figure III-9. Différents isomères possibles en tenant compte de la régiosélectivité pour une séquence ELL. Pour le tripeptide au milieu, qui peut être formé à partir de  $\alpha$ EL et de  $\gamma$ EL, le groupement NH<sub>2</sub> terminal est bien porté par le résidu Glu ; dans ce cas, les deux groupements carboxyl sont impliqués dans une liaison peptidique.

Or, rappelons que les  $\gamma$ -peptides (peptides dans lesquels le COOH en  $\gamma$  est impliqué dans la liaison peptidique) sont présents dans la nature (voir Chapitre I). Le glutathion en est un exemple, il est synthétisé de manière intracellulaire à partir d'acide glutamique, de cystéine et de glycine<sup>216</sup>. La question de la régiosélectivité est importante, entre autres parce que les peptides  $\alpha$ - et  $\gamma$ - formés n'ont pas le même comportement vis à vis d'une activation thermique. Comme évoqué au Chapitre I (§ III.2.2), Kasai<sup>191</sup> a montré que le chauffage du dipeptide  $\alpha$ Glu-Leu menait à la formation de pyroGlu-Leu alors que le chauffage de  $\gamma$ Glu-Leu aboutissait uniquement à la formation de pyroGlu.

D'autres acides aminés trifonctionnels peuvent réagir à différentes positions et ainsi poser la question de la régiosélectivité. C'est le cas par exemple, de la lysine avec ses deux groupements amine. Dans l'étude de l'oligomérisation d'un système lysine + acide diamino-butyrique, Frenkel-Pinter et al.<sup>217</sup> ont montré par RMN que la lysine était oligomérisée régiosélectivement *via* le groupe  $\alpha$ -amine, et gardait des  $\epsilon$ -amines libres.

Afin d'étudier la régiosélectivité lors de la condensation de notre échantillon (L+E)/SiO<sub>2</sub>, des condensations partant de dérivés d'acide glutamique bloqués par un groupement tertio-butyle (l'un avec le  $\alpha$ -COOH bloqué, l'autre le  $\gamma$ -COOH bloqué, voir Figure III-10) ont été étudiées. Les spectres de masse ESI en mode positif obtenus pour les échantillons (L+E $\alpha$ bloqué)/SiO<sub>2</sub> et (L+E $\gamma$ bloqué)/SiO<sub>2</sub> sont reportés dans la Figure III-10. Ils sont identiques aux spectres obtenus précédemment pour l'échantillon de référence (L+E)/SiO<sub>2</sub> (voir Figure III-3), à savoir que seuls des oligopeptides de type ExLy sont détectés.

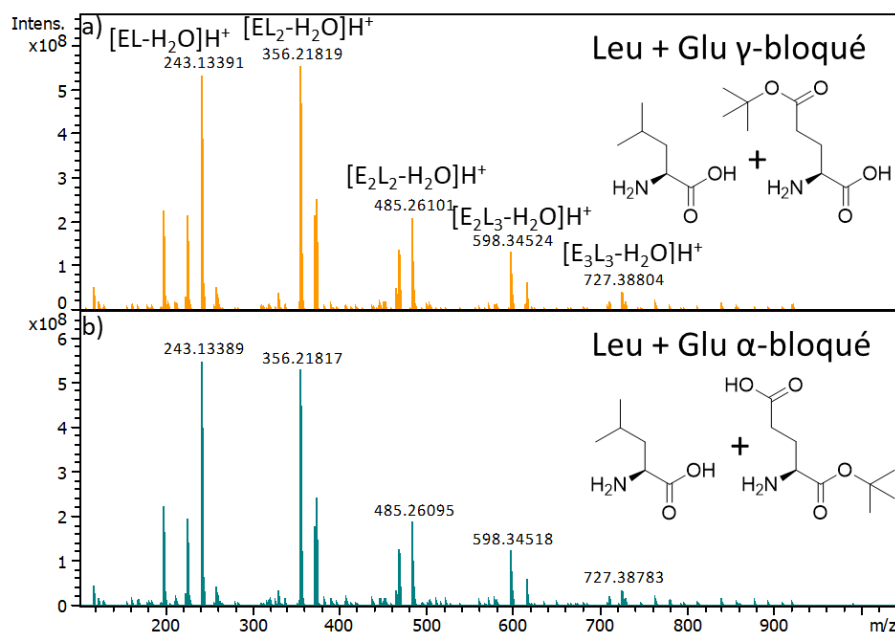


Figure III-10. Spectres de masse ESI en mode positif des échantillons a) (L+E) $\gamma$ -bloqué/SiO<sub>2</sub> et b) (L+E) $\alpha$ -bloqué/SiO<sub>2</sub> (instrument FTICR). Les structures des acides aminés bloqués utilisés sont reportées dans chaque spectre.

La perte inattendue du groupement protecteur tertibutyle et la production d'espèces de stœchiométries E<sub>x</sub>L<sub>y</sub> lors de l'oligomérisation peut avoir deux explications. La première est que la déprotection se fasse en amont de l'oligomérisation en raison d'un possible effet catalytique de la surface sur les réactions d'hydrolyse. Or une réaction d'hydrolyse paraît improbable dans nos conditions d'activation car l'eau initialement adsorbée sur la surface est éliminée à des températures supérieures à 100 °C. La seconde hypothèse est la perte du groupement protecteur lors des processus d'oligomérisation, par exemple par attaque nucléophile de l'amine sur le carbonyl portant le groupement tertibutyle. La préférence pour le groupement ester (vis-à-vis de la fonction acide) paraît contradictoire avec l'utilisation du tertibutyle comme groupement protecteur, même si la réactivité sur surface minérale peut être différente de la réaction en solution aqueuse. Une autre solution qui peut être proposée pour expliquer l'attaque préférentielle du groupement tertibutyle est qu'il serait le seul carbonyl disponible, le groupement COOH non bloqué pouvant être impliqué dans une liaison avec la surface.

Cette approche n'a donc pas permis d'étudier la régiosélectivité, en ce sens qu'il n'a pas été possible de confirmer ou d'infirmier la formation préférentielle de  $\alpha$ -peptides ou de  $\gamma$ -peptides. En outre, la spectrométrie de masse en tandem (MS/MS) opérée sur ces deux

échantillons ne montre pas de différences entre les voies de fragmentation pour une espèce donnée.

## IV. Etude de l'énantiosélectivité de la polymérisation de (L+E)/SiO<sub>2</sub>

Un dernier aspect important dans notre étude concerne l'énantiosélectivité. La Vie n'utilise que des acides aminés énantiomères L alors que les modèles prébiotiques connus comme les expériences de Miller produisent des mélanges racémiques<sup>25</sup>. Une question fondamentale en exobiologie est de savoir à quelle étape du scénario de l'origine de la vie la sélection d'un énantiomère dans les (bio)molécules est apparue. Nous avons décidé d'utiliser l'activation d'un mélange racémique d'acides aminés L et D pour étudier l'apparition d'une énantiosélectivité lors de la polymérisation. Afin de distinguer les stéréoisomères par spectrométrie de masse, l'un des énantiomères a été marqué par des atomes lourds, en l'occurrence par le <sup>13</sup>C.

Ainsi, quatre échantillons ont été préparés : (L<sub>L</sub>+(E<sub>L</sub>\*+E<sub>D</sub>))/SiO<sub>2</sub>, (L<sub>L</sub>+(E<sub>L</sub>+E<sub>D</sub>\*))/SiO<sub>2</sub>, (E<sub>L</sub>+(L<sub>L</sub>\*+L<sub>D</sub>))/SiO<sub>2</sub> et (E<sub>L</sub>+(L<sub>L</sub>+L<sub>D</sub>\*))/SiO<sub>2</sub>. Chaque échantillon est constitué d'un mélange entre un des deux acides aminés sous forme L (L-Leu ou L-Glu) et un mélange racémique de l'autre acide aminé, où l'un des deux énantiomères est marqué. Par exemple, l'échantillon (L<sub>L</sub>+(E<sub>L</sub>\*+E<sub>D</sub>))/SiO<sub>2</sub> contient 1.10<sup>-4</sup> mol de L-Leu avec 0,5.10<sup>-4</sup> mol de L-Glu et 0,5.10<sup>-4</sup> mol de D-Glu (voir Chapitre II).

Dans les quatre cas, les mêmes oligopeptides de stœchiométrie ExLy sont formés. Le marquage n'entraîne pas d'influence notable sur la taille maximum des oligopeptides formés ni sur les compositions des oligopeptides.

L'observation des massifs isotopiques correspondant à chaque oligopeptide permet d'étudier l'énantiosélectivité de la polymérisation. En effet, l'acide aminé L-Glu\* de l'échantillon (L<sub>L</sub>+(E<sub>L</sub>\*+E<sub>D</sub>))/SiO<sub>2</sub> étant marqué <sup>13</sup>C (le carbone 5 est un <sup>13</sup>C), tous les oligopeptides du mélange contenant des résidus L-Glu auront une masse augmentée de  $n \cdot 1,003355$  u (où  $n$  est le nombre d'énantiomères L) par rapport aux mêmes oligopeptides ne contenant que D-Glu.

On peut ici aussi comparer avec une polymérisation aléatoire correspondant à l'absence d'énantio-sélectivité. Par exemple, pour un térapeptide de formule (E<sub>2</sub>L<sub>2</sub>-H<sub>2</sub>O), les intensités des pics de peptides contenant 0, 1 et 2 résidu(s) L-Glu devraient dans l'hypothèse zéro être dans un rapport 1 : 2 : 1. La comparaison avec la répartition expérimentale permet donc de déterminer une possible énantiosélectivité lors de la polymérisation.

La Figure III-11 montre l'aspect des massifs isotopiques pour les térapeptides [E<sub>2</sub>L<sub>2</sub>-H<sub>2</sub>O]H<sup>+</sup>.

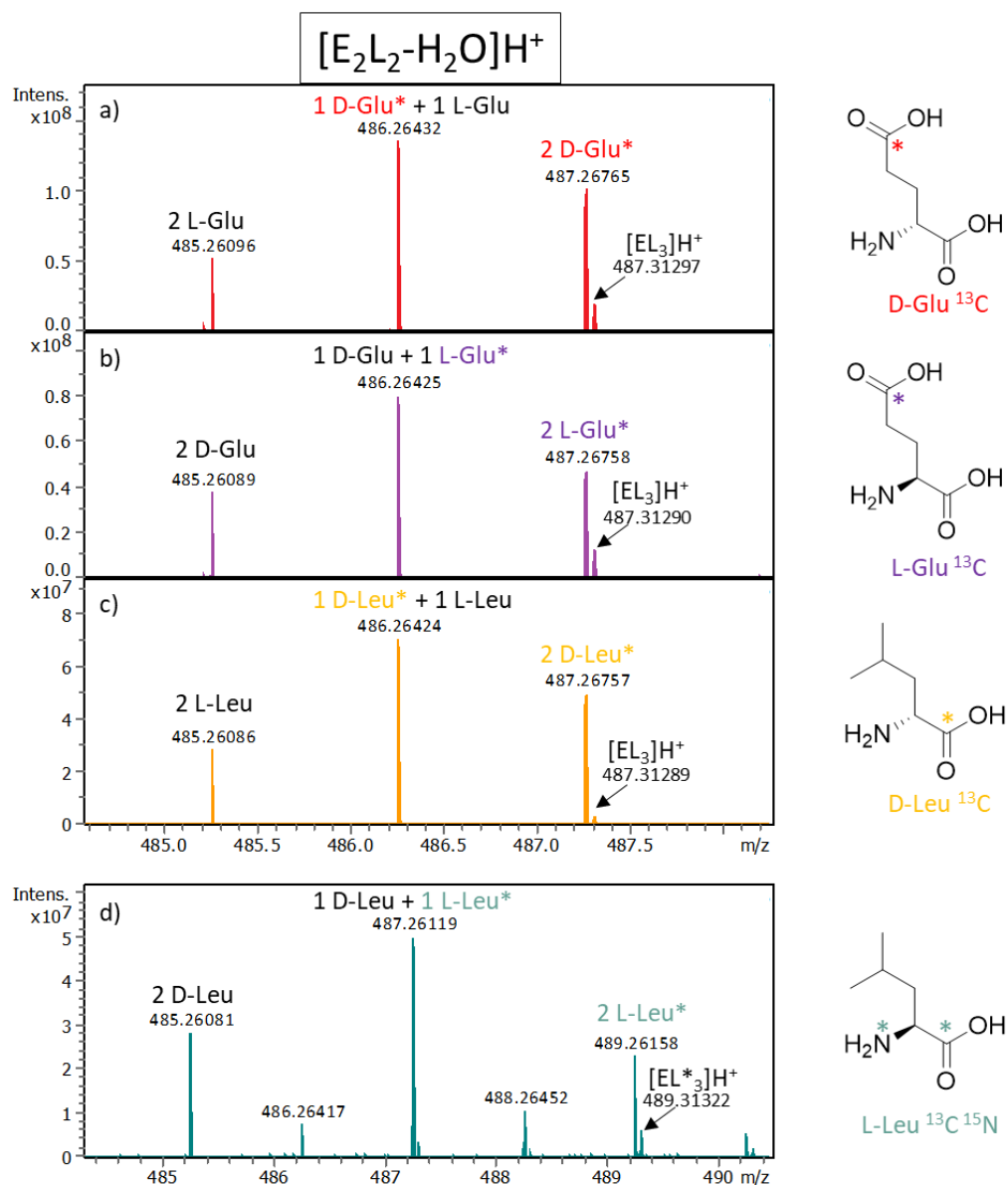


Figure III-11 : Comparaison des massifs isotopiques du tétramère [E<sub>2</sub>L<sub>2</sub>-H<sub>2</sub>O]H<sup>+</sup> (m/z 485.2608) pour les échantillons a) (L<sub>L</sub> + (E<sub>L</sub>+E<sub>D</sub>\*))/SiO<sub>2</sub>, b) (L<sub>L</sub>+(E<sub>L</sub>\*+E<sub>D</sub>))/SiO<sub>2</sub>, c) (E<sub>L</sub>+(L<sub>L</sub>+L<sub>D</sub>\*))/SiO<sub>2</sub> et d) (E<sub>L</sub>+(L<sub>L</sub>\*+L<sub>D</sub>))/SiO<sub>2</sub>. Théoriquement les intensités relatives de ce massif devraient être de 1 : 2 : 1.

La formation de différents diastéréoisomères (2 D-Glu, 1 D-Glu+1 L-Glu ou 2 L-Glu) est observée sur le spectre de masse Figure III-11a. Les intensités semblent tout d'abord différer légèrement de la distribution aléatoire 1 :2 :1 : le pic correspondant au rapport m/z le plus élevé a une intensité supérieure que celui à m/z plus faible. Mais si l'on interprète ce résultat comme indiquant une préférence des tétrapeptides pour l'isomère D-Glu, on aboutit à une incohérence en comparant la Figure III-11a et la Figure III-11b. En effet, la seconde correspond au système où D-Glu est non marqué, donc plus léger. On devrait donc, si l'interprétation était correcte, observer cette fois une intensité supérieure pour le pic de plus faible m/z. Or, au contraire, on observe à nouveau que c'est le pic de masse élevée qui a la plus forte intensité. C'est aussi le cas pour le spectre 11c, où l'entité marquée est cette fois le D-Leu. Il faut se souvenir toutefois que ces différents pics ne sont séparés que par une unité de masse atomique. Or, la quantité naturellement présente de l'isotope <sup>13</sup>C est suffisante pour contribuer aux pics de plus grands m/z (m+1 et m+2), on parle alors de massif isotopique.

Le spectre de masse d correspondant à l'étude de l'échantillon comportant L-Leu\* permet de vérifier ces hypothèses. En effet, l'acide aminé L-Leu est doublement marqué : sur <sup>13</sup>C et sur <sup>15</sup>N, les pics sont donc décalés de 2,0004 u pour chaque unité L-Leu. Le massif isotopique naturel (pics à 486,26417 et 488,26452 u) ne se recouvre donc pas avec les pics isotopiques et les deux contributions sont séparées. On peut donc affirmer que le léger biais par rapport à la théorie est dû au marquage et non à la configuration L ou D des acides aminés de départ. La comparaison des quatre échantillons est bien nécessaire pour s'affranchir de possibles effets isotopiques (l'acide aminé marqué, plus lourd, peut avoir une réactivité différente). Ces mêmes observations ont été faites pour tous les oligopeptides détectés dans le mélange.

Au vu de ces résultats expérimentaux, la polymérisation sur surface ne semble pas induire d'énantiosélectivité significative.

En conclusion, cette étude préliminaire de l'échantillon (L+E)/SiO<sub>2</sub> permet de tirer plusieurs conclusions intéressantes sur la polymérisation des acides aminés Leu et Glu soumis à une activation thermique. En premier lieu, la formation de liaisons peptidiques menant à des oligopeptides de composition ExLy a été confirmée par différentes méthodes analytiques. Le scénario prébiotique impliquant une polymérisation sur surface minérale grâce à une source de chaleur est donc parfaitement plausible (voir Chapitre I). Il ne se limite pas à la formation de dipeptides cycliques, mais aboutit à la production de peptides linéaires relativement longs.

Les espèces produites démontrent une certaine chimiosélectivité au niveau de la stœchiométrie des oligopeptides formés. Ainsi, la formation d'hétéropeptides apparaît comme privilégiée par rapport à une distribution aléatoire.

Il est plus difficile de statuer pour d'autres types de sélectivité. Au-delà de la stœchiométrie, l'identification de séquences d'acides aminés préférentielles nécessite des techniques de caractérisation avancée comme la spectrométrie de masse en tandem (MS/MS) éventuellement en couplage avec des techniques séparatives ; ce point sera étudié au Chapitre VI.

Il n'a pas été possible de confirmer ou d'infirmer une régiosélectivité au niveau de la polymérisation de l'échantillon (L+E)/SiO<sub>2</sub> en utilisant des acides aminés modifiés contenant des groupements acides bloqués. Pour étudier cette notion de régiosélectivité, d'autres approches analytiques comme les méthodes en couplage LC/MS et IM-MS ainsi que des expériences MS/MS ont été, par la suite, appliquées aux différents systèmes étudiés (voir Chapitres V et VI). Par ailleurs, aucune énantiosélectivité n'est observée pour ce système, dans nos conditions de polymérisation.

Une observation expérimentale d'importance est la formation d'espèces déshydratées majoritaires. Leurs origines et leurs rôles dans la polymérisation posent question. Sont-elles des produits intermédiaires de la polymérisation, ou à l'inverse, des produits de dégradation ? Dans cette optique, une étude cinétique visant à révéler les étapes mécanistiques de la polymérisation a été réalisée. Par ailleurs, une étude structurale a été opérée par MS/MS pour déterminer une sélectivité au niveau de la séquence peptidique. Ces études seront discutées dans les chapitres suivants.





Chapitre IV. Etude du mécanisme de  
polymérisation : changement des  
conditions opératoires

La formation de polypeptides linéaires dans notre système de référence ayant été établie au Chapitre III, le présent chapitre regroupe des études complémentaires de ce système destinées à mieux comprendre son mécanisme, afin de pouvoir évaluer le potentiel prébiotique de la condensation d'acides aminés déposés sur surface et ses limites.

## I. Etude cinétique de la polymérisation

Une question primordiale dans cette étude concerne le mécanisme réactionnel de la réaction de condensation. Dans cette optique, une étude cinétique a été entreprise et pour ce faire, la polymérisation du système Leu + Glu a été opérée à différentes températures et/ou à des temps variables. Nous cherchions avant tout à comprendre dans quel ordre les espèces polymériques apparaissent et ainsi à mettre en évidence les étapes du mécanisme réactionnel.

### I.1. Activation thermique de (L+E)/SiO<sub>2</sub> à différentes températures et différents temps de réaction

Les contraintes imposées par le « scénario » prébiotique sélectionné (dépôt à partir d'une solution aqueuse, séchage puis activation thermique) rendent difficile une étude cinétique classique. En effet, on ne peut pas amener le réacteur à la température voulue, puis introduire les réactifs seulement à ce moment (chose qui serait possible si on utilisait un dépôt en phase gazeuse ou « CVD », voir Chapitre I). Dans nos conditions, le système atteint progressivement la température cible, et a déjà commencé à réagir lorsque celle-ci est atteinte.

Dans un premier temps, nous avons activé le mélange (L+E)/SiO<sub>2</sub> à des températures variables, et notamment des températures inférieures à 160 °C. Cette étude est également pertinente par rapport à l'intérêt prébiotique du scénario, car des températures aussi élevées que 160° C peuvent être jugées peu plausibles, ou bien seulement dans des circonstances exceptionnelles (par exemple à proximité d'un volcan).

Nous avons ensuite considéré l'effet d'un plateau de durée variable après chauffage jusqu'à une même température cible. Les huit différentes activations réalisées sont regroupées dans le Tableau IV-1. Pour chacune de ces huit activations, deux échantillons ont été préparés pour vérifier la répétabilité de la réaction. Tous les échantillons ont été analysés le même jour par FTICR avec une injection automatique (mode FIA, voir Chapitre II).

*Tableau IV-1. Plan d'expériences mis en œuvre pour l'étude de la cinétique de polymérisation. En bleu, expériences à température finale variable ; en orange, expérience avec plateaux de température de durées variables.*

Température cible	Durée du plateau (en min)	Durée de l'activation $T_i = 25^\circ\text{C}$
100	0	1h15
115	0	1h30
130	0	1h45
145	0	2h
160	0	2h15
115	30	2h
115	120	3h30
115	420	5h30

On observe que les oligopeptides contenant plus de trois résidus n'apparaissent qu'à une température supérieure à  $130^\circ\text{C}$  : en-dessous de cette température, seuls les acides aminés non polymérisés et des di- et tripeptides sont détectés.

Dans le but de mettre en évidence des tendances semi-quantitatives, les abondances relatives des différents oligopeptides détectés ont été calculées. Pour une longueur d'oligopeptide donnée, les intensités des pics des différentes espèces (protonées ou cationisées, intactes ou déshydratées) ont été sommées. Ensuite, la valeur obtenue pour chaque oligopeptide a été normalisée par rapport à la somme des intensités de toutes les espèces. Il est important de souligner que les valeurs ainsi calculées ne correspondent pas à des concentrations car les facteurs de réponse (efficacités d'ionisation) en spectrométrie de masse ne sont pas connus et sont difficilement prédictibles sans composés standards. Nous avons vu qu'ils peuvent différer considérablement entre différentes séquences (Chapitre III, § II.1).\*\*

La Figure IV-1 montre l'évolution des oligopeptides de longueurs différentes, en fonction de la température d'activation du système. Les signaux des monomères diminuent d'intensité

\*\* Toutefois, il semble qu'il y ait une certaine compensation des erreurs induites par la non-considération des facteurs de réponse, comme le montrent les calculs rapportés en Annexe 3.

lorsque la température d'activation augmente. Il ne reste quasiment plus de monomères à 160°C. En ce qui concerne les dipeptides, on observe une augmentation de l'abondance relative jusqu'à 130 °C suivie d'une légère diminution. Ce type d'évolution suggère que les dipeptides sont des intermédiaires de réaction. Quant aux espèces plus longues, elles sont absentes au début de la polymérisation et leur intensité augmente avec la température. Les tripeptides apparaissent d'abord à partir de 115 °C, puis les espèces plus longues, suggérant une formation séquentielle des oligopeptides. Les peptides les plus courts sont donc formés à plus faible température et doivent servir de « briques » pour former les plus longs.

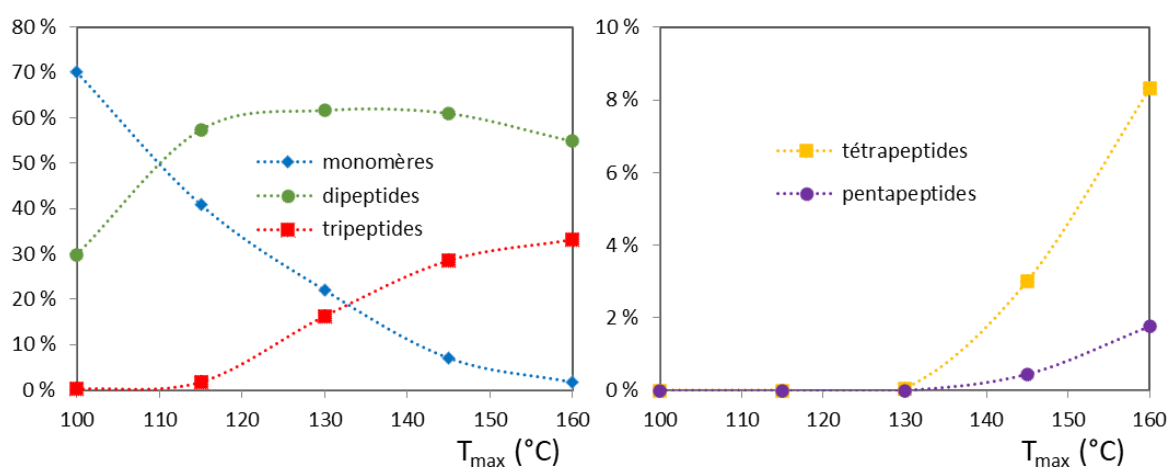


Figure IV-1. Evolution des oligopeptides de différentes tailles en fonction de la température maximum d'activation thermique. L'abondance relative est portée en ordonnée. Des hexapeptides et heptapeptides sont détectés avec de très faibles abondances à 160 °C (non représenté). Les courbes en pointillé sont de simples guides optiques.

On s'est également intéressé à la formation des espèces déshydratées. Par exemple pour le tripeptide de stœchiométrie E<sub>2</sub>L, la forme intacte [E<sub>2</sub>L]H<sup>+</sup> est observée à partir de 115 °C, puis augmente progressivement tandis que la forme déshydratée [E<sub>2</sub>L-H<sub>2</sub>O]H<sup>+</sup> apparaît seulement après 130 °C (Figure IV-2). Cette tendance est vérifiée pour tous les oligopeptides, indépendamment de leur taille. Les intensités de toutes les espèces déshydratées ont été sommées puis normalisées par rapport à l'ensemble des oligopeptides détectés comme précédemment. Les abondances relatives en fonction de la température d'activation sont reportées dans la Figure IV-3. Elles augmentent de façon quasi linéaire avec la température ; ainsi les espèces déshydratées deviennent majoritaires pour une température supérieure à 130 °C.

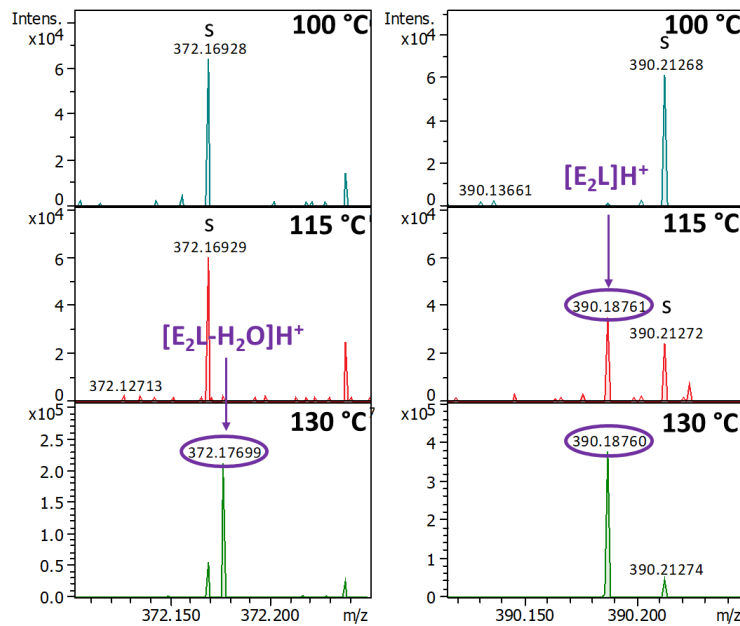


Figure IV-2. Zoom autour des ions  $[E_2L-H_2O]H^+$  et  $[E_2L]H^+$  des spectres de masse ESI obtenus pour les échantillons (L+E)/SiO<sub>2</sub> activés à différentes températures. Les pics annotés s correspondent à des contaminants du solvant.

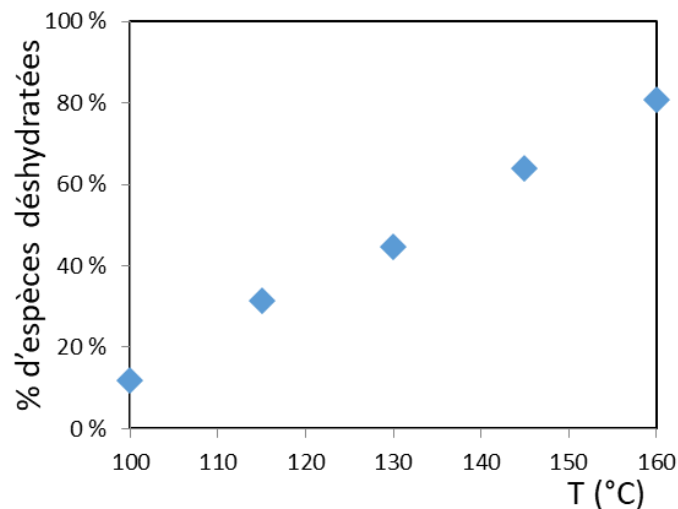


Figure IV-3. Evolution de l'abondance relative des espèces déshydratées en fonction de la température d'activation maximum.

On doit se poser la question de savoir si ces formes déshydratées sont perdues pour l'évolution biochimique, ce qui revient à se demander si la déshydratation est réversible. Nous tenterons donc au Chapitre VI de déterminer plus précisément la nature de ces formes ; elles peuvent résulter d'une cyclisation globale du peptide et/ou de cyclisations internes impliquant la fonction acide carboxylique des chaînes latérales des résidus Glu.

Nous verrons au Chapitre V que des réactions de déshydratation se produisent également pour d'autres systèmes testés.

Le suivi en cinétique isotherme à 115 °C est détaillé en Figure IV-4. Sachant que la polymérisation a déjà commencé lorsque la température cible est atteinte, l'évolution ultérieure en fonction du temps montre des tendances similaires à celles qui sont observées en fonction des températures finales croissantes (Figure IV-1). Les dipeptides ont déjà atteint un plateau ; leur évolution ne montre pas de tendance claire. Les monomères disparaissent tandis que les tripeptides augmentent ; aux temps les plus élevés, de faibles quantités de tétrapeptides et pentapeptides sont observées. Les oligopeptides plus longs ne sont pas détectés.

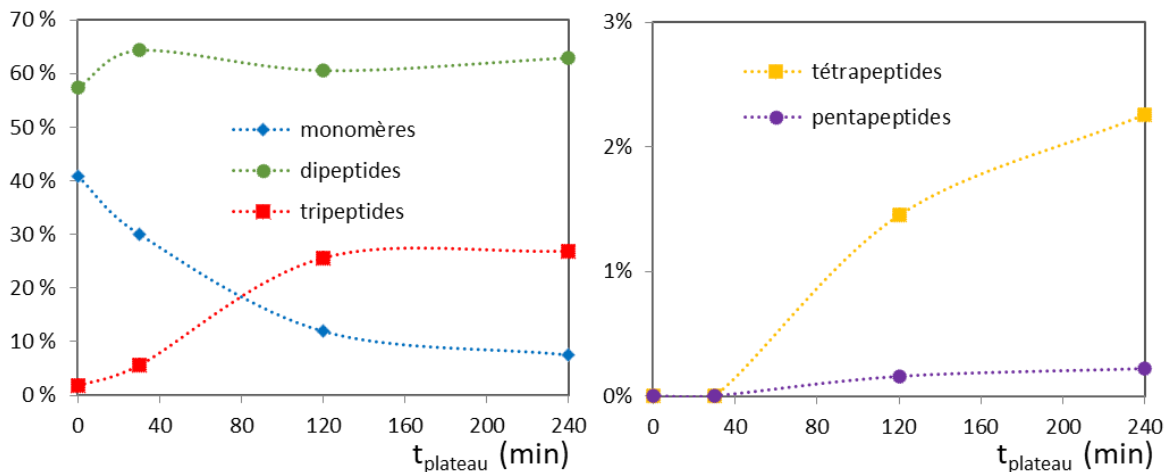


Figure IV-4. Evolution de la taille des oligopeptides en fonction du temps de plateau à 115 °C. Les courbes en pointillé sont des guides optiques.

Cette étude cinétique montre à nouveau l'ordre séquentiel d'apparition des peptides. Les tendances observées en fonction de la température de chauffage sont les mêmes qu'en fonction du temps de réaction en isotherme. Ainsi, il ne reste presque plus de monomères libres à 160 °C, ou après un plateau de quatre heures à 115 °C. Les résultats obtenus à température élevée peuvent essentiellement être reproduits à température plus basse, à condition d'augmenter le temps de réaction. On peut « échanger le temps contre la température », comme le suggère une interprétation intuitive de la loi d'Arrhénius. Le modèle le plus simple de polymérisation consisterait en une croissance de la chaîne peptidique par agrégation d'un monomère à la fois, ce qui signifie que la croissance devrait

s'arrêter lorsque les monomères libres sont épuisés. Des condensations entre deux oligopeptides préalablement formés pourraient également être envisagées ; cette question sera explorée plus bas.

Pour illustrer ces observations préliminaires, on définit l'indice de polymérisation comme la somme des abondances relatives pondérées par la longueur de l'oligopeptide :  $\sum_i i \cdot x_i$ , avec  $i$  la longueur de l'oligopeptide et  $x_i$  l'abondance relative correspondante. L'indice de polymérisation permet d'avoir une idée des longueurs moyennes obtenues pour chaque échantillon et/ou expérience. La Figure IV-5 montre l'évolution des valeurs de cet indice pour les différentes activations réalisées et confirme un avancement de la polymérisation avec la température et/ou la durée de réaction. Par ailleurs, en parallèle avec la polymérisation, des espèces déshydratées se forment à partir des oligopeptides intacts et deviennent largement majoritaires dans les conditions d'avancement le plus élevé, tout comme nous l'avons observé dans l'étude à température variable.

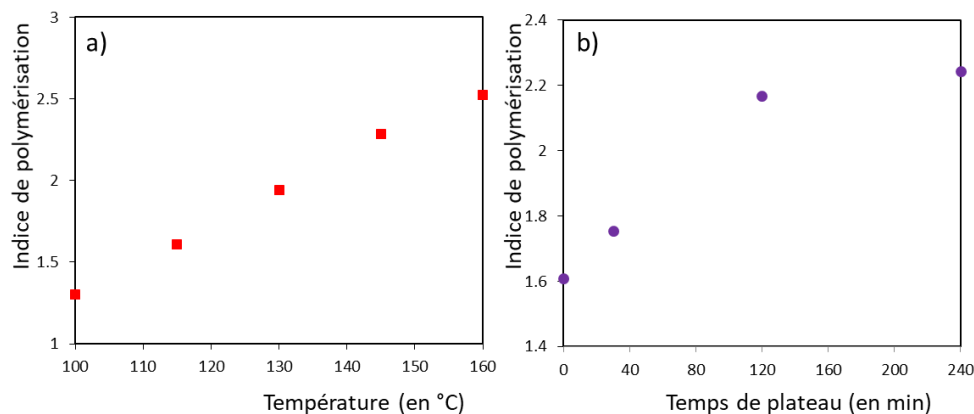


Figure IV-5. Evolution des indices de polymérisation en fonction a) de la température maximale atteinte ; b) du temps de plateau appliqué pour une activation à 115 °C.

Enfin, dans le but d'évaluer le potentiel prébiotique d'un scénario d'activation thermique simple, une activation de longue durée avec un plateau de 24 h à la température de 160 °C, a été réalisée. La Figure IV-6 compare la distribution en taille des peptides obtenus pour cet échantillon avec un plateau de 24 h à celle des peptides obtenus pour l'échantillon de référence avec un plateau de 30 min. Les analyses ont été faites le même jour sur le FTICR en mode FIA (voir Chapitre II).



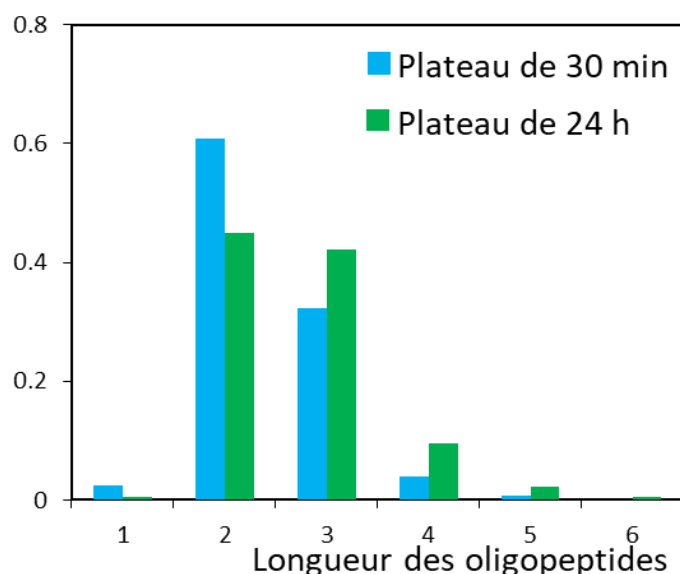


Figure IV-6. Comparaison de l'abondance relative des différentes longueurs d'oligopeptides pour deux activations thermiques à 160 °C, avec un plateau de 30 min et de 24 h.

Les longueurs maximales des oligopeptides détectés ne sont pas modifiées, en revanche un décalage de la distribution vers les oligopeptides les plus longs est observé. Les indices de polymérisation sont respectivement de 2,40 et de 2,69 pour l'échantillon avec plateau de 30 min et celui avec un plateau de 24 h. Il sera montré plus loin (cf § IV.1) que l'application de cycles hydratation/déshydratation est plus efficace pour augmenter la longueur des peptides.

## I.2. Etude de la polymérisation à partir de dipeptides

L'étude cinétique a montré la formation initiale de dipeptides, puis d'oligopeptides de plus en plus longs. Pour observer plus clairement la seconde étape de cette réactivité, des synthèses à partir non pas de deux acides aminés mais d'un acide aminé et d'un dipeptide ont été réalisées. En outre, pour évaluer la réactivité potentielle des peptides déshydratés, des activations (dipeptide linéaire + monomère) et (dipeptide cyclique + monomère) ont été comparées. Les dipeptides cycliques utilisés sont des formes DKP (cf. Chapitre I, § I.3.2). Les six échantillons étudiés sont les suivants : (LE+L)/SiO<sub>2</sub>, (LE+E)/SiO<sub>2</sub>, (EE+L)/SiO<sub>2</sub>, (AA+E)/SiO<sub>2</sub>, (cycloEE+L)/SiO<sub>2</sub> et (cycloAA+E)/SiO<sub>2</sub>. Chaque échantillon a été activé à deux températures

différentes, 130 °C et 160 °C, pour tester des degrés d'avancement différents de la réaction. Les analyses ont été faites par FTICR en mode FIA (Chapitre II).

### I.2.1. Polymérisation à partir de dipeptides linéaires

La polymérisation des échantillons (EE+L)/SiO<sub>2</sub>, (LE+L)/SiO<sub>2</sub> et (LE+E)/SiO<sub>2</sub> a été étudiée à 130 °C et 160 °C. La Figure IV-7 montre les spectres de masse des activations à 130 °C (ainsi que l'activation à 160 °C de l'échantillon (LE+E)/SiO<sub>2</sub>) car à cette température, les espèces majoritaires sont les dipeptides et les monomères résiduels, même si une quantité significative de tripeptides est déjà présente. En activant à 130 °C un monomère et un dipeptide, on souhaite ainsi observer de manière contrôlée les étapes ultérieures de polymérisation.

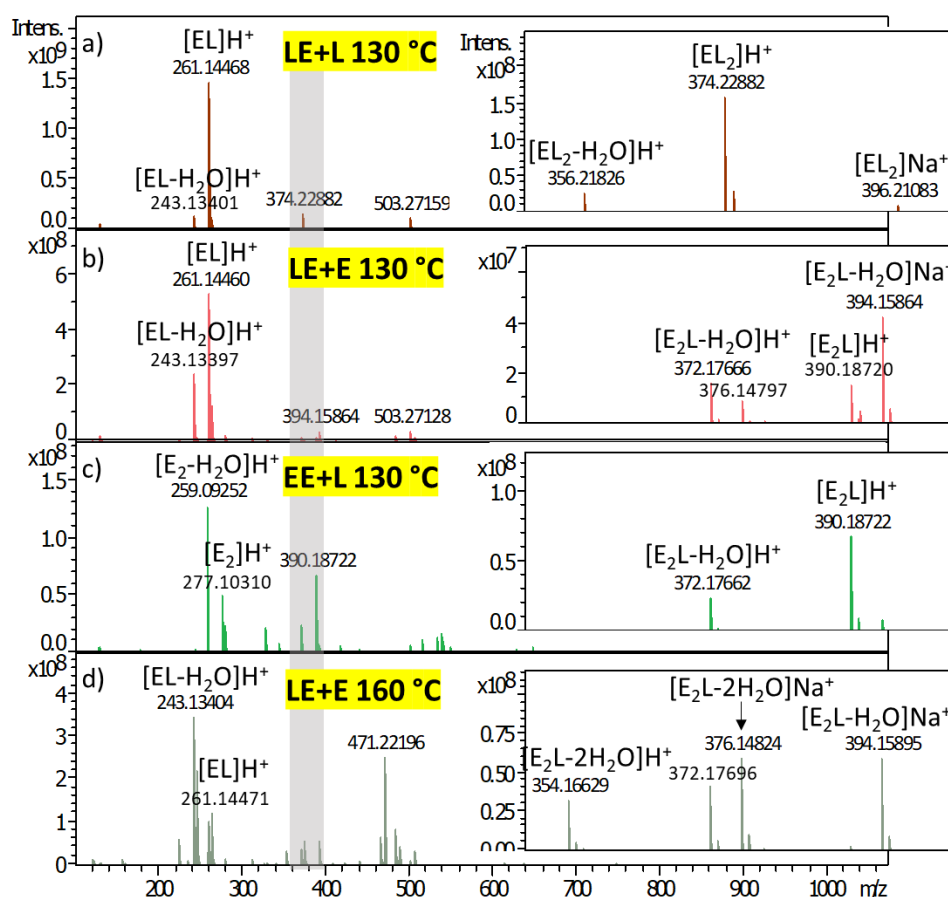


Figure IV-7. Spectres de masse obtenus pour la polymérisation des échantillons a) (LE+L)/SiO<sub>2</sub> 130 °C, b) (LE+E)/SiO<sub>2</sub> 130°C, c) (EE+L)/SiO<sub>2</sub> 130 °C et d) (LE+E)/SiO<sub>2</sub> 160 °C (ESI positif, instrument FTICR). Un zoom sur la zone des tripeptides (grisée) est montré pour chaque spectre de masse.

Pour les trois échantillons activés à 130 °C, des oligopeptides allant jusqu'aux heptapeptides sont détectés, alors que pour l'échantillon (L+E)/SiO<sub>2</sub> activé à 130 °C, la polymérisation dépassait à peine le stade des tripeptides, traduisant la capacité des dipeptides de se condenser efficacement avec des monomères. Néanmoins, les produits principaux sont des tripeptides, et essentiellement ceux qui résultent d'une simple condensation dipeptide/monomère, c'est à dire des tripeptides de stœchiométrie EL<sub>2</sub> à partir de (LE+L)/SiO<sub>2</sub>, et E<sub>2</sub>L aussi bien à partir de (LE+E)/SiO<sub>2</sub> que de (EE+L)/SiO<sub>2</sub>. Une étude plus détaillée de ces deux derniers échantillons, qu'il est possible de comparer car ils contiennent exactement la même « quantité » de chaque acide aminé, indique que les condensations ne montrent pas la même efficacité. En effet, le pic correspondant au tripeptide E<sub>2</sub>L représente 27 % de l'intensité globale des oligopeptides dans le cas de l'échantillon (EE+L)/SiO<sub>2</sub> et seulement 7 % pour (LE+E)/SiO<sub>2</sub>.

En revanche, la comparaison des échantillons (LE+L)/SiO<sub>2</sub> et (LE+E)/SiO<sub>2</sub> indique une réactivité équivalente des deux monomères par rapport au dipeptide de séquence LE. D'une part, les deux systèmes mènent à la formation d'oligopeptides de même longueur ; d'autre part, les pourcentages des pics du tripeptide correspondant à la condensation dipeptide/monomère ont des valeurs similaires (respectivement 9 % et 7 % de l'intensité globale pour EL<sub>2</sub> et E<sub>2</sub>L).

De façon générale, pour une activation à 130 °C, les réarrangements interchaînes (c'est-à-dire des ruptures de liaisons peptidiques suivies de recombinaisons aboutissant à des séquences différentes), théoriquement envisageables, sont quasi inexistantes en pratique. En effet, concernant les tripeptides, le rapport E<sub>2</sub>L/EL<sub>2</sub> des intensités des espèces de stœchiométrie E<sub>2</sub>L et EL<sub>2</sub> vaut respectivement 0,03 et 596 pour les échantillons (LE+L)/SiO<sub>2</sub> et (LE+E)/SiO<sub>2</sub> activés à 130 °C alors qu'il est de 0,76 pour la polymérisation à partir des monomères. Concernant les oligopeptides de longueurs égale et supérieure aux tripeptides, la plupart d'entre eux sont issus d'additions soit des monomères, soit des dipeptides de départ à des espèces formées antérieurement. Ainsi, dans l'échantillon (EE+L)/SiO<sub>2</sub>, 98 % des oligopeptides présents peuvent s'expliquer de cette façon (en l'occurrence, ce sont simplement ceux qui ont un nombre pair de E). Pour les échantillons (LE+L)/SiO<sub>2</sub> et (LE+E)/SiO<sub>2</sub>, ce pourcentage est respectivement de 99 % et de 97 %.

En ce qui concerne la déshydratation, il apparaît une différence de comportement entre l'échantillon (EE+L)/SiO<sub>2</sub> et ceux comportant le dipeptide LE : le pic majoritaire des

dipeptides correspond à l'espèce déshydratée  $[E_2-H_2O]H^+$  dans le premier cas tandis qu'il correspond au dipeptide intact  $[EL]H^+$  dans les deux autres. La déshydratation est donc corrélée à la présence de Glu, sûrement en raison de la capacité que possède Glu de faire une cyclisation interne (pour former du pyroGlu ou d'autres cycles). En revanche, la tendance est plus complexe au niveau des tripeptides (voir zoom de la Figure IV-7). En effet, le rapport intact/déshydraté pour la stœchiométrie  $E_2L$  varie en fonction de l'échantillon étudié ( $(LE+E)/SiO_2$  ou  $(EE+L)/SiO_2$ ). Il est possible que la (les) séquence(s) présente(s) varient dans les deux échantillons, car nous avons seulement accès à la stœchiométrie des oligopeptides. Considérant uniquement des processus de condensation entre les réactifs mis en présence, l'échantillon  $(LE+E)/SiO_2$  devrait conduire pour les tripeptides aux séquences LEE et/ou ELE suivant que la condensation a lieu par l'extrémité C-terminale ou N-terminale du dipeptide, tandis que l'échantillon  $(EE+L)/SiO_2$  devrait donner les séquences EEL et/ou LEE.

Activés à 160 °C, les trois échantillons (dipeptide linéaire + monomère) montrent des polymérisations équivalentes : le dipeptide précurseur est principalement présent sous sa forme déshydratée et des oligopeptides allant jusqu'aux octapeptides sont détectés, traduisant une progression de la polymérisation. Pour comparaison, l'espèce la plus longue observée pour l'échantillon  $(L+E)/SiO_2$  dans les mêmes conditions d'analyse était un hexapeptide (voir Figure IV-11).

La Figure IV-8a compare les distributions des oligopeptides de l'échantillon  $(EE+L)/SiO_2$  pour les activations à 130 et 160 °C avec un indice de polymérisation augmentant de 2,66 à 3,21. Si l'on considère de plus près la stœchiométrie des oligopeptides formés, on constate que ceux qui s'expliquent par l'addition de dipeptides et/ou de monomères sont toujours majoritaires à 160 °C, mais de façon non exclusive. Pour  $(EE+L)/SiO_2$ , ils représentent ainsi 84,2 % des oligopeptides totaux contre 98 % à 130 °C (voir Annexe 4). Il est donc probable que les réarrangements de chaînes, inexistant à 130 °C, commencent à se produire à 160 °C.

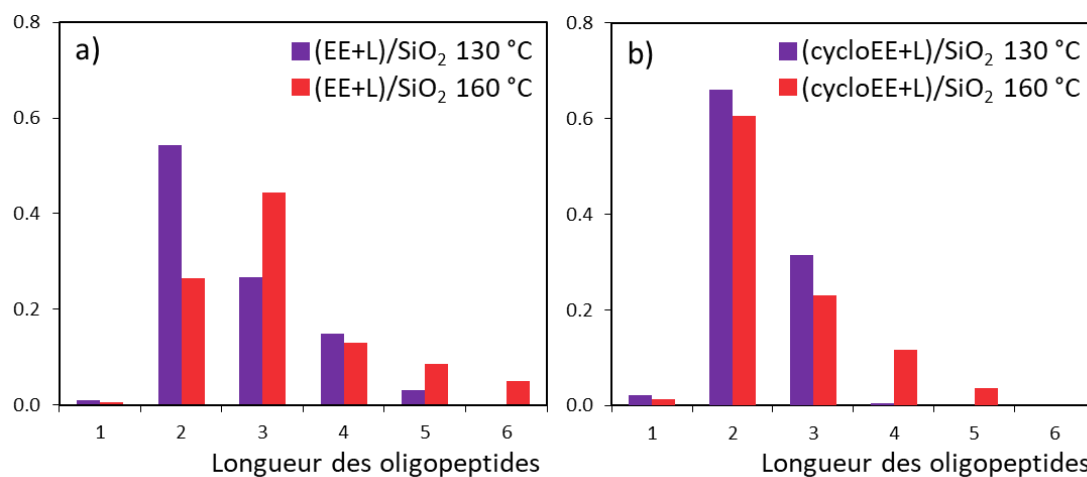


Figure IV-8. Distribution des abondances relatives des oligopeptides en fonction de leur longueur pour les échantillons a) (EE+L)/SiO<sub>2</sub> activés à 130 °C et 160 °C et b) (cycloEE+L)/SiO<sub>2</sub> activés à 130 °C et 160 °C.

### I.2.2. Comparaison de l'efficacité de polymérisation des dipeptides cycliques par rapport aux dipeptides linéaires

Nous nous sommes intéressés à la polymérisation à partir de dipeptides cycliques. En effet, les espèces déshydratées se forment en quantité importante quand la température augmente. Comme il a été dit au § I.1, on peut se demander si ces formes constituent une « impasse évolutive » ou si elles peuvent continuer le processus de polymérisation.

Nous avons donc effectué des synthèses à partir de dipeptides cycliques et de monomères. Toutefois, alors que le cyclo(EE) est disponible commercialement, ce n'est pas le cas du cyclo(LL). Nous avons tenté de remplacer celui-ci par le cyclo(AA), dipeptide cyclique de l'alanine, un acide aminé appartenant à la même famille que la leucine, avec une chaîne latérale hydrophobe. Nous n'avons observé qu'une faible polymérisation pour l'échantillon (cycloAA+E)/SiO<sub>2</sub>. Toutefois, il s'est révélé que l'échantillon correspondant avec un dipeptide linéaire (AA+E)/SiO<sub>2</sub> montre lui aussi une faible polymérisation (voir Annexe 5) et donc ne se comporte pas comme les autres systèmes « dipeptide + monomère ». Nous ne discuterons donc pas en détail les résultats correspondants et nous nous concentrerons sur la comparaison entre les échantillons (EE+L)/SiO<sub>2</sub> et (cycloEE+L)/SiO<sub>2</sub>, dont les spectres de masse sont regroupés sur la Figure IV-9. Par ailleurs, la Figure IV-8a et la Figure IV-8b permettent de comparer les distributions de polypeptides à 130 et 160 °C pour ces deux échantillons.

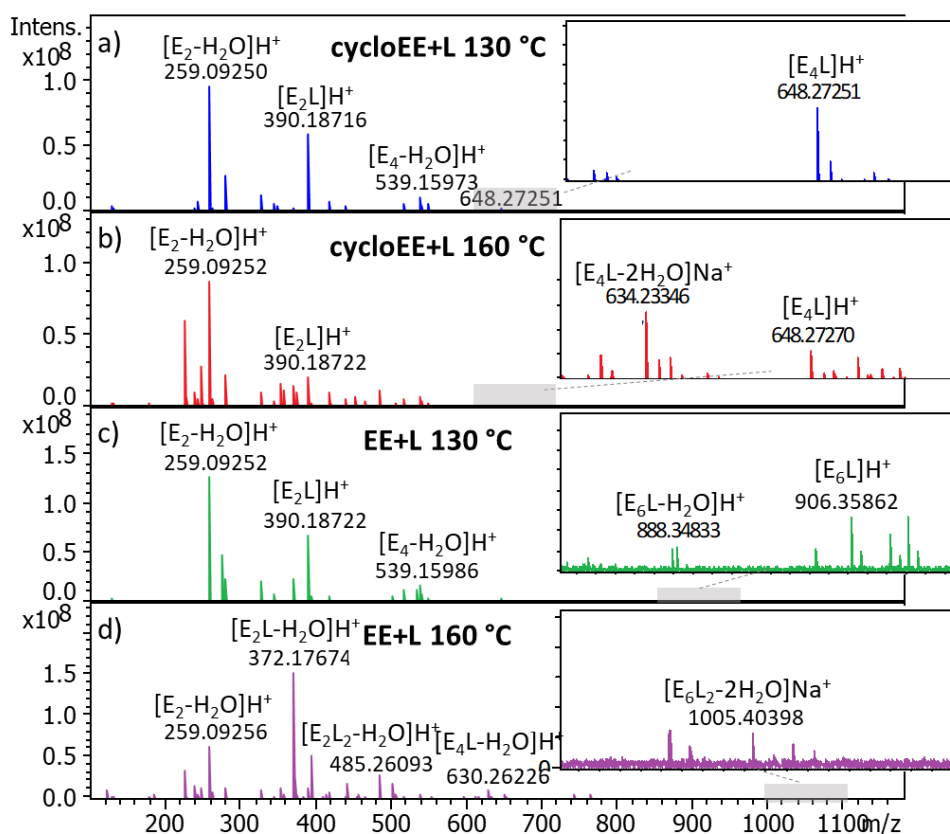


Figure IV-9. Comparaison des spectres de masse obtenus pour les polymérisations des échantillons a) (cycloEE+L)/SiO<sub>2</sub> 130°C, b) (cycloEE+L)/SiO<sub>2</sub> 160 °C, c) (EE+L)/SiO<sub>2</sub> 130 °C et d) (EE+L)/SiO<sub>2</sub> 160 °C (ESI positif, instrument FTICR). Un zoom sur la gamme des oligopeptides les plus longs détectés est reporté pour chaque spectre de masse. La gamme donnée est grisée sur le spectre.

Il apparaît que les longueurs des oligopeptides formés varient avec la température d'activation. Ainsi, les échantillons (cycloEE+L)/SiO<sub>2</sub> activés à 130 °C et 160 °C (Figure IV-9a et b) mènent à la formation de pentapeptides tandis que les échantillons (EE+L)/SiO<sub>2</sub> activés à 130 °C et 160 °C permettent respectivement la formation d'heptapeptides et octapeptides (Figure IV-9c et d). Concernant l'échantillon (cycloEE+L)/SiO<sub>2</sub>, parmi les tripeptides, l'espèce majoritaire est de stœchiométrie E<sub>2</sub>L, indiquant la capacité d'une espèce cyclique à participer à l'élongation peptidique. Cependant, la polymérisation d'une espèce cyclique apparaît moins efficace que celle de l'espèce intacte correspondante. Ainsi, à 130 °C, les tétrapeptides sont présents en quantité significative pour (EE+L)/SiO<sub>2</sub>, mais très minoritaires pour (cycloEE+L)/SiO<sub>2</sub>. Quant à l'indice de polymérisation, il vaut 2,36 à 130 °C et 2,58 à 160 °C quand on part du dipeptide cyclique (à comparer avec 2,66 et 3,21 à partir du dipeptide linéaire).

Il faut aussi souligner que le dipeptide cyclo(EE) comporte deux chaînes latérales avec des fonctions acides qui pourraient permettre de continuer la polymérisation sans ouverture de cycle (formation de  $\gamma$ -peptides), produisant ainsi des espèces déshydratées. Or, il est visible sur la Figure IV-9 que les espèces  $E_2L$  formées par activation thermique de (cycloEE+L)/SiO<sub>2</sub> ne sont pas plus déshydratées que celles formées par (EE+L)/SiO<sub>2</sub>. L'ouverture de cycle lors de la condensation paraît donc probable.

Nous avons aussi étudié l'activation thermique du système (pE+L)/SiO<sub>2</sub> (voir Annexe 6). La polymérisation reste possible en partant de la forme lactame pyroGlu. Les oligopeptides produits sont essentiellement identiques à ceux obtenus pour (L+E)/SiO<sub>2</sub>, ce qui signifie que la formation du lactame n'empêche pas le pyroGlu de réagir par son terminus azoté.

Après avoir étudié la polymérisation à partir d'un dipeptide et d'un monomère, nous avons testé la polymérisation d'un tripeptide avec un monomère. Nous avons choisi le système (ELL+D)/SiO<sub>2</sub> contenant le monomère Asp, afin de pouvoir identifier avec certitude les espèces ayant incorporé le monomère. Le spectre de masse obtenu est reporté sur la Figure IV-10. Plusieurs observations peuvent être faites à partir de cette seule expérience. Tout d'abord, les seuls peptides identifiables résultent d'une condensation sans réarrangement, entre le tripeptide ELL et un à trois résidus Asp (stœchiométries  $DEL_2$ ,  $D_2EL_2$  et  $D_3EL_2$ ), entre deux résidus Asp, mais aussi entre deux tripeptides (stœchiométrie  $E_2L_4$ ). D'autre part, il apparaît que les intensités des oligopeptides résultant d'une condensation tripeptide/tripeptide sont supérieures à celles des espèces résultant d'une condensation tripeptide/monomère. Cela signifie, que les rencontres entre tripeptides sont possibles et que leur condensation est efficace.

Par ailleurs, la présence du pic majoritaire correspondant au tripeptide ELL déshydraté indique que la déshydratation est nettement favorisée.

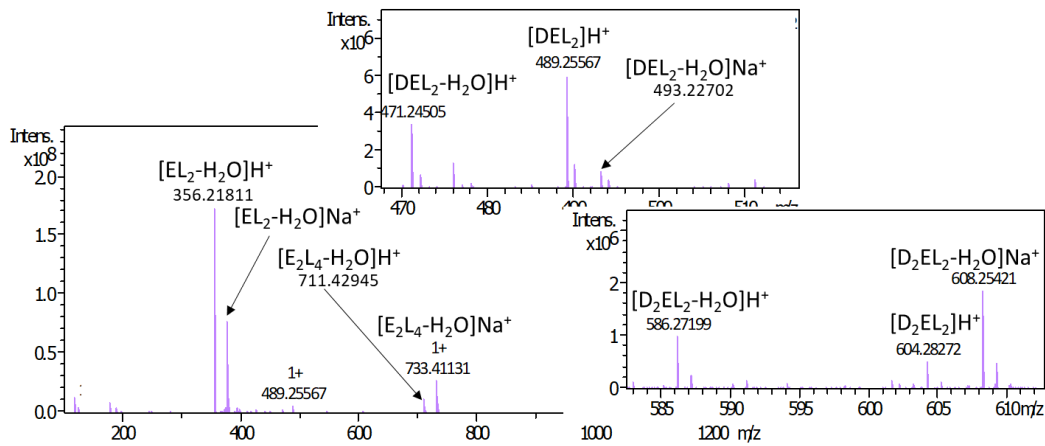


Figure IV-10. Spectre de masse de l'échantillon (ELL+D)/SiO<sub>2</sub> activé à 160 °C (ESI positif, instrument FTICR). Des zooms sur les espèces contenant le résidu Asp sont montrés.

## II. Variation des quantités d'acides aminés déposés sur la silice

### II.1. Effet de la densité d'acides aminés déposés : importance de la diffusion de surface

Pour étudier l'influence de la quantité d'acides aminés déposés sur la silice sur l'efficacité de la polymérisation, les échantillons (L+E)<sub>(0,25 : 0,25)</sub>/SiO<sub>2</sub> et (L+E)<sub>(3 : 3)</sub>/SiO<sub>2</sub> ont été étudiés. Les acides aminés sont toujours en proportions équimolaires mais la quantité totale déposée est différente : alors qu'elle était de  $1 \cdot 10^{-4}$  mol par gramme de silice en chaque acide aminé pour l'échantillon (L+E)/SiO<sub>2</sub>, elle est respectivement de  $0,25 \cdot 10^{-4}$  mol/g et de  $3 \cdot 10^{-4}$  mol/g pour (L+E)<sub>(0,25 : 0,25)</sub>/SiO<sub>2</sub> et (L+E)<sub>(3 : 3)</sub>/SiO<sub>2</sub>. La Figure IV-11 montre la comparaison de la polymérisation de ces deux échantillons avec celle de l'échantillon de référence (L+E)/SiO<sub>2</sub>.



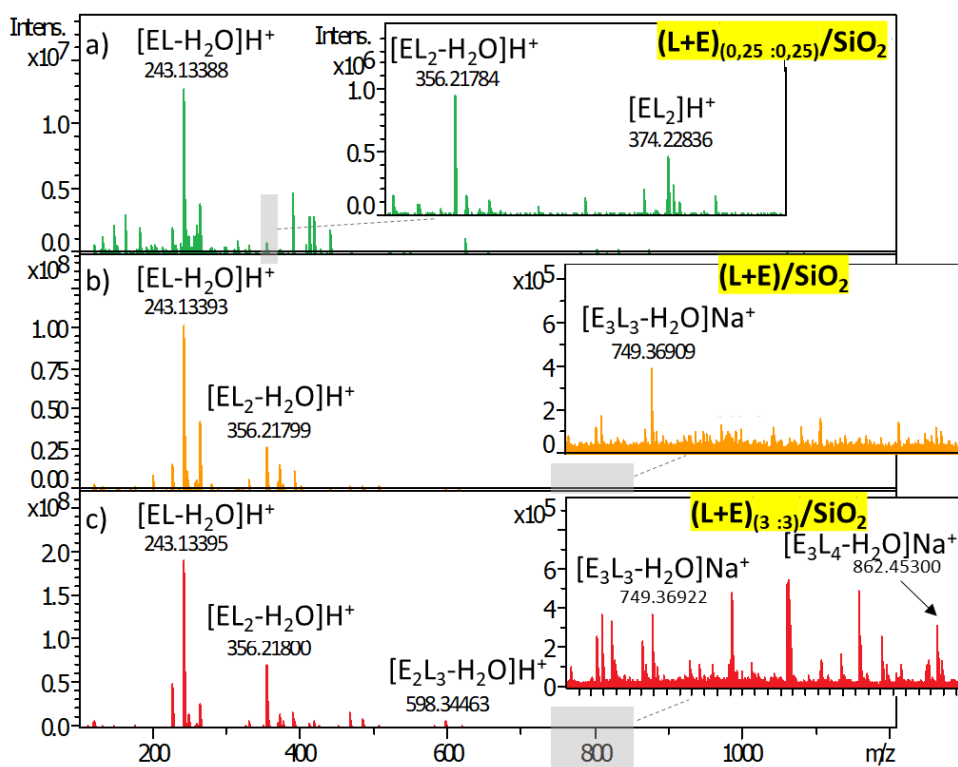


Figure IV-11. Spectres de masse obtenus pour les échantillons a)  $(L+E)_{(0,25 : 0,25)}/SiO_2$ , b)  $(L+E)/SiO_2$  et c)  $(L+E)_{(3 : 3)}/SiO_2$  (ESI positif, instrument FTICR). Un zoom sur les oligopeptides les plus longs est montré pour chaque spectre ; les gammes de  $m/z$  correspondantes sont grisées.

La polymérisation se produit bien dans les trois cas. Pour des quantités déposées de  $0,25 \cdot 10^{-4}$  mol,  $1 \cdot 10^{-4}$  mol et  $3 \cdot 10^{-4}$  mol, les espèces les plus longues détectées sont respectivement des tripeptides, des hexapeptides et des heptapeptides dans les mêmes conditions d'analyse. Néanmoins, il faut noter qu'après l'étape d'extraction, les concentrations dans les solutions analysées sont différentes dans les trois cas, ce qui peut introduire un biais, les peptides les plus longs étant sans doute perdus dans le bruit de fond. Une comparaison de la longueur maximale des oligopeptides aurait été possible si le facteur de dilution avait compensé la différence de quantité de matière déposée.

Même déposés avec une faible densité sur la surface de silice, les acides aminés ont polymérisé. Nous avons déjà souligné que dans l'expérience de référence, la densité superficielle d'acides aminés est inférieure à une monocouche physique : elle ne représente qu'environ 11 % de celle-ci (Chapitre II, § I.4). En d'autres termes, si les acides aminés sont répartis de façon uniforme sur la surface, ils ne sont théoriquement pas au contact les uns des autres. Dans le cas de l'échantillon  $(L+E)_{(0,25 : 0,25)}/SiO_2$ , la densité est quatre fois plus

basse, soit 2,75 % d'une monocouche physique. Toujours dans le cas d'une répartition uniforme, deux acides aminés seraient séparés d'une distance égale à quatre à cinq fois leur taille, soit plusieurs nanomètres. Or, pour réagir l'un avec l'autre, deux acides aminés doivent d'abord se rencontrer, et donc se déplacer l'un vers l'autre. Dire que les acides aminés se déplacent correspond à une vision au niveau moléculaire du phénomène physique qu'on appelle la diffusion. De façon générale, tout mécanisme de réaction doit donc inclure une étape de diffusion. Cette étape n'est souvent pas écrite explicitement en cinétique car la diffusion est rarement limitante pour des réactions en solution, mais elle peut l'être dans le cas d'une catalyse hétérogène. Dans notre scénario, où la réaction a lieu entre des molécules préalablement adsorbées, il pourrait s'agir d'une diffusion sur la surface de silice – ou éventuellement d'une diffusion par l'intermédiaire de la phase gazeuse. Ce dernier point sera étudié plus loin. L'efficacité de la diffusion d'une espèce donnée est caractérisée par son coefficient de diffusion, aussi appelé diffusivité.

La Figure IV-12 montre les répartitions des peptides en fonction de leur longueur pour les trois densités de surface testées.

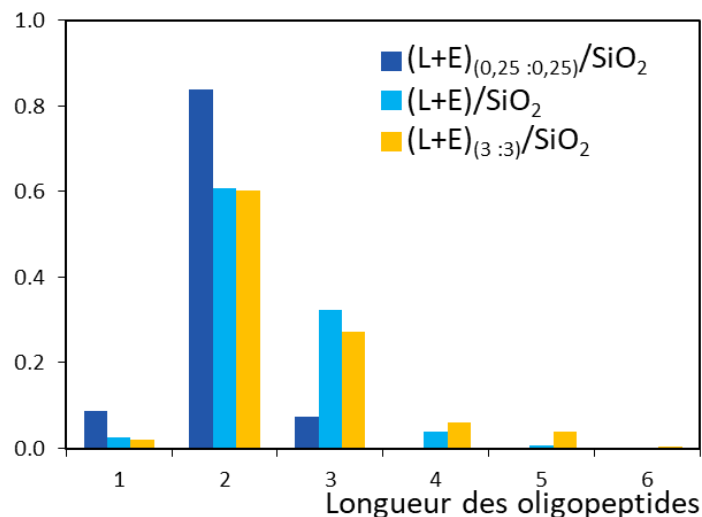


Figure IV-12. Répartition des peptides formés après activation à 160 °C pour les échantillons  $(L+E)_{(0,25:0,25)}/SiO_2$ ,  $(L+E)/SiO_2$  et  $(L+E)_{(3:3)}/SiO_2$ . Les abondances relatives sont reportées en fonction de la longueur des oligopeptides.

La répartition de l'échantillon  $(L+E)_{(0,25:0,25)}/SiO_2$  est significativement différente des deux autres : les dipeptides prédominent et les tripeptides ne représentent que quelques pourcents de l'abondance totale des pics d'oligopeptides. L'indice de polymérisation est de

1,99. Cela suggère que la diffusion est l'étape limitante dans ce cas. Nous avons, dans ce cas aussi, étudié l'effet d'une augmentation du temps de réaction : si l'on maintient un palier de 24 h à 160 °C, l'indice de polymérisation passe de 1,99 à 2,06, une différence minime.

La différence entre les échantillons (L+E)<sub>(1:1)</sub>/SiO<sub>2</sub> et (L+E)<sub>(3:3)</sub>/SiO<sub>2</sub> est nettement plus faible, essentiellement limitée aux tétra- et pentapeptides. L'indice de polymérisation est respectivement de 2,40 et 2,55 pour les deux échantillons. Les limitations diffusionnelles constituent un facteur moins important pour ces densités de surface.

## II.2. Variation du rapport Leu :Glu déposé sur la silice

Il a été établi dès les premières études sur nos systèmes que la présence d'un acide aminé trifonctionnel comme Glu est nécessaire pour obtenir des polypeptides linéaires, tandis qu'en leur absence les dipeptides cycliques prédominent. Dans nos conditions standards, les deux acides aminés sont introduits dans des proportions équimolaires. Cependant, la probabilité que Glu soit impliqué dans une liaison peptidique devrait être plus importante. On peut alors se demander si une faible quantité d'acide glutamique pourrait suffire à amorcer la polymérisation, et si au contraire une plus grande quantité de Glu permettrait de former des oligopeptides plus longs. Pour répondre à ces interrogations, des échantillons présentant des rapports Leu :Glu différents ont été préparés : (L+E)<sub>(5:1)</sub>/SiO<sub>2</sub>, (L+E)<sub>(2:1)</sub>/SiO<sub>2</sub>, (L+E)<sub>(1:2)</sub>/SiO<sub>2</sub>, (L+E)<sub>(1:5)</sub>/SiO<sub>2</sub> et (L+E)<sub>(0,5:5,5)</sub>/SiO<sub>2</sub>. Il convient de préciser que tous les échantillons n'ont pas la même quantité totale déposée d'acides aminés. Ainsi, l'échantillon (L+E)<sub>(2:1)</sub>/SiO<sub>2</sub> a la même quantité de Glu que l'échantillon de référence (L+E)/SiO<sub>2</sub> (de rapport molaire 1 :1) mais une quantité de Leu deux fois plus importante ; la quantité totale d'acides aminés sur silice est donc de  $3 \cdot 10^{-4}$  mol/g au lieu de  $2 \cdot 10^{-4}$  mol/g. De même (L+E)<sub>(1:2)</sub>/SiO<sub>2</sub> a une quantité totale en acides aminés de  $3 \cdot 10^{-4}$  mol/g tandis que (L+E)<sub>(5:1)</sub>/SiO<sub>2</sub>, (L+E)<sub>(1:5)</sub>/SiO<sub>2</sub> et (L+E)<sub>(0,5:5,5)</sub>/SiO<sub>2</sub> contiennent une quantité totale de  $6 \cdot 10^{-4}$  mol/g. En toute rigueur, seuls les échantillons contenant la même quantité d'acides aminés totale peuvent être comparés. Ainsi, la Figure IV-13 compare les spectres de masse obtenus pour les échantillons (L+E)<sub>(2:1)</sub>/SiO<sub>2</sub> et (L+E)<sub>(1:2)</sub>/SiO<sub>2</sub>. Pour rappel, celui de l'échantillon de référence (L+E)/SiO<sub>2</sub> est également montré. Tous les échantillons ont été analysés le même jour par FTICR avec une injection automatique (mode FIA, voir Chapitre II).

Des heptapeptides sont détectés pour  $(L+E)_{(2:1)}/SiO_2$  et  $(L+E)_{(1:2)}/SiO_2$  alors que les plus longues espèces détectées sont des hexapeptides pour  $(L+E)/SiO_2$ . Ces différences de longueur maximale pourraient, comme dans le paragraphe précédent, être dues plus à la concentration des solutions de désorption (dépendante de la quantité d'acides aminés déposés) qu'à une différence d'efficacité de la polymérisation. Cette hypothèse est confirmée par la comparaison des spectres de masse des échantillons  $(L+E)_{(5:1)}/SiO_2$ ,  $(L+E)_{(1:5)}/SiO_2$  et  $(L+E)_{(0,5:5,5)}/SiO_2$ , pour lesquels le rapport varie sans changer la quantité totale d'acides aminés, montrée sur la Figure IV-14 ; les espèces les plus longues obtenues sont dans les trois cas des octapeptides.

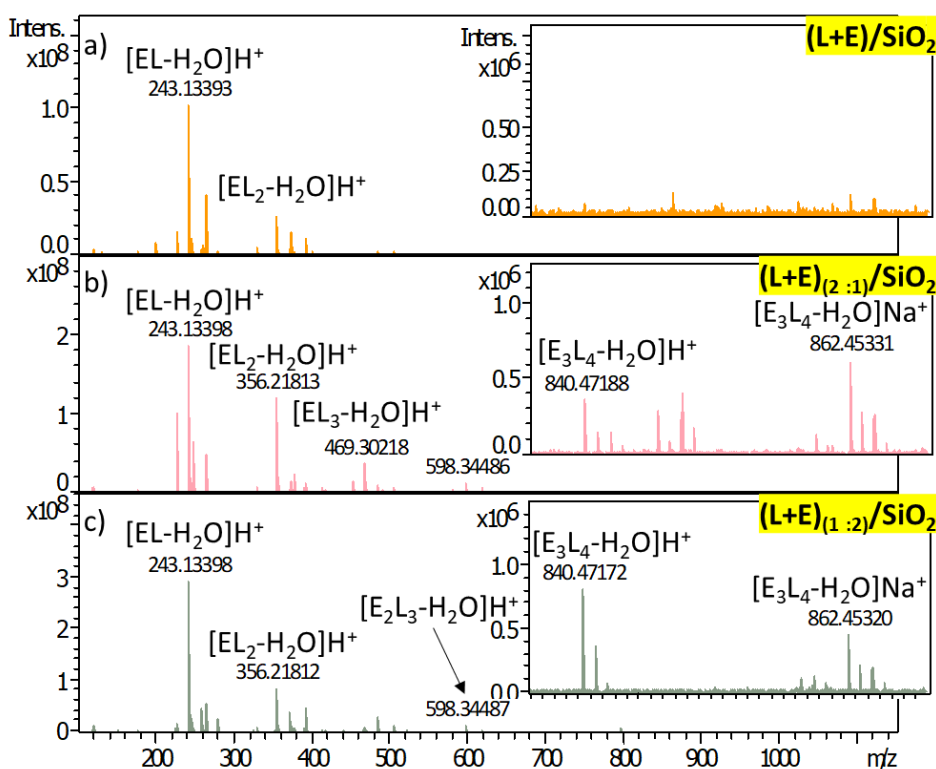


Figure IV-13. Spectres de masse obtenus pour les échantillons a)  $(L+E)/SiO_2$ , b)  $(L+E)_{(2:1)}/SiO_2$  et c)  $(L+E)_{(1:2)}/SiO_2$  (dilution 1/25, ESI positif, instrument FTICR). Un zoom sur la gamme  $m/z$  835-870 est montré pour chaque spectre, permettant de voir les heptapeptides détectés dans les spectres b et c.

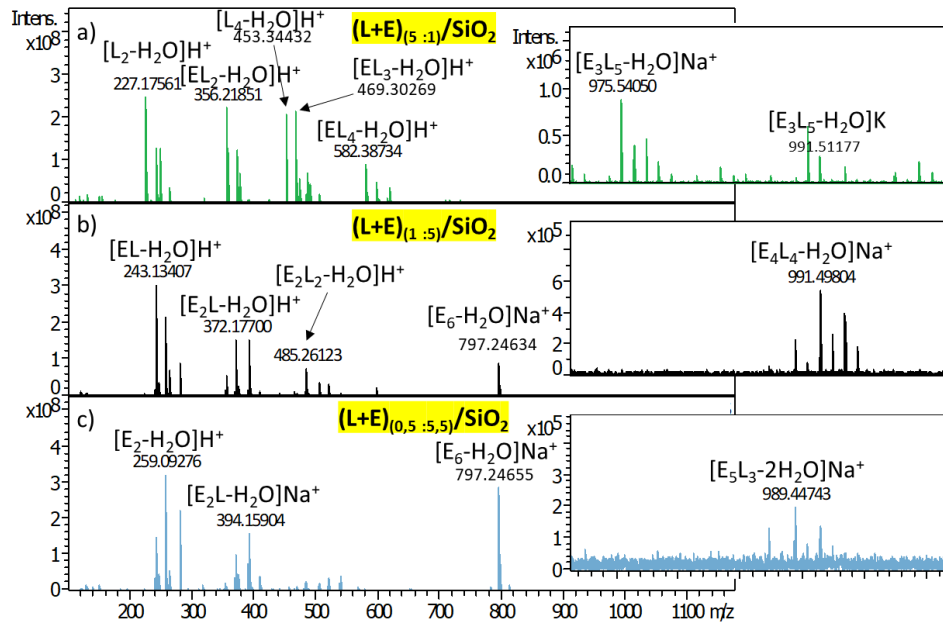


Figure IV-14. Spectres de masse obtenus pour les échantillons a)  $(L+E)_{(5:1)}/SiO_2$ , b)  $(L+E)_{(1:5)}/SiO_2$  et c)  $(L+E)_{(0,5:5,5)}/SiO_2$  (dilution 1/25, ESI positif, instrument FTICR). Un zoom sur la gamme m/z 970-1000 (octapeptides) est montré pour chaque spectre.

La Figure IV-15 représente les indices de polymérisation de chaque expérience en fonction de la densité totale d'acides aminés déposés pour chaque échantillon. On remarque que pour une même quantité d'acides aminés, les échantillons contenant un excès de Leu ont des indices de polymérisation plus importants.

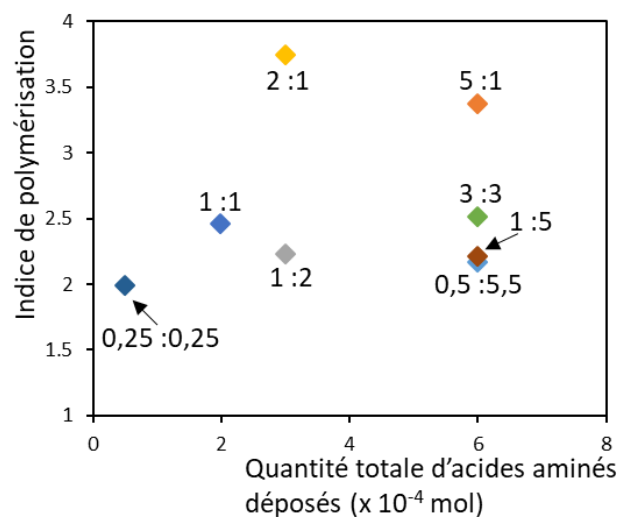


Figure IV-15. Indice de polymérisation en fonction de la quantité d'acides aminés et du rapport molaire Leu :Glu.

Par ailleurs, pour une longueur d'oligopeptide donnée, la distribution des différentes stœchiométries est également impactée par le rapport Leu :Glu de l'échantillon. Ainsi le tripeptide majoritaire est l'espèce  $[EL_2-H_2O]H^+$  pour  $(L+E)_{(5:1)}/SiO_2$  tandis que c'est l'espèce  $[E_2L-H_2O]H^+$  pour l'échantillon  $(L+E)_{(1:5)}/SiO_2$  (voir spectres de masse a et b de la Figure IV-14). Pour étudier ce phénomène, nous nous focaliserons ici sur le cas des tripeptides.

L'entropie de Shannon permet de caractériser l'écart global à l'aléatoire d'une distribution expérimentale. Les calculs, appliqués aux tripeptides, montrent que celle-ci diminue de façon marquée quand le rapport Leu/Glu augmente : les répartitions sont proches de l'aléatoire pour les échantillons riches en Glu, mais plus sélectives au contraire pour les échantillons riches en Leu. Ainsi, pour l'échantillon  $(L+E)_{(5:1)}/SiO_2$ , l'entropie de Shannon est égale à 49 % d'une distribution aléatoire, à comparer à 89 % pour l'échantillon  $(L+E)_{(1:5)}/SiO_2$ . Si l'écart global à l'aléatoire peut ainsi être évalué, cette méthode n'indique pas quelles stœchiométries sont favorisées. Pour étudier la sélectivité de plus près, nous avons calculé pour chaque stœchiométrie de tripeptides le rapport entre l'abondance relative expérimentale et l'abondance relative attendue dans le modèle aléatoire, selon la même méthode qu'au Chapitre III, en corrigeant l'abondance théorique par les efficacités d'ionisation mesurées expérimentalement. La Figure IV-16 montre le résultat de ce calcul pour les différents rapports L : E. Un résultat supérieur à 1 indique que la stœchiométrie est favorisée par rapport à l'hypothèse aléatoire, et inversement.

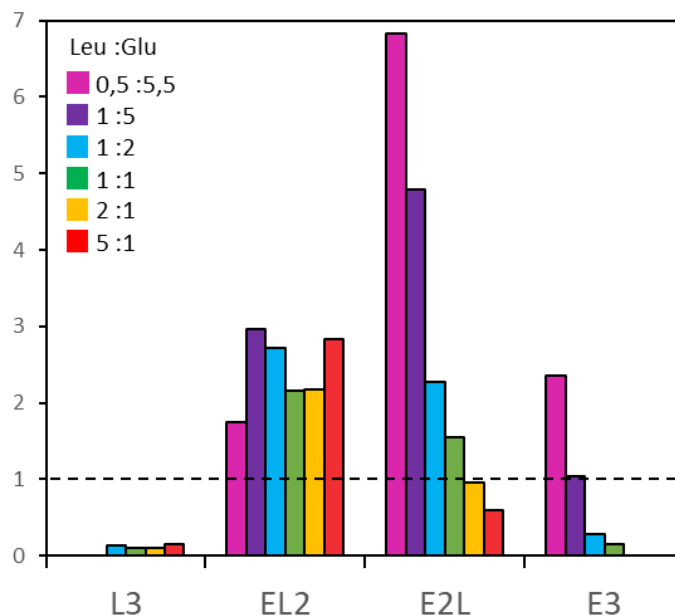


Figure IV-16. Rapport entre les abondances expérimentales relatives et les abondances attendues dans le modèle aléatoire corrigées pour les différents tripeptides pour les différents rapports molaires Leu :Glu des deux acides aminés déposés.

Les homopeptides sont généralement fortement défavorisés, à l'exception des tripeptides E<sub>3</sub> dans les échantillons très riches en Glu ; les L<sub>3</sub> quant à eux sont en faible abondance quel que soit l'échantillon étudié. En ce qui concerne les hétéropeptides, ceux de stœchiométrie E<sub>2</sub>L sont favorisés dans les échantillons riches en Glu. Ces résultats sont intéressants car ils suggèrent que la polymérisation de surface pourrait exacerber une dissymétrie de composition préexistante dans le mélange d'acides aminés. Il serait souhaitable de pouvoir généraliser aux peptides de longueur plus élevée, mais il est difficile de le faire dans l'état actuel des connaissances, en l'absence d'informations précises sur les efficacités d'ionisation de tous les oligopeptides.

### III. Condensation sur d'autres surfaces minérales

L'étude de l'échantillon (L+E)/SiO<sub>2</sub> a démontré l'influence de la surface sur les oligopeptides formés (par exemple en comparant avec une polymérisation en solution, voir Chapitre III). Quelques essais de polymérisation en déposant ces mêmes acides aminés sur d'autres types de surfaces ont été réalisés. Notre choix s'est porté sur la montmorillonite et sur l'alumine,

deux minéraux présents sur la Terre primordiale et ayant déjà été considérés par certains auteurs pour la polymérisation (voir Chapitre I). Le protocole utilisé est similaire à celui de (L+E)/SiO<sub>2</sub> (voir Chapitre II). Les échantillons sont activés à 160 °C avec une rampe de 1 °C/min puis un plateau de 30 min à 160 °C est appliqué. Ils sont ensuite analysés par spectroscopie infrarouge et par spectrométrie de masse à haute résolution.

### III.1. Réactivité de Leu + Glu sur montmorillonite

La surface de montmorillonite a déjà été utilisée dans des études de polymérisation dans un contexte prébiotique (voir Chapitre I). Par exemple, Bujdák<sup>45</sup> a montré une sélectivité de polymérisation sur montmorillonite.

Les spectres infrarouge (opérées en mode ATR) avant activation thermique, après activation thermique et après désorption sont reportés en Annexe 7. Avant activation thermique, des bandes correspondant aux acides aminés monomériques sont observées. Aucune bande amide, caractéristique de la formation d'une liaison peptidique, n'est observée après activation thermique ; notamment la bande amide I, attendue vers 1671 cm<sup>-1</sup>, est absente.

Ces observations sont confirmées par l'analyse par spectrométrie de masse (spectre en Annexe 8). La solution de désorption de l'échantillon montre surtout les pics correspondant aux monomères (avec des intensités faibles, comparables à celles des pics du solvant). Toutefois, la présence du pic à m/z 245,18642 correspondant au dipeptide linéaire [L<sub>2</sub>]H<sup>+</sup> pose question. Des liaisons peptidiques auraient donc été formées en faible quantité, sans doute trop faible pour être détectées en spectroscopie IR. Par ailleurs, un échantillon de montmorillonite vierge a subi la même procédure d'extraction et le désorbat a été analysé par FTICR. Aucune molécule organique n'a été détectée, indiquant que la montmorillonite ne relargue pas de telles molécules.

Une activation thermique à 160 °C ne permet donc pas de mener à une polymérisation efficace. Ces résultats semblent compatibles avec la littérature : dans les études de Jaber et al.<sup>93</sup> et de Castro Silva et al.<sup>95</sup>, la formation d'une liaison peptidique sur des argiles de type smectites nécessite une température supérieure à 200 °C, et dans celle de Fuchida et al.<sup>218</sup> un chauffage à 150 °C de longue durée. Bujdák<sup>45</sup> chauffait à une température encore plus faible, 80 °C, mais en appliquant de nombreux cycles impliquant un chauffage de 24 h, et



cette procédure ne fournissait que des rendements inférieurs à 1 % de polymérisation. La montmorillonite est donc un minéral moins intéressant du point de vue prébiotique que la silice car des températures importantes sont nécessaires pour conduire à une polymérisation.

La faible intensité des pics correspondant aux acides aminés observés sur le spectre de masse est paradoxale. L'étape de désorption est peut-être inefficace pour extraire les acides aminés et/ou les peptides de la surface de montmorillonite. Mais si c'était le cas, ceux-ci devraient être encore apparents sur le spectre infrarouge de la montmorillonite après désorption ; or, celui-ci ne montre qu'une bande à  $1628\text{ cm}^{-1}$  qui est due à de la vibration de déformation de l'eau. Nous n'avons pas cherché à résoudre ce problème puisque la montmorillonite ne semblait pas constituer un minéral prometteur.

### III.2. Réactivité de Leu + Glu sur alumine

Plusieurs auteurs ont affirmé que l'alumine gamma est l'un des catalyseurs les plus efficaces pour la formation de liaisons peptidiques<sup>67,219,220</sup>. De plus, Bujdak<sup>58</sup> a observé une sélectivité de polymérisation sur un support d'alumine non poreuse (cf. Chapitre I).

Les spectres IR avant et après activation thermique d'un échantillon (L+E)/Al<sub>2</sub>O<sub>3</sub> sont reportés en Figure IV-17.

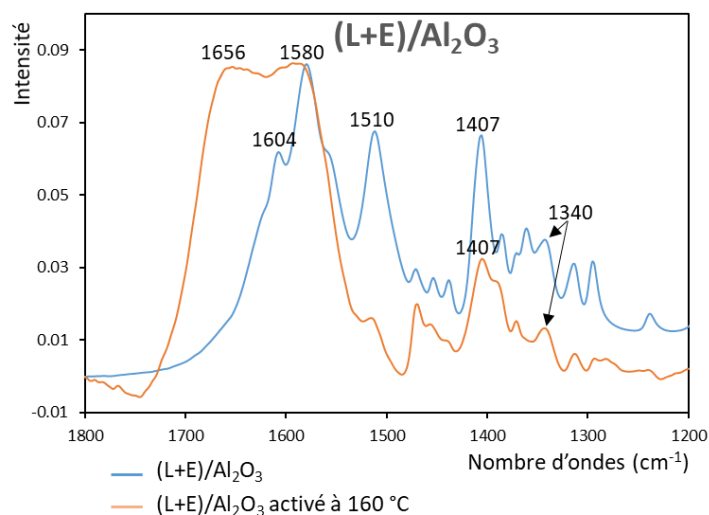


Figure IV-17. Comparaison des spectres IR de l'échantillon (L+E)/Al<sub>2</sub>O<sub>3</sub> avant et après activation thermique. Une normalisation par rapport à la bande à  $1580\text{ cm}^{-1}$  a été effectuée.

Dans ce cas, les spectres avant et après activation thermique sont différents. Avant activation, la bande à  $1604\text{ cm}^{-1}$  pourrait correspondre aux vibrations asymétriques  $\nu_a(\text{CO}_2^-)$  des groupements acides carboxyliques liés à la surface des atomes d'aluminium<sup>67</sup>. La bande à  $1407\text{ cm}^{-1}$  correspond à l'élongation symétrique  $\nu_s(\text{CO}_2^-)$  de l'acide carboxylique, la bande à  $1340\text{ cm}^{-1}$  à la rotation des  $\text{CH}_2$ , celles à  $1580\text{ cm}^{-1}$  et  $1510\text{ cm}^{-1}$  correspondent respectivement aux déformations asymétrique et symétrique des  $\text{NH}_3^+$ . Après activation thermique, une nouvelle bande apparaît à  $1656\text{ cm}^{-1}$ . Cette bande correspond certainement aux vibrations  $\nu(\text{C=O})$  de la fonction amide (bande amide I)<sup>67</sup>. Mais d'autres bandes attribuables aux acides aminés monomériques restent présentes, notamment à  $1510$  et  $1407\text{ cm}^{-1}$ . Néanmoins, la bande à  $1510\text{ cm}^{-1}$  diminue considérablement après traitement thermique ; une bonne partie des monomères sont donc consommés.

Sur alumine, la formation de liaisons peptidiques semble donc possible à la température de  $160\text{ }^\circ\text{C}$  et cela est confirmé par les analyses de spectrométrie de masse. En effet, des oligopeptides de leucine sont observés sur le spectre de masse obtenu (voir Figure IV-18). Il a été vérifié que ces espèces ne sont pas des dimères non covalents. Une observation importante est également la quasi-absence de formes déshydratées, contrairement aux polymérisations effectuées sur un support de silice. L'espèce prédominante est le dipeptide linéaire  $\text{L}_2$ , alors que sur silice la polymérisation de la leucine donne essentiellement le dipeptide cyclique. Par ailleurs, aucun polypeptide contenant Glu n'est détecté, et le monomère Glu est également absent du spectre de masse. On peut donc penser que la procédure de désorption utilisée ne permet pas d'extraire l'acide aminé Glu (et les peptides contenant ce résidu) de l'alumine, probablement en raison d'interactions irréversibles de ce résidu avec la surface. Le spectre IR après désorption est de mauvaise qualité mais montre une absorption non négligeable vers  $1590\text{ cm}^{-1}$ , correspondant à de la matière organique (Annexe 9).

Dans l'étude de Basiuk et al<sup>67</sup>, une solution de  $\text{CaCl}_2$  à  $100\text{ mM}$  était utilisée pour la désorption des produits de polymérisation. Cette procédure est calquée sur celle qu'on utilise pour désorber les cations compensateurs d'une argile. Or, il est improbable que l'acide glutamique soit retenu sous forme cationique – la forme cationique ne prédomine en solution que pour des pH inférieurs à 2. Il est plus plausible que le mécanisme de rétention de Glu fasse intervenir une coordination aux ions  $\text{Al}^{3+}$  de surface par les fonctions

carboxylate, comme proposé par Greiner et al.<sup>221</sup>. Une telle adsorption serait difficilement réversible, surtout si elle est polydentate.

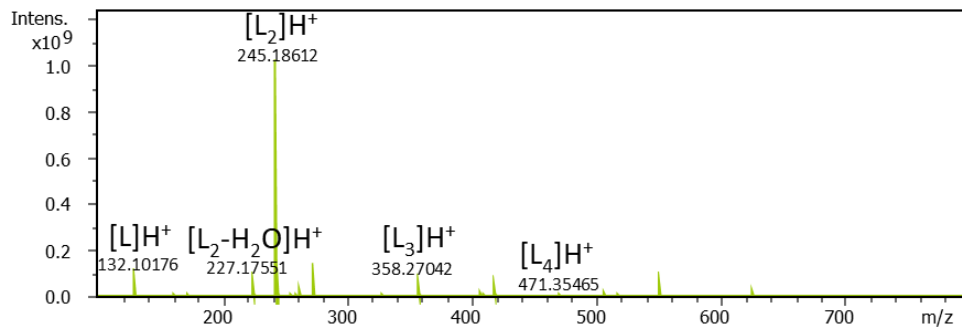


Figure IV-18. Spectre de masse de l'échantillon (L+E)/Al<sub>2</sub>O<sub>3</sub> activé à 160 °C (ESI positif, instrument FTICR).

En résumé, l'étude de polymérisation du même système Leu + Glu sur d'autres surfaces minérales a mis en évidence l'influence du support utilisé. Pour une même température d'activation de 160 °C, la polymérisation se produit sur silice mais pas sur montmorillonite. Cela confirme que la surface de silice a un rôle catalytique pour la formation des liaisons peptidiques ; les sites responsables de la catalyse ne sont pas précisément identifiés, mais ils ne se trouvent pas sur une surface d'argile smectite. Quant à l'alumine, elle montre une réactivité encore différente : en présence du mélange Leu + Glu, seul le second acide aminé semble adsorbé fortement, tandis que les molécules de Leu réagissent entre elles. Ainsi, on n'obtient pas sur alumine les oligopeptides linéaires mixtes qui sont formés sur silice et ont constitué la motivation de notre étude. Cela ne veut pas dire que cette surface soit dépourvue d'intérêt : en effet, la procédure d'activation thermique + désorption aboutit à une séparation des deux constituants du mélange, causant de la sorte une sélection des acides aminés qui pourrait aussi trouver sa place dans un processus prébiotique.

Ces résultats, même incomplets, soulignent que chaque surface minérale a sa propre spécificité : on ne peut pas généraliser les conclusions obtenues sur une surface à n'importe quelle autre. Le monde minéral qui a préexisté à l'apparition de la vie est un monde complexe, et cette complexité se retrouve dans son interaction avec les « briques du vivant ».

## IV. Evaluation d'autres scénarios de polymérisation

### IV.1. Cycles d'hydratation /déshydratation

Comme discuté dans le Chapitre I, l'utilisation de cycles hydratation/déshydratation pour la polymérisation est couramment rencontrée dans des études de chimie prébiotique, l'objectif étant de mimer les changements climatiques de l'environnement géochimique. L'étape d'activation thermique à 160 °C (séchage en conditions sèches) étant bien caractérisée (voir parties précédentes avec étude cinétique), l'étude discutée ici consiste à poursuivre la polymérisation par un cycle en réhumidifiant l'échantillon activé thermiquement à 160 °C puis en l'activant une seconde fois, cette fois dans des conditions où l'activité de l'eau est élevée.

Pour ce faire, l'échantillon est d'abord activé classiquement à 160 °C sous flux d'argon puis 50 mg de l'échantillon sont remouillés avec 500 µL d'eau MilliQ (10 mL/g) pour former un gel. Le gel ainsi obtenu est directement mis à l'étuve à 70 °C ou 100 °C pendant 48 h dans un flacon hermétiquement fermé. Les échantillons résultants sont appelés (L+E)/SiO<sub>2</sub> 160+70 et (L+E)/SiO<sub>2</sub> 160+100. Comme pour la procédure classique, une étape d'extraction est ensuite opérée : l'ensemble de l'échantillon est désorbé dans 5 mL d'eau MilliQ puis centrifugé.

Selon Marshall-Bowmann<sup>40</sup>, pour des températures inférieures à 70 °C les cinétiques de réaction de condensation et d'hydratation sont très lentes, même en présence de surfaces minérales. Nous avons donc choisi de procéder à deux activations différentes, une à 70 °C, qui est la température caractéristique de sources hydrothermales comme le site actuel de Lost City<sup>222</sup> et, d'après certains auteurs, la température de l'océan archéen<sup>223</sup>. La deuxième température choisie est de 100 °C pour espérer former plus d'oligopeptides tout en restant dans des conditions réalistes (donc pas au-dessus de 100 °C dans l'eau).

Les spectres de masse obtenus pour les trois échantillons (L+E)/SiO<sub>2</sub>, (L+E)/SiO<sub>2</sub> 160+70 et (L+E)/SiO<sub>2</sub> 160+100 sont comparés en Figure IV-19. Les deux premiers sont identiques, les espèces les plus longues détectées sont des hexapeptides. En revanche, si les oligopeptides majoritaires sont également les mêmes, l'échantillon (L+E)/SiO<sub>2</sub> 160+100 montre la formation d'heptapeptides nouveaux. De plus, la quantité d'oligopeptides de longueur supérieure à trois résidus est plus élevée que dans l'expérience de référence.

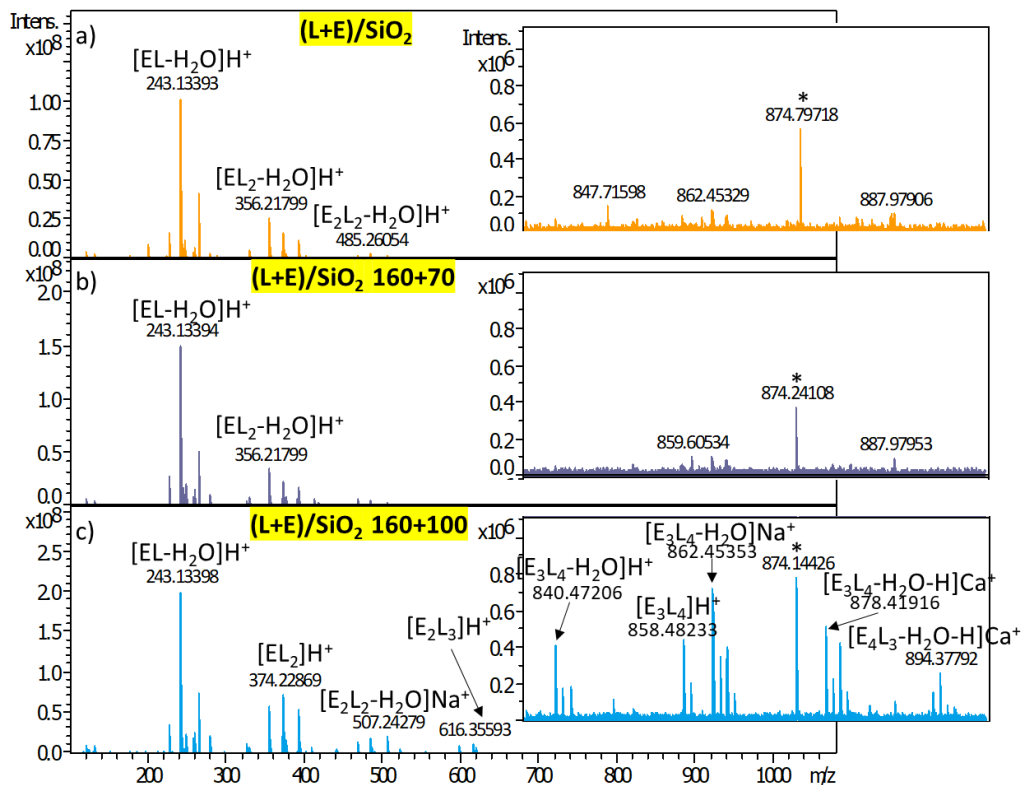


Figure IV-19. Spectres de masse des échantillons a) (L+E)/SiO<sub>2</sub>, b) (L+E)/SiO<sub>2</sub> 160+70 et c) (L+E)/SiO<sub>2</sub> 160+100 (ESI positif, instrument FTICR). Un zoom sur la zone des heptapeptides est montré pour chaque spectre. Cette longueur d'oligopeptides n'est présente que dans l'échantillon où la seconde activation est faite à 100 °C.

L'application de conditions d'humidité fluctuantes (séchage suivi de réhydratation) aboutit donc bien à un effet positif pour la formation de longues chaînes peptidiques, même s'il reste assez limité dans nos conditions. On peut en proposer une interprétation si on accepte que la diffusion des espèces réactives peut constituer l'étape délimitant la vitesse, comme il a été proposé plus haut. Que la diffusion se fasse en surface ou par l'intermédiaire de la phase gazeuse, il est probable que les monomères sont les espèces les plus mobiles, et que la mobilité diminue avec la longueur des peptides. Quand les espèces plus longues deviennent prédominantes, la faible diffusivité (coefficient de diffusion) de celles-ci stoppe les réactions de condensation. Si le système est remis au contact d'une phase aqueuse, les oligopeptides, désorbant facilement, pourront diffuser en solution et réagir ensemble lorsque l'étape de séchage suivante sera appliquée. On observe en effet que les oligopeptides longs formés dans l'échantillon (L+E)/SiO<sub>2</sub> 160+100 (Figure IV-19c) pourraient être obtenus par condensation des deux espèces majoritaires qui avaient été formées lors

de la première activation, soit EL et EL<sub>2</sub>, selon : EL + EL = E<sub>2</sub>L<sub>2</sub>, EL + EL<sub>2</sub> = E<sub>2</sub>L<sub>3</sub> puis E<sub>2</sub>L<sub>2</sub> + EL<sub>2</sub> = E<sub>3</sub>L<sub>4</sub>.

## IV.2. Dépôt de Leu et Glu sur des particules de silice différentes

Pour tous les échantillons étudiés précédemment, les acides aminés étaient mélangés en solution aqueuse avant d'être déposés sur une surface. Afin d'étudier la mobilité des acides aminés, nous avons préparé un échantillon dans lequel les deux acides aminés, déposés séparément, sont adsorbés sur des nanoparticules différentes de silice. Les deux acides aminés ont simplement été déposés en solution aqueuse puis séchés sur deux échantillons de silice différents. Les deux poudres ont ensuite été mélangées physiquement avant d'être activés thermiquement ensemble à 160 °C selon la procédure habituelle. De cette façon, chaque nanoparticule de silice a initialement sa surface recouverte soit de Leu, soit de Glu, mais pas des deux acides aminés ensemble. Les analyses ont été faites sur FTICR en injection directe. Le spectre de masse obtenu est montré Figure IV-20.

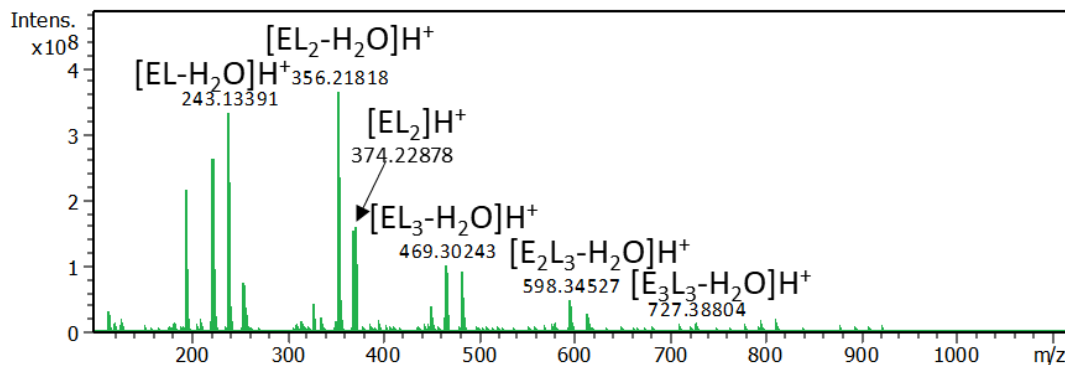


Figure IV-20. Spectre de masse de l'échantillon (L+E)/SiO<sub>2</sub> dans lequel les deux acides aminés ont été déposés séparément, puis activé à 160 °C (ESI positif, instrument FTICR).

Ce spectre de masse est identique à celui de référence pour un dépôt classique. Dans les conditions d'analyse utilisées, des oligopeptides de différentes longueurs allant jusqu'aux octapeptides sont détectés dans les deux cas. Ces résultats indiquent donc que les acides aminés seraient mobiles entre deux grains de silice différents, au moins pour une activation

à 160 °C. Il y a donc très probablement un passage par la phase gazeuse, par désorption d'un (ou des deux) acide(s) aminé(s), suivi d'une réadsorption.

Deux autres activations thermiques, à 130 °C et 145 °C, ont été testées pour étudier l'influence de la température.

Pour chacune des trois expériences, les intensités de tous les pics correspondant respectivement à des oligopeptides et à des pics de solvant ont été sommées pour les douze pics majoritaires du spectre de masse. Les rapports  $\frac{\sum \text{intensités pics oligopeptides}}{\sum \text{intensités pics solvant}}$  ont été calculés pour les deux modes de dépôt et sont reportés dans le Tableau IV-2.

*Tableau IV-2. Comparaison des rapports oligopeptide/solvant pour différentes températures pour les deux modes de dépôt.*

	Dépôt classique	Dépôt séparé
130 °C	0,23	0,30
145 °C	2,50	0,59
160 °C	8,19	4,88

Pour une activation à 130 °C, les rapports sont identiques pour les deux techniques de dépôt ; les pics de solvant sont largement majoritaires. En revanche, le rapport oligopeptide/solvant est quatre fois plus important pour le dépôt classique que pour le dépôt séparé pour une activation thermique à 145 °C. A cette température, donc, les deux acides aminés peuvent réagir ensemble s'ils ont été déposés sur la même particule, mais nettement moins s'ils ont été déposés sur des particules séparées.

Il y aurait donc une diffusion sur la surface de silice, plus rapide que la diffusion par l'intermédiaire de la phase gazeuse. Il existe quelques données sur la diffusion d'acides aminés sur des surfaces métalliques<sup>224</sup>, mais pas à notre connaissance sur des surfaces d'oxydes minéraux.

Pour l'activation à 160 °C, le rapport oligopeptide/solvant du dépôt classique est 1,67 fois plus important que celui du dépôt séparé mais les mêmes espèces y sont présentes. A cette température, il y aurait donc diffusion sur la surface de silice et diffusion en phase gazeuse.

## V. Discussion générale

### V.1. Rappels des données bibliographiques

Il n'existe à notre connaissance aucune tentative dans la littérature de comprendre globalement l'ensemble des étapes menant à la polymérisation des acides aminés sur les surfaces minérales. On trouve en revanche des réflexions théoriques approfondies sur la voie réactionnelle catalytique menant à la formation des liaisons peptidiques, y compris la nature exacte des sites catalytiques sur surfaces de silice, et un certain nombre de données expérimentales concernant des réactions entre peptides particuliers, que nous allons brièvement passer en revue.

Comme nous l'avons signalé au Chapitre I, § I.3.3, les acides aminés déposés sur surface à partir d'une solution aqueuse sont sous forme zwitterionique, comme ils le sont en solution<sup>89</sup>. La forme zwitterionique est plus stable que la forme neutre tant que la surface comporte des molécules d'eau coadsorbées<sup>91,92</sup>, alors que la forme neutre est plus stable sur une surface complètement anhydre.

Les condensations peptidiques, qui nécessitent une attaque nucléophile par un groupe amine non ionisé, commencent seulement à des températures supérieures à 100 °C. Elles doivent impliquer une catalyse par des sites de surface de la silice car à l'état gazeux par exemple, les acides aminés ne se condensent pas à ces températures. Dès les premières démonstrations expérimentales de la polymérisation de surface, des hypothèses ont été formulées sur le mécanisme catalytique, basées sur une vision du mécanisme d'adsorption. En 1984, White et al.<sup>90</sup> ont postulé la formation « d'acyl silicates », aussi appelés « esters de surface » ou « anhydrides de surface » par formation d'une liaison covalente entre le groupe acide carboxylique et un silanol, entraînant une diminution de la densité électrique du carbone de l'acide carboxylique, ce qui faciliterait l'attaque nucléophile par le groupement amine d'un autre acide aminé.

Ultérieurement, Rimola et al.<sup>225</sup> ont montré, par modélisation moléculaire, que la formation d'esters de surface n'est thermodynamiquement possible que par ouverture d'un cycle contraint à 3 ou 4 siloxanes, et donc que ce mécanisme n'est possible que sur les silices possédant initialement ce type de défauts ; sur d'autres silices (dont le support que nous



avons utilisé, l'Aerosil 380), la catalyse se fait alors probablement par un mécanisme complexe impliquant un réseau de liaisons H avec les réactifs<sup>226</sup>. L'étape de condensation catalytique est donc déjà extrêmement complexe et encore mal comprise.

Nous l'avons dit, les scénarios d'activation thermique sur surfaces minérales ont été peu appliqués aux mélanges d'acides aminés. Certains résultats pertinents ont été obtenus dans des études assez anciennes portant sur l'activation à température élevée de mélanges d'acides aminés massiques, notamment par Fox et Harada<sup>51,52</sup>. Un modèle mécanistique menant à la formation de tripeptides a été proposé par Hartmann et al<sup>56</sup>. Ces études ont été mentionnées au Chapitre I (§ I.2.3). Il faut souligner toutefois que ce mécanisme n'est pas supporté par des données précises, qu'elles soient spectroscopiques ou cinétiques : il s'agit seulement d'une hypothèse conforme à l'intuition chimique générale.

## V.2. Conclusions de la présente étude

### V.2.1. Identification des étapes élémentaires de la polymérisation

Les résultats expérimentaux montrent que le mécanisme impliqué dans la condensation de l'échantillon (L+E)/SiO<sub>2</sub> ne suit pas celui en trois étapes proposé par Hartmann et al.<sup>56</sup> expliqué au Chapitre I (§ I.2.3). En effet, le dipeptide cyclique cyclo(LL) ne semble pas être un intermédiaire de réaction nécessaire en ce sens où le tripeptide ELL apparaît d'abord sous forme intacte. Les étapes successives de condensation peuvent être illustrées par le schéma de la Figure IV-21, dont les différentes parties sont expliquées ci-dessous.

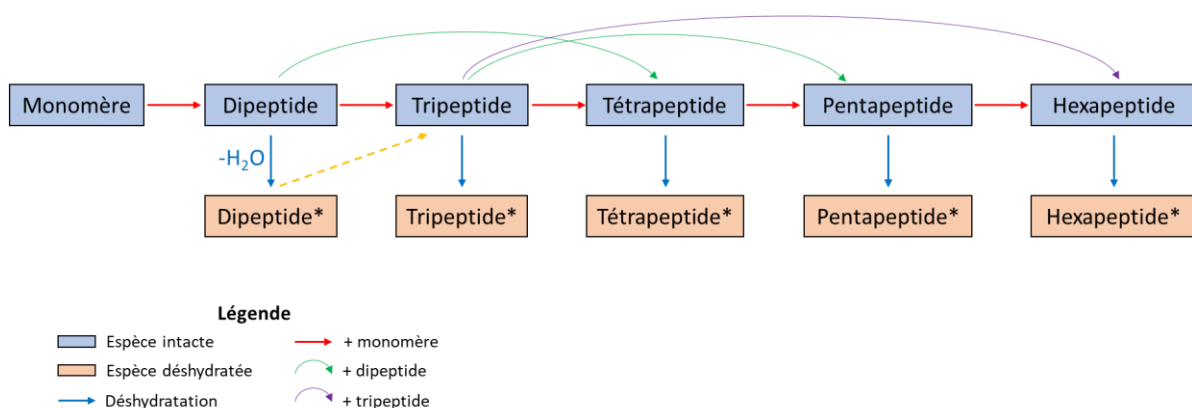


Figure IV-21. Etapes élémentaires proposées pour l'oligomérisation lors de l'activation thermique de l'échantillon (L+E)/SiO<sub>2</sub>.

D'après l'étude cinétique, une apparition successive d'oligopeptides de plus en plus longs a lieu, ce qui suggère une croissance des chaînes par condensation d'un résidu à la fois, visible la première ligne du schéma de la Figure IV-21. Les flèches rouges correspondent ainsi à l'ajout d'un résidu acide aminé. La croissance des chaînes peptidiques « s'essouffle » lorsque le réservoir de monomères libres est épuisé, comme l'indiquent les expériences avec un plateau de longue durée (24 h). Mais elle ne s'arrête pas complètement ; les oligopeptides, du moins les plus courts, peuvent réagir entre eux. Les flèches courbes vertes sur le schéma indiquent les réactions impliquant une condensation avec un dipeptide, comme déterminé expérimentalement par l'étude des systèmes (dipeptide + monomère) ; la flèche courbe violette indique quant à elle la condensation entre deux tripeptides, prouvée par l'étude de l'échantillon (ELL+D)/SiO<sub>2</sub>.

L'étude cinétique a également montré que des espèces déshydratées (notées d'un astérisque, sur la seconde ligne du schéma) étaient formées ultérieurement aux peptides intacts. Pour les échantillons activés à 160 °C, de telles espèces déshydratées sont majoritaires. L'étape de déshydratation est représentée par les flèches bleues indiquant la perte de H<sub>2</sub>O sur la Figure IV-21. Elle apparaît comme irréversible car dans le cas d'une activation thermique simple, le système ne contient plus de molécules d'eau libres, celles-ci étant emportées par le flux gazeux.

Il est important de savoir si les oligopeptides déshydratés (dont les DKP) constituent des impasses évolutives, une question plusieurs fois posée dans la littérature où différentes idées s'opposent. Des études avec activation thermique en conditions sèches considèrent que la formation de DKP est une voie sans issue pour l'élongation de peptides en raison de la stabilité thermodynamique du cycle DKP<sup>67,227</sup>. En revanche, Nagayama et al.<sup>228</sup> ont montré que l'activation d'un DKP(Gly-Gly) en présence de Gly ou d'oligopeptides linéaires, en solution aqueuse dans des conditions hydrothermales, entraîne des réactions d'allongement de chaîne. Dans ces conditions, le DKP n'est pas une impasse mais un intermédiaire de réaction vers l'obtention de chaînes peptidiques plus complexes.

Nos expériences, limitées jusqu'à présent au DKP de formule cyclo(EE), (cf § 1.2.2), montrent que la réaction « DKP + monomère » peut former des tripeptides, bien que de façon moins efficace que la réaction « dipeptide linéaire + monomère ». Cette étape est illustrée par la flèche jaune en pointillés sur le schéma de la Figure IV-21. Par ailleurs, ce schéma ne montre pas de réarrangements interchaînes ; ils sont en effet très minoritaires à température

modérée (130 °C) et restent limités à 160 °C, comme indiqué par les résultats des systèmes (dipeptide + monomère).

Une question importante est de savoir s'il existe des limitations diffusionnelles à la polymérisation. Celles-ci semblent bien présentes dans les systèmes à faible charge en acides aminés, où la distance moyenne entre les réactifs adsorbés est importante. Elles le sont moins lorsqu'on s'approche d'une monocouche. L'étude du dépôt séparé des deux acides aminés sur des particules de silice différentes indique que la diffusion est possible par l'intermédiaire de la phase gazeuse, mais moins efficace que la diffusion de surface. Enfin, l'élongation des peptides lorsqu'on applique des cycles hydratation/déshydratation est sans doute due en partie à la possibilité d'une diffusion des réactifs par l'intermédiaire de la phase aqueuse.

### V.2.2. Signification prébiotique

De nombreuses questions restent ouvertes avant d'arriver à une compréhension complète de la polymérisation des acides aminés sur un support minéral. Nous pouvons d'ores et déjà souligner quelques conséquences des résultats obtenus pour les scénarios de polymérisation des acides aminés en conditions prébiotiques.

Des oligopeptides allant parfois jusqu'aux octapeptides sont détectés après une seule étape d'activation ; ils ne représentent toutefois qu'une minorité des produits formés, la majeure partie étant des di- et tripeptides. Il est possible d'obtenir des peptides plus longs en augmentant la concentration de surface, ou la température de réaction, ce qui rend toutefois les scénarios moins probables – il est difficile de défendre l'idée que des solutions très concentrées d'acides aminés se soient déposées sur des surfaces minérales, et celles-ci n'étaient amenées à des températures de l'ordre de 160 °C que dans des circonstances exceptionnelles. Toutefois, on peut arriver au même résultat en augmentant le temps de réaction, et il est certainement permis d'imaginer des scénarios impliquant une activation de longueur supérieure à 24 heures, le maximum que nous ayons considéré. La direction sans doute la plus prometteuse est l'application de cycles hydratation/déshydratation, capables de favoriser l'élongation de chaînes préformées. Le temps nous a manqué pour étudier la succession de cycles multiples, mais la littérature suggère que la longueur moyenne des peptides pourrait être significativement améliorée *via* cette approche.

La quasi-absence de réarrangements de chaînes peptidiques pour des températures d'activation moyennes signifie que dans ces conditions, une sélectivité de composition acquise pour les peptides de longueur  $n$  sera maintenue pour ceux de longueur  $n+1$  (ou  $n+2$  si on condense un dipeptide, etc). L'évolution du système est probablement orientée dans une direction déterminée par les premières étapes de condensation : la sélectivité est d'origine essentiellement cinétique, en d'autres termes il n'y a pas de réarrangement aléatoire des peptides lors de leur allongement. Cela est en accord avec la discussion basée sur les entropies de Shannon (voir Chapitre III, § II.3). Les conditions idéales d'étude de ces systèmes sont donc des températures basses ( $\leq 130$  °C), des temps de réaction longs ( $\geq 1$  jour) et l'application de cycles hydratation/déshydratation.

Enfin, les conclusions obtenues concernent une surface minérale bien définie, une silice amorphe – et plus particulièrement la silice Aerosil 380. Les résultats préliminaires sur alumine et montmorillonite prouvent bien qu'on ne peut pas généraliser les conclusions obtenues à un support minéral différent. Le monde minéral est d'une complexité plus grande qu'on ne le reconnaît d'ordinaire, et cette complexité oriente son interaction avec les molécules du vivant.



# Chapitre V. Vers une généralisation du scénario de polymérisation sur silice

## I. Choix de l'acide aminé trifonctionnel : étude de différents systèmes binaires (L+X)/SiO<sub>2</sub>

### I.1. Présentation des systèmes binaires étudiés

Nous avons étudié les réactions de polymérisation de plusieurs acides aminés trifonctionnels présentant une fonction trivalente de type (C=O)-X avec X un groupement amine ou hydroxyle sur la chaîne latérale. L'objectif est de voir si, selon la nature de ce groupement X, la formation de produits de polymérisation analogues au cas de (L+E)/SiO<sub>2</sub> est toujours possible. Les échantillons (L+D)/SiO<sub>2</sub>, (L+Q)/SiO<sub>2</sub> et (L+βE)/SiO<sub>2</sub>, D étant l'acide aspartique et Q la glutamine, sont ainsi comparés à l'échantillon modèle (L+E)/SiO<sub>2</sub> étudié dans les Chapitres III et IV. Tous les échantillons (L+X)/SiO<sub>2</sub> (ici X désignant un des AA trifonctionnels testés) ont été réalisés dans les mêmes conditions, avec le protocole classique (voir Chapitre II).

Les structures des acides aminés étudiés sont montrées dans la Figure V-1. L'acide aspartique (Asp) a la même fonction acide que l'acide glutamique mais la chaîne est plus courte. Cette structure défavorise sans doute la cyclisation interne de type pyroAsp (voir Figure V-1). Néanmoins d'autres produits de cyclisation interne peuvent être envisagés (voir discussion ci-dessous).

La glutamine (Gln) est un acide aminé trifonctionnel neutre, ayant de nombreuses similarités avec l'acide glutamique d'un point de vue biologique (voir Chapitre I). Ainsi, l'un peut être formé à partir de l'autre, la transformation étant catalysée par une enzyme, la glutamine synthétase (GS). En outre, la glutamine peut également subir une cyclisation interne pour former un pyroGlu, en perdant une molécule de NH<sub>3</sub>. L'étude de l'échantillon (L+Q)/SiO<sub>2</sub> devrait permettre de mettre en évidence l'origine des formes déshydratées, les pertes de NH<sub>3</sub> et de H<sub>2</sub>O se traduisant par des pertes de masses distinctes.

Le dernier AA trifonctionnel étudié est l'acide β-glutamique (aussi appelé β-glutamate ou acide β-aminoglutarique). En préambule, il convient de souligner que dans la suite, les dénominations αGlu et βGlu ne correspondent pas aux formes polymorphes de l'acide

glutamique solide<sup>229</sup>, mais à deux molécules distinctes, qui diffèrent par le carbone portant le groupement carboxyle. L'acide  $\beta$ -glutamique a un intérêt biologique malgré la faible occurrence des  $\beta$ -acides aminés dans la nature et leur absence au sein des protéines<sup>230</sup>. La glutamine synthétase (GS), qui catalyse la principale réaction métabolique d'utilisation de l'acide glutamique, n'est pas efficace sur la forme bêta. C'est pourquoi l'acide  $\beta$ -glutamique s'accumule comme osmolyte dans plusieurs archées, au lieu d'être métabolisé<sup>231</sup>. La glutamate 2,3-aminomutase est une enzyme permettant la transformation de  $\alpha$ Glu en  $\beta$ Glu<sup>232</sup>. A notre connaissance, il n'y a pas d'information spécifique sur sa présence sur l'ère prébiotique, mais les synthèses de type Miller produisent autant d'acides  $\beta$ -aminés que  $\alpha$ -aminés.



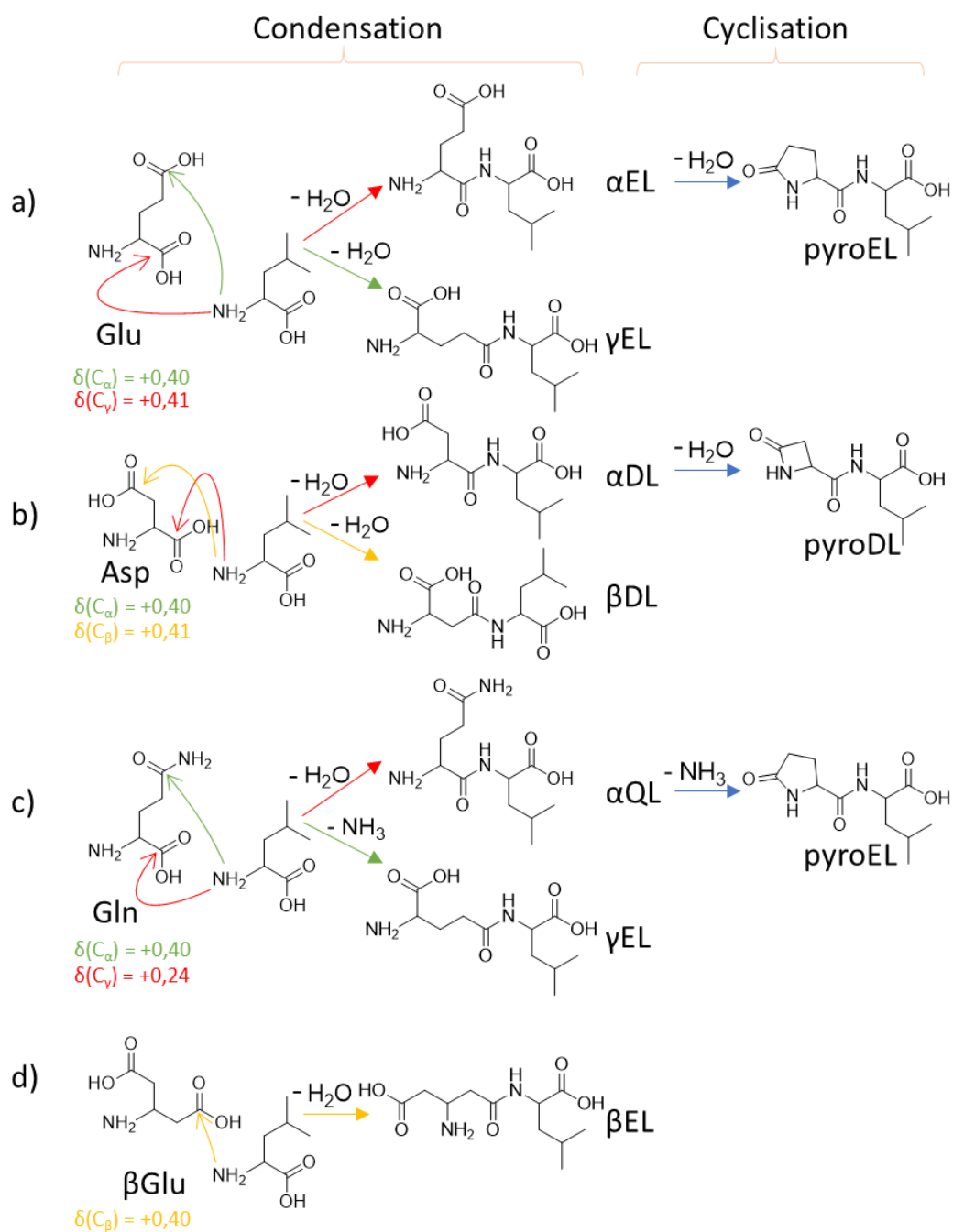


Figure V-1. Structures des acides aminés trifonctionnels testés et des dipeptides de séquence XL formés entre un résidu trifonctionnel X et Leu (une séquence LX est aussi possible). La possibilité de former une espèce déshydratée par cyclisation interne en forme pyroX est également montrée. Les charges partielles des positions électrophiles (valeurs obtenues à partir du logiciel Molcalc) sont indiquées pour chaque acide aminé trifonctionnel.

## I.2. Comparaison de différents acides aminés trifonctionnels

Les échantillons (L+D)/SiO<sub>2</sub> et (L+Q)/SiO<sub>2</sub> ont été analysés par FTICR en injection directe le même jour, et l'échantillon (L+βE)/SiO<sub>2</sub> en utilisant le passeur automatique (mode FIA, voir Chapitre II). Les spectres de masse obtenus sont montrés dans la Figure V-2. Des oligopeptides allant jusqu'aux octapeptides, heptapeptides et hexapeptides sont respectivement détectés pour les échantillons (L+D)/SiO<sub>2</sub>, (L+Q)/SiO<sub>2</sub> et (L+βE)/SiO<sub>2</sub> activés à 160 °C, indiquant une condensation efficace pour les trois acides aminés testés. Les formes déshydratées ([DxLy-H<sub>2</sub>O]H<sup>+</sup> et [ExLy-H<sub>2</sub>O]H<sup>+</sup>) sont majoritaires dans le cas des échantillons (L+D)/SiO<sub>2</sub> et (L+Q)/SiO<sub>2</sub> tout comme elles l'étaient dans l'échantillon de référence. Des expériences ont été également menées pour l'échantillon (L+D)/SiO<sub>2</sub> à une température d'activation plus faible (130 °C). Elles ont démontré que les formes déshydratées sont des produits de dégradation des formes intactes, ces dernières (toutefois de tailles limitées) étant majoritaires à une température de 130 °C.

Cependant, des différences sont constatées entre les systèmes, et comparativement à (L+E)/SiO<sub>2</sub>. En premier lieu, pour les deux systèmes comportant βGlu et Gln, un assemblage des acides aminés par perte de NH<sub>3</sub> est observé. En outre, les peptides formés à partir de βGlu ne présentent que peu de produits de déshydratation (peu d'espèces obtenues par perte de H<sub>2</sub>O et/ou de NH<sub>3</sub> sont détectées, voir discussion ci-dessous).

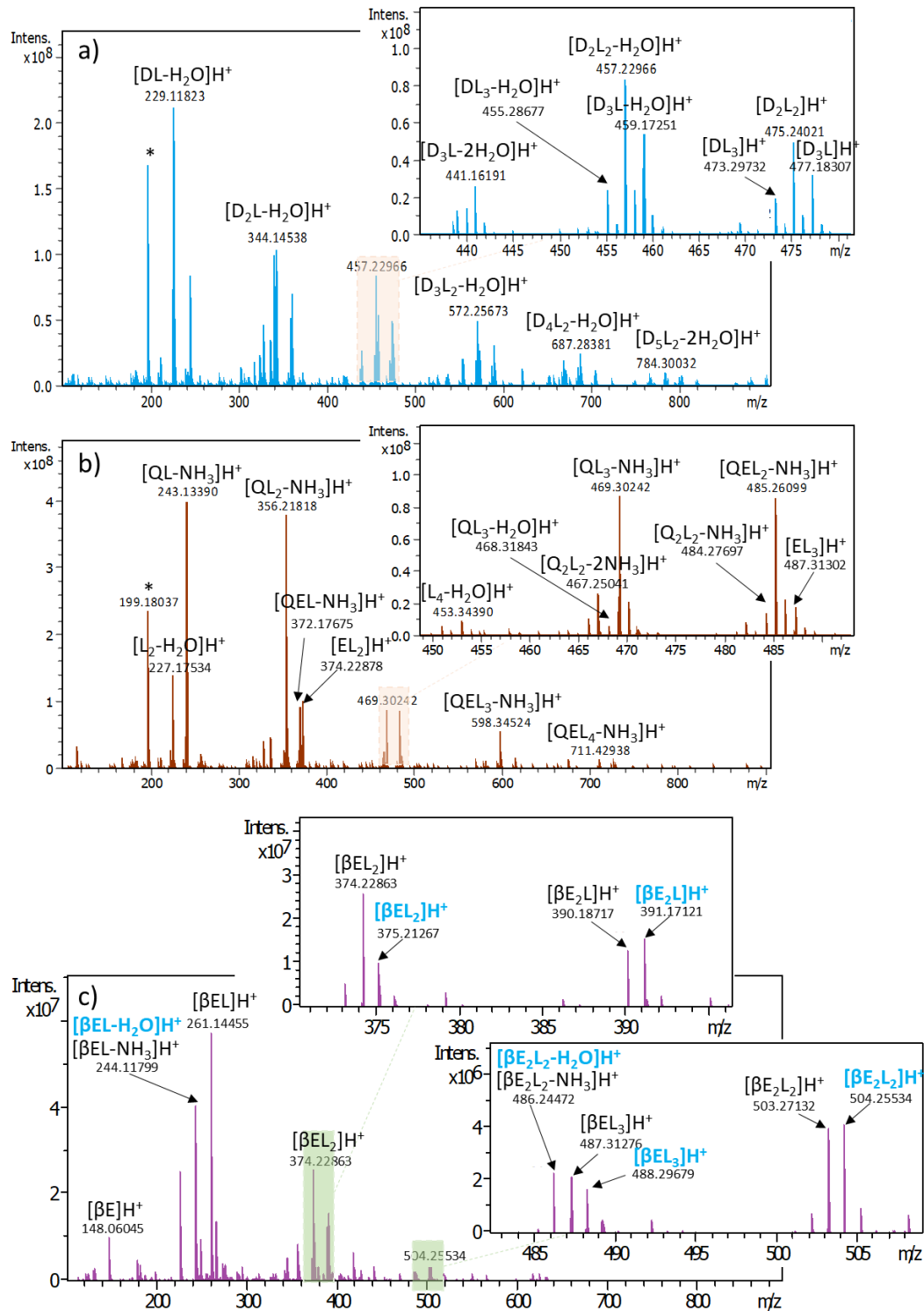


Figure V-2. Spectres de masse ESI obtenus sur FTICR des trois systèmes binaires déposés sur silice et activés à 160 °C. a) (L+D)/SiO<sub>2</sub>, b) (L+Q)/SiO<sub>2</sub> et c) (L+βE)/SiO<sub>2</sub>, sur la gamme m/z 100-900. Un zoom sur la zone des tétrapeptides est montré. Pour le spectre c un zoom sur la zone des tripeptides est ajouté. L'ion à m/z 199 noté avec un astérisque est un contaminant du solvant. Les oligopeptides notés en bleu sont formés par libération de NH<sub>3</sub> et correspondent à un rapport m/z décalé de + 0,98406 u par rapport aux oligopeptides attendus (voir discussion plus bas).

### I.2.1. Système Leu + Asp

Pour l'échantillon (L+D)/SiO<sub>2</sub> activé à 130 °C, des formes déshydratées majoritaires sont observées (voir Figure V-2a). Or, si la cyclisation de Glu en pyroGlu est attendue, en revanche, à notre connaissance, aucune étude ne considère la possibilité de cycliser Asp en pyroAsp (voir Figure V-1), une espèce qui contiendrait un cycle contraint. La déshydratation d'un peptide contenant Asp peut toutefois s'expliquer par la cyclisation d'un résidu acide aspartique sur l'amine secondaire de l'acide aminé suivant pour former un succinimide (voir un exemple, pour une séquence αDL, sur la Figure V-3). Cette déshydratation se produit spontanément dans des conditions physiologiques<sup>233</sup>. Une telle cyclisation interne d'un résidu trifonctionnel sur l'amine du résidu suivant peut, en principe, également avoir lieu dans le cas de systèmes impliquant le résidu Glu, produisant une forme glutarimide (voir la Figure V-4).

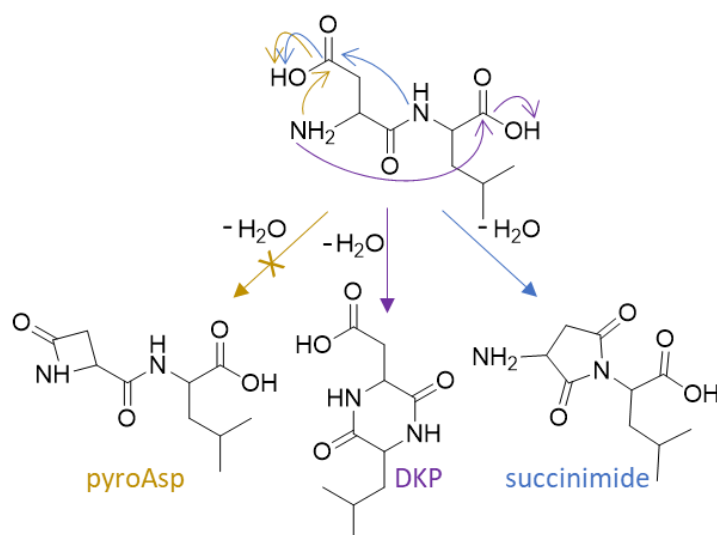


Figure V-3. Mécanismes de cyclisation interne possibles pour un dipeptide αDL. Les voies menant à une forme succinimide et au peptide cyclique (DKP) semblent plus favorables que la forme pyro, contrairement au système avec un résidu Glu.

En outre, en biologie, la réaction inverse se produit spontanément. Le cycle succinimide pouvant s'ouvrir à deux positions, son hydrolyse donne soit un résidu isoaspartate, soit un résidu aspartate (voir Figure V-4), dans un rapport d'environ 3 : 1<sup>234,235</sup>. En biologie, la formation de liaisons peptidiques isoaspartyl est l'une des dégradations non-enzymatiques les plus communes dans les conditions physiologiques<sup>236</sup> pour les peptides et protéines

contenant des résidus d'asparagine ou d'aspartate. Dans nos conditions d'activation, les réactions de condensation sont favorisées, donc les réactions inverses d'hydrolyse avec ouverture de cycle doivent être défavorisées (voir discussion pour Leu + Gln, ci-dessous).

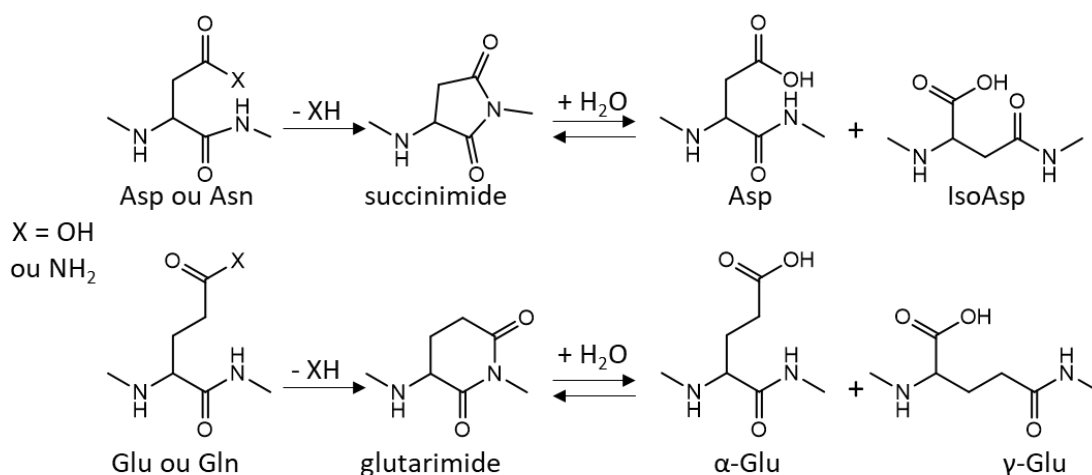


Figure V-4. Réactions de désamination (et/ou déshydratation) de peptides contenant un résidu Asx (Asn ou Asp), Glu ou Gln et réactions d'hydrolyse dans des conditions physiologiques.

Une autre différence apparaît dans la stœchiométrie des oligopeptides de l'échantillon (L+D)/SiO<sub>2</sub> par rapport aux oligopeptides formés dans l'échantillon (L+E)/SiO<sub>2</sub>. Un taux d'incorporation de l'acide aminé trifonctionnel plus important est constaté pour les oligopeptides du système Leu + Asp. Dans leurs travaux précurseurs, Fox et Harada<sup>237,238</sup> avaient également mis en évidence une plus grande incorporation de l'acide aspartique par rapport à l'acide glutamique dans des peptides activés thermiquement (sans surface minérale).

### 1.2.2. Système Leu + Gln

Ce système montre la formation d'espèces polypeptidiques de différentes stœchiométries. Certaines sont attribuables à des espèces contenant le résidu Gln (QxLy, (QxLy-NH<sub>3</sub>) ou (QxLy-H<sub>2</sub>O)), d'autres ne peuvent s'expliquer que par la présence d'un résidu Glu. On trouve ainsi des peptides de formule ExLy, QxEyLz et leurs équivalents partiellement déshydratés/désaminés (voir Figure V-2b).

La présence d'un résidu Glu dans ces stœchiométries peut s'interpréter de différentes façons. Premièrement, on pourrait penser qu'une hydrolyse de Gln en Glu est possible sur la surface de silice. Pour réaliser cette hydrolyse, il faut que de l'eau soit présente dans le système, ce qui n'est le cas qu'avant déshydratation complète de la silice, c'est-à-dire avant 100 °C. Afin d'étudier cette possibilité, nous avons activé thermiquement l'échantillon (L+Q)/SiO<sub>2</sub> à 100 °C. et analysé par spectrométrie de masse le mélange obtenu après extraction. Il n'y a pas de condensation à cette température : seuls des monomères Leu et Gln sont détectés. De plus, aucune trace d'acide aminé Glu n'est présente, suggérant qu'aucune hydrolyse n'a eu lieu dans nos conditions.

On peut exclure également des réactions consécutives de cyclisation-désamination et d'hydrolyse, semblables aux modifications post-traductionnelles rencontrées fréquemment dans les conditions physiologiques et illustrées en Figure V-4, et cela pour la même raison. En effet, dans nos conditions d'activation, la silice est déjà déshydratée lorsque la température devient assez élevée pour catalyser les réactions d'hydrolyse, ces dernières ne sont donc *a priori* pas envisageables.

Deuxièmement, nous avons vérifié que l'hydrolyse Gln→Glu n'a pas lieu lors de l'étape d'extraction par désorption en solution aqueuse. Nous avons effectué une désorption, non pas dans de l'eau, mais dans du méthanol. Les mêmes espèces étant détectées sur le spectre de masse, cette hypothèse a également été rejetée.

La seule explication envisageable à la présence d'espèces de stœchiométrie ExLy est donc la condensation entre la fonction amine d'un résidu et la fonction CO-NH<sub>2</sub> en position  $\gamma$  d'un autre résidu Gln. Cette condensation implique la libération d'une molécule de NH<sub>3</sub> et la formation de  $\gamma$ -peptides (voir l'exemple du dipeptide  $\gamma$ EL de la Figure V-1c). L'étude des charges partielles indique une différence d'électronégativité entre les deux fonctions trivalentes mais elle va dans le sens d'une régiosélectivité pour les formes  $\alpha$  lors de la condensation, à l'inverse de nos observations expérimentales. En revanche, l'amine est considérée comme étant un meilleur nucléofuge que l'eau. Il est donc possible que des peptides se forment préférentiellement par attaque du carbone  $\gamma$  d'un Gln, conduisant aux espèces de stœchiométrie ExLy (voir Figure V-2c). L'utilisation de Gln est donc une voie intéressante pour discriminer les différents régioisomères, les peptides  $\alpha$  et les peptides  $\gamma$  n'ayant pas la même masse.

Il faut souligner que la stœchiométrie d'une espèce ( $QL_\gamma-NH_3$ ) est exactement la même que celle de ( $EL_\gamma-H_2O$ ) (à titre d'illustration, l'espèce correspondant au pic à  $m/z$  243.13390 peut s'écrire comme  $[EL-H_2O]H^+$  ou  $[QL-NH_3]H^+$ ). En revanche, les espèces non déshydratées telles que  $ExLy$  et  $QxLy$  ne sont attribuables qu'à une stœchiométrie. Au vu des intensités de ces espèces observées sur le spectre de masse (Figure V-2b), il semble que la formation de peptides  $\gamma$  (équivalents aux stœchiométries  $ExLy$ ) soit plus favorable que celle des peptides  $\alpha$  (équivalents aux stœchiométries  $QxLy$ ). En effet, le pic de  $[EL_2]H^+$  est dix fois plus intense que celui de  $[QL_2]H^+$ . De même, celui de  $[EL]H^+$  est cinq fois plus intense que celui de  $[QL]H^+$ . A la lumière de ces observations, nous pouvons postuler que les espèces attribuables à deux stœchiométries ( $(QL_\gamma-NH_3)$  et  $(EL_\gamma-H_2O)$ ) sont plutôt produites à partir de forme  $ExLy$ .

Que les espèces soient de stœchiométrie  $QxLy$ ,  $ExLy$  ou  $QxEyLz$ , des formes correspondantes déshydratées (et/ou désaminées) sont observées. Les espèces majoritaires de formule ( $QxLy-NH_3$ ) sont compatibles avec la formation d'un résidu pyroGlu par cyclisation de la glutamine à la polymérisation. En effet, en biochimie, la glutamine, surtout quand elle est incorporée en position N-terminale d'une chaîne peptidique, a tendance à former pyroGlu par perte de  $NH_3$  (voir Chapitre I, et Figure V-1c). Ces dégradations peuvent également être dues à différentes réactions de cyclisation interne au sein de la chaîne peptidique, comme montré pour le tripeptide  $[QEL_3-NH_3]H^+$  sur la Figure V-5.

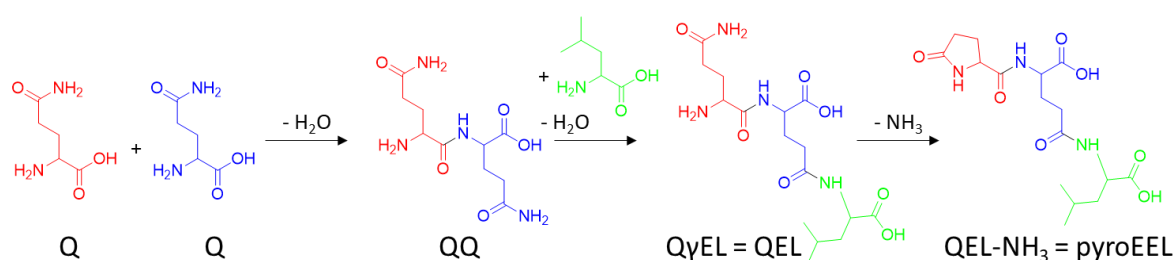


Figure V-5. Exemple des réactions successives de condensation et de cyclisation interne produisant une espèce  $[QEL-NH_3]H^+$ , détectée à 372 u.

### I.2.3. Système Leu + $\beta$ Glu

Nous avons décidé d'étudier un autre type d'acide aminé, *a priori* non protéinogène, l'acide  $\beta$ -glutamique. Dans cette molécule, les groupements -COOH sont portés par les carbones en  $\beta$  et  $\gamma$  de l'amine, et non  $\alpha$  et  $\gamma$  comme dans les acides aminés étudiés précédemment (et tous les acides aminés protéinogènes). La molécule est symétrique et n'a donc ni forme L ni forme D, ni la possibilité de former différents régioisomères. A notre connaissance, aucun travail de recherche n'a montré la possibilité de former un lactame par cyclisation interne d'un acide aminé  $\beta$ Glu. Comme dans le cas pyroAsp, il s'agirait d'un cycle à quatre atomes de carbone, donc sans doute instable.

Dans l'échantillon (L+ $\beta$ E)/SiO<sub>2</sub>, on observe expérimentalement des produits de polymérisation allant jusqu'aux hexamères. A l'inverse de ce qui est observé pour les autres systèmes, les espèces intactes (stœchiométrie E<sub>x</sub>L<sub>y</sub>) sont majoritaires. Plus important, deux ensembles de pics sont détectés. Certains correspondent à la stœchiométrie attendue (de composition E<sub>x</sub>L<sub>y</sub>), par condensation d'un groupement amine et d'un groupement acide carboxylique, et ont bien une liaison peptidique. D'autres espèces sont détectées à un rapport m/z décalé de +0,98406 u par rapport aux oligopeptides classiques (Figure V-2c, espèces notées en bleu)<sup>††</sup>. Or, cette valeur correspond à la différence de masse entre H<sub>2</sub>O et NH<sub>3</sub>. Elles pourraient être formées par une attaque de l'amine d'un acide aminé sur le carbone en  $\alpha$  du groupement amine du second résidu. Dans l'exemple de la Figure V-6, le dimère linéaire ainsi formé  $\beta$ EL + 0,98 (détecté sous forme protonée à m/z 262) possède trois fonctions acides (contrairement à un dipeptide EL classique qui en possède deux). Il ne s'agit pas d'un peptide car il n'y a pas de liaison amide -CO-NH-.

---

<sup>††</sup> Par soucis de clarté ces espèces seront notées +0,98 dans la suite du manuscrit.



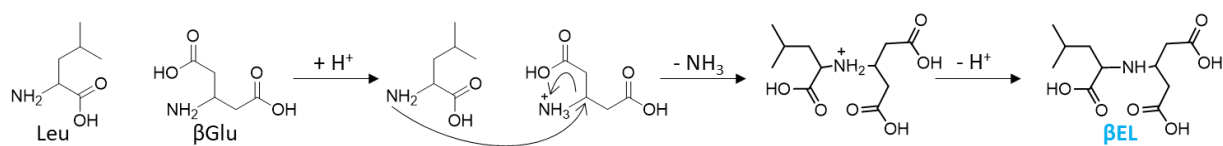


Figure V-6. Hypothèse de mécanisme pour la formation d'un oligopeptide de stœchiométrie  $\beta EL + 0,98$  par libération de  $NH_3$ . L'espèce protonée correspondante est annotée sur le spectre de masse  $[\beta EL]H^+$  à  $m/z$  262. La condensation classique menant à une liaison peptidique est visible en Figure V-1d.

Deux réactions bien connues donnent lieu à une élimination d'amine, l'élimination de Hofmann<sup>239</sup> et la substitution d'un ammonium en phase gazeuse<sup>240</sup>, mais dans aucun des cas l'amine n'est en compétition avec un acide carboxylique. Pour un acide aminé en solution aqueuse, cette réaction n'est pas compétitive, le carbone  $\alpha$  étant moins électrophile que les deux carboxyles. En revanche, avec une activation de l'amine par un catalyseur acide de Brønsted, une telle addition peut-être envisageable,  $NH_3$  étant un bon groupement partant. Or, il est important de noter ici que l'acide aminé  $\beta Glu$  étant très peu soluble, quelques gouttes d'acide formique ont été ajoutées pour permettre la dissolution du mélange  $Leu + \beta Glu$  en solution aqueuse avant dépôt sur silice. La différence de polymérisation entre  $\alpha Glu$  et  $\beta Glu$  peut donc s'expliquer par l'ajout d'acide formique, qui entraîne une catalyse acide, à moins que celle-ci ne soit le fait de la surface de silice, catalyseur hétérogène faiblement acide.

Si notre hypothèse est correcte, la formation des oligopeptides  $ExLy + 0,98$  serait théoriquement possible aussi pour une condensation à partir de  $\alpha Glu$ , mais elle n'est pas observée. Cela peut être dû à l'absence d'acide formique additionnel dans le système, ou au fait que seule la structure symétrique de  $\beta Glu$  pourrait donner lieu à des interactions spécifiques de cette espèce avec la surface.

Les longueurs des oligopeptides formés sont semblables entre les deux systèmes  $Leu + \alpha Glu$  et  $Leu + \beta Glu$  (pour une analyse le même jour dans les mêmes conditions). En revanche, la condensation avec  $\beta Glu$  forme des oligopeptides plus riches en résidus  $\beta Glu$ .

### I.3. Comparaison de la réactivité des acides aminés trifonctionnels

Si la formation d'oligopeptides a été observée dans tous les systèmes étudiés, certaines différences ont été constatées au niveau des formes produites. La question alors soulevée est la suivante : les acides aminés trifonctionnels présentent-ils une réactivité plus ou moins importante selon leur structure ? Pour y répondre, nous avons mené une polymérisation en mélangeant, en présence de Leu, les AA trifonctionnels deux à deux de façon à révéler une éventuelle sélectivité de polymérisation pour un des deux résidus trifonctionnels en compétition. Pour ce faire, nous avons à nouveau utilisé des AA marqués (ici Glu\*, voir Chapitre III) afin de comparer les réactivités respectives des deux isomères Glu et  $\beta$ -Glu. Nous avons testé quatre échantillons : L+(E\*+ $\beta$ E)/SiO<sub>2</sub>, L+(D+E)/SiO<sub>2</sub>, L+(D+ $\beta$ E)/SiO<sub>2</sub> et L+(E\*+Q)/SiO<sub>2</sub>. On peut considérer que Glu et Glu\* ont la même réactivité (les éventuels effets isotopiques sont *a priori* mineurs, voir Chapitre III). Sachant que dans le cas du système Leu + Gln, les oligopeptides de type  $\gamma$  ont une stœchiométrie ExLy, nous avons choisi, afin d'éviter toute confusion, d'utiliser un Glu marqué avec un atome <sup>13</sup>C pour comparer les réactivités de Glu et Gln.

Nous avons, en parallèle, comparé les efficacités d'ionisation des quatre monomères Gln,  $\beta$ Glu, Glu et Asp en utilisant la méthode cinétique de Cooks (voir Annexe 10). L'ordre relatif des efficacités d'ionisation est le suivant : Gln >  $\beta$ Glu > Glu > Asp. Il correspond à l'ordre des affinités protoniques établi dans la littérature<sup>241-243</sup>. *A priori*, les peptides correspondants présenteront des efficacités d'ionisation suivant cet ordre, bien que d'autres aspects comme les conformations secondaires entrent en compte pour expliquer la réponse MS des différents oligopeptides.

#### I.3.1. Comparaison des réactivités de Glu, $\beta$ Glu et Asp

Pour comparer les réactivités de ces acides aminés trifonctionnels qui diffèrent par un CH<sub>2</sub> (entre Glu et Asp) et par la position de la chaîne latérale (entre Glu et  $\beta$ Glu), les échantillons L+(D+E)/SiO<sub>2</sub> et L+(D+ $\beta$ E)/SiO<sub>2</sub> ont été étudiés. Pour les deux, un zoom sur la zone des tétrapeptides est montré dans la Figure V-7. Comme pour l'échantillon (L+ $\beta$ E)/SiO<sub>2</sub>, des espèces à +0,98 u sont présentes : la réaction de condensation particulière de  $\beta$ Glu se

produit dans ce système également. Des espèces contenant les trois acides aminés (hétéropeptides  $D_x\beta E_yL_z$ ) sont également détectées. Dans les deux cas, les espèces contenant Asp et celles contenant Glu ou  $\beta$ Glu ont sensiblement les mêmes intensités. Or, le résidu Asp ayant une moins bonne efficacité d'ionisation que Glu et  $\beta$ Glu, l'acide aspartique présenterait plutôt une importante efficacité de réaction dans nos conditions d'activation.

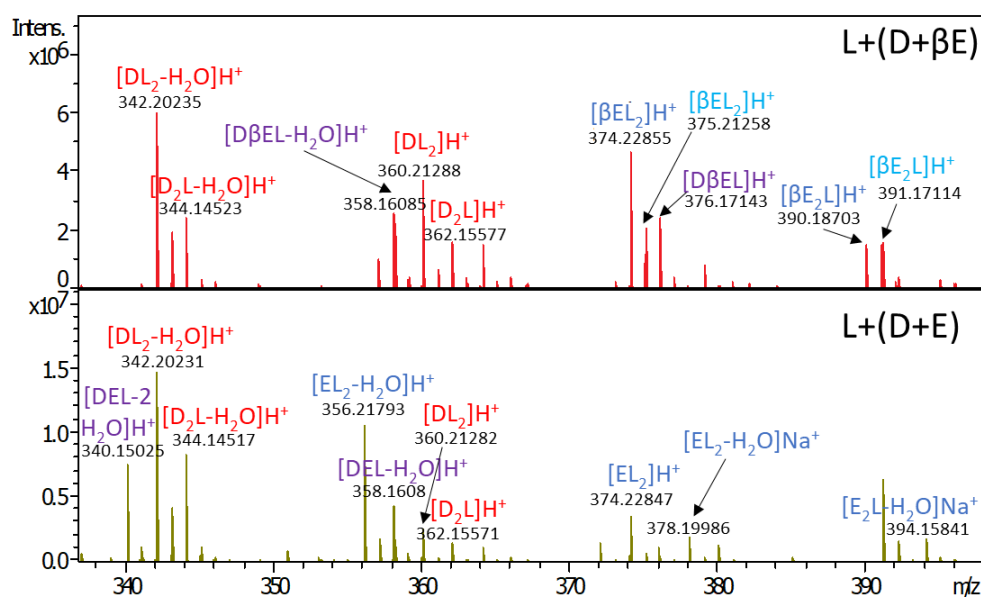


Figure V-7. Zoom sur la zone des tripeptides pour les échantillons L+(D+βE)/SiO<sub>2</sub> et L+(D+E\*)/SiO<sub>2</sub> (ESI positif, instrument FTICR). Les espèces contenant respectivement Asp, (β)Glu ou mixtes sont en rouge, bleu et violet. Les espèces en bleu clair correspondent à celles formées par perte de NH<sub>3</sub>.

L'analyse de ces deux échantillons ne permet pas de déterminer qui de Glu ou  $\beta$ Glu est le plus réactif. En effet, les oligopeptides contenant Glu sont majoritairement sous forme déshydratée tandis que ceux contenant  $\beta$ Glu sont des oligopeptides sous forme intacte. Pour comparer les réactivités de Glu et  $\beta$ Glu, l'échantillon L+(E\*+βE)/SiO<sub>2</sub> a donc été étudié, la réactivité de Glu étant assimilée à celle de Glu\*. Le spectre de masse ESI obtenu correspond strictement à celui de l'échantillon (L+βE)/SiO<sub>2</sub>, à l'exception de deux pics à m/z 249.14774 et m/z 362.23178 que nous n'avons pas réussi à annoter (marqués d'un astérisque sur la Figure V-8). Aucun pic correspondant à des oligopeptides formés à partir du résidu Glu\* n'est observé, suggérant une sélectivité beaucoup plus importante pour  $\beta$ Glu : L-Leu préfère se condenser avec  $\beta$ Glu qu'avec  $\alpha$ Glu.

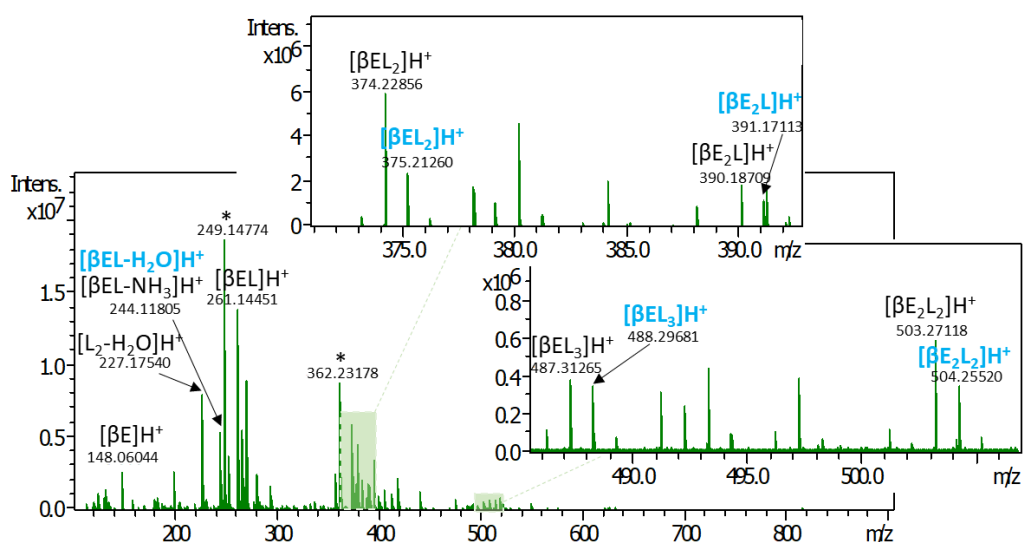


Figure V-8. Spectre de masse de l'échantillon  $L+(E^*+\beta E)/SiO_2$  (ESI positif, instrument FTICR). Les deux pics à  $m/z$  249 et  $m/z$  362 annotés d'un astérisque n'ont pas pu être attribués. Les autres pics sont les mêmes que ceux de l'échantillon  $(L+\beta E)/SiO_2$ .

Par ailleurs, un échantillon composé uniquement de Glu\* et de  $\beta$ Glu en proportions équimolaires a été analysé. Dans ce cas, peu d'oligopeptides sont formés (au maximum un tripeptide) et ils ne correspondent qu'à des résidus  $\beta$ Glu. Il faut noter que des oligopeptides formés par libération de  $NH_3$  sont également détectés à  $+0,98$  u. Cela confirme que dans le cas d'un mélange contenant Glu et  $\beta$ Glu (avec ou sans Leu), seul  $\beta$ Glu entre en jeu dans la polymérisation. Cette observation pourrait s'expliquer par une interaction plus importante de  $\beta$ Glu avec la surface.

Ces trois échantillons permettent d'établir l'ordre de réactivité suivant : Asp >  $\beta$ Glu >  $\alpha$ Glu. La situation est toutefois complexe puisque  $\beta$ Glu exclut pratiquement  $\alpha$ Glu de la polymérisation quand ils sont présents dans le même échantillon. En revanche, dans les échantillons où ils sont en présence de Asp, ils sont à peu près aussi incorporés l'un que l'autre.

### I.3.2. Comparaison des réactivités de Glu et Gln

Pour comparer les réactivités de Glu et Gln, l'échantillon  $(L+(E^*+Q))/SiO_2$  a été étudié. Nous avons montré que, conformément à la littérature<sup>241-243</sup>, l'acide aminé Gln a une meilleure efficacité d'ionisation que Glu. Donc dans le cas où les deux acides aminés auraient la même

réactivité, on s'attend à détecter des signaux plus intenses pour les oligopeptides formés à partir de Gln. La polymérisation de cet échantillon entraîne la formation d'oligopeptides de longueurs et stœchiométries variées, y compris la formation d'oligopeptides mixtes. Sur la Figure V-9 montrant un zoom sur le tétrapeptide  $[E_2L_2-H_2O]H^+$ , on observe que cette espèce a la même probabilité de contenir un résidu venant de Glu (et donc marqué isotopiquement) qu'un résidu venant de Gln. Ce dernier est noté  $\gamma E$ , car la stœchiométrie correspond à la perte d'une molécule de  $NH_3$ , donc sûrement à la formation d'un  $\gamma$ -oligopeptide (voir discussion du § 1.2). En effet, la distribution des différents tétrapeptides  $[E_2L_2-H_2O]H^+$  est approximativement 1 : 2 : 1, ce qui correspond à la répartition aléatoire. Comme attendu, l'intensité de l'espèce  $[E^*_2L_2-H_2O]H^+$  est légèrement supérieure à celle de  $[\gamma E_2L_2-H_2O]H^+$  (formée à partir de résidu Gln) car l'intensité de l'espèce  $E^*$  s'additionne au massif isotopique naturel de l'espèce contenant  $E$  (voir Chapitre III). Finalement, une réactivité équivalente est démontrée entre les systèmes avec Glu et Gln.

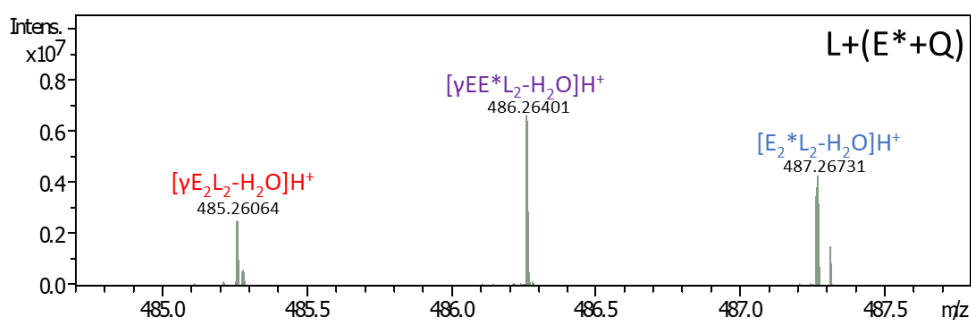


Figure V-9. Zoom du spectre de masse des produits d'activation de l'échantillon  $L+(E^*+Q)/SiO_2$  sur l'espèce  $[E_2L_2-H_2O]H^+$  à  $m/z$  485 (ESI positif, instrument FTICR).

#### I.4. Conclusion sur les systèmes binaires

Tous les acides aminés trifonctionnels au sein des systèmes binaires étudiés ont, *a priori*, une réactivité proche : ils donnent lieu à la formation d'oligopeptides intacts et de leurs produits de déshydratation. En raison de la fonctionnalité de leur chaîne latérale, ils sont sujets à plusieurs modifications spontanées telles que des réactions de déshydratation ou de désamination. Cependant, des différences significatives sont observées entre les systèmes.

Tout d'abord, en termes d'efficacité de condensation, l'ordre suivant de réactivité peut être établi sur base des comparaisons par couples : Asp > βGlu > Glu = Gln.

Le cas de βGlu est clairement différent des autres puisque la formation de liaisons peptidiques y est en concurrence avec une réactivité originale aboutissant à la perte de NH<sub>3</sub>. Les synthèses prébiotiques comme celles de Miller, produisent une quantité importante d'acides β-aminés en plus des acides α-aminés<sup>6</sup>. Les résultats obtenus avec βGlu laissent penser que la présence de ceux-ci pourrait être problématique puisqu'elle cause la formation de polymères contenant des liaisons non peptidiques.

En ce qui concerne la réactivité supérieure de Asp par rapport à Glu, on peut noter que Glu a une enthalpie de sublimation plus élevée que celle de Asp<sup>244</sup>. Il aurait donc plus de difficulté à diffuser par l'intermédiaire de la phase gazeuse. Toutefois, cette remarque concerne la sublimation à partir d'acides aminés purs et non sur silice. Une autre explication à la plus grande incorporation du résidu Asp par rapport au résidu Glu aurait pu être que la cyclisation du monomère Glu en pyroGlu rende cet acide aminé plus inerte, tandis qu'aucune cyclisation n'est possible pour un monomère Asp. Toutefois, cette idée est contredite par les résultats obtenus en faisant réagir directement l'acide pyroglutamique avec la leucine (Chapitre IV, § 1.2.2). Ceux-ci indiquent que la cyclisation du résidu Glu en pyroGlu ne diminue pas sa réactivité.

Tous les résultats discutés jusqu'ici portent sur l'effet d'une substitution de l'acide aminé trifonctionnel. Nous avons aussi modifié l'acide aminé aliphatique en testant un échantillon (V+E)/SiO<sub>2</sub> (voir résultats en Annexe 11). Nous avons constaté que la substitution de Val à Leu ne modifie ni les stœchiométries ni la taille des oligopeptides produits.

Les systèmes que nous avons discutés jusqu'à présent sont binaires ou ternaires. L'activation de systèmes plus complexes est pertinente si l'on veut se rapprocher des conditions prébiotiques réelles, où plusieurs acides aminés étaient simultanément présents en solution (« soupe primordiale »).

## II. Etude de systèmes plus complexes : polymérisation des mélanges GAVD et GAVE

Nous avons choisi d'étudier le système GAVD (glycine (Gly, G), alanine (Ala, A), valine (Val, V), acide aspartique (Asp, D)) car ces quatre acides aminés sont les plus abondants dans la météorite de Murchison et dans l'expérience de synthèse prébiotique par décharge électrique<sup>156</sup>. De façon intéressante, ils sont aussi suggérés comme primordiaux par une approche « top-down »<sup>157</sup> qui reconstruit le protéome ancien par comparaison de séquences conservées dans les protéines (voir Chapitre I). L'intérêt de cette étude est également de comparer la réactivité de différents AA aliphatiques dans un même système. En outre, la polymérisation de GAVD/SiO<sub>2</sub> a été comparée au système GAVE/SiO<sub>2</sub> (substitution Asp/Glu, cf Chapitre V, § I.2.1). La Figure V-10 montre les spectres de masse obtenus pour ces deux systèmes.

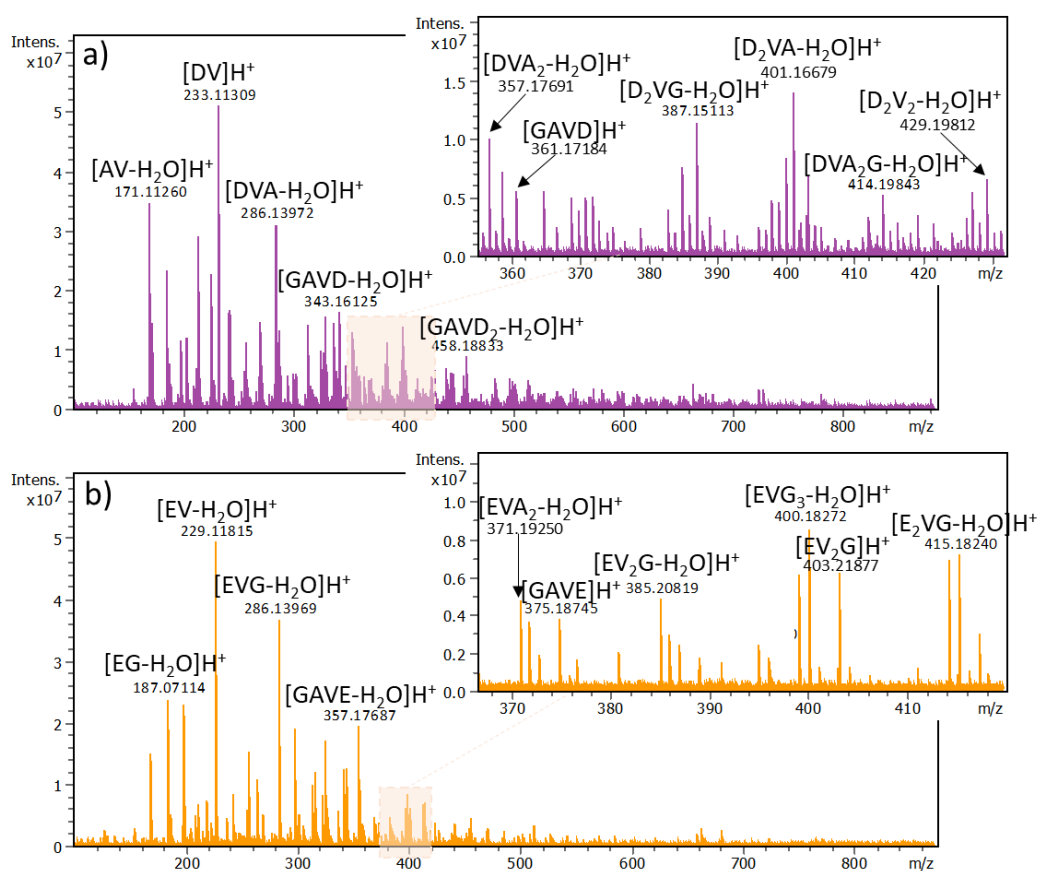


Figure V-10. Spectres de masse des échantillons a) GAVD/SiO<sub>2</sub> et b) GAVE/SiO<sub>2</sub> (ESI positif, instrument FTICR). Un zoom sur les tétrapeptides est montré pour chacun.

Les deux systèmes GAVD et GAVE forment, par activation thermique à 160°C, des oligopeptides. Les plus longs détectés dans nos conditions FTICR sont des nonapeptides pour GAVD et des heptapeptides pour GAVE. Une grande variété de stœchiométries est observée à partir du mélange d'acides aminés, dont la plupart contiennent un résidu Asp ou Glu (voir discussion ci-dessous). Comme observé précédemment pour les systèmes binaires, les formes déshydratées sont majoritaires, même pour le système GAVD (pour la nature de ces formes, voir § I.2.1 et Chapitre VI).

## II.1. Réactivité similaire des deux systèmes : le rôle essentiel du résidu trifonctionnel

### II.1.1. Evaluation de la réactivité des systèmes GAVX (X=E/D) : utilisation des représentations de Van Krevelen et de Kendrick

Les spectres de masse obtenus étant relativement complexes, nous avons choisi, pour comparer les deux ensembles de peptides obtenus, de recourir à des représentations basées sur la détection en ultra haute résolution et les mesures de masses précises. Les approches basées sur les masses précises, telles que les analyses dites de Van Krevelen, de Kendrick ou l'utilisation de filtres de défaut de masse<sup>246,247</sup>, permettent une visualisation et une comparaison plus claires des données complexes fournies par une analyse en haute résolution. Dans le présent travail, les méthodes de Van Krevelen et de Kendrick ont été utilisées.

#### II.1.1.1. Analyse de Van Krevelen

Les diagrammes de Van Krevelen ont été développés pour caractériser des mélanges complexes en fonction des compositions élémentaires, et plus précisément des rapports atomiques entre deux éléments tels que O/C, N/C ou H/C. La connaissance des formules brutes des composés est requise. Ces diagrammes sont utilisés dans plusieurs domaines scientifiques. Ils permettent de visualiser la diversité des composés d'un mélange complexe,



par exemple des échantillons de matière organique naturelle (« NOM » : natural organic matter)<sup>138,248</sup>, de pétrole et charbon<sup>248,249</sup>, mais également de vins<sup>250</sup>, de métabolites de plantes<sup>251</sup> ou encore en chimie interstellaire pour étudier la composition de météorites<sup>141,252</sup>.

Les diagrammes de Van Krevelen classiques représentent les rapports O/C ou N/C en fonction du rapport H/C. Le rapport H/C sépare les composés selon leur nombre d'insaturations tandis que les rapports O/C et N/C classent les composés selon leur teneur en O et N respectivement. Ainsi, les composés d'une même classe seront regroupés sur une même ligne/diagonale. On peut expliquer l'existence de telles tendances dans un diagramme de Van Krevelen. Pour la simplicité, le principe est ici expliqué pour l'échantillon (L+E)/SiO<sub>2</sub>, bien que de telles méthodes soient plus adaptées à des systèmes complexes comme GAVE ou GAVD, systèmes pour lesquels nous les utilisons.

Dans l'échantillon (L+E)/SiO<sub>2</sub>, tous les oligopeptides intacts de longueur n ont une formule E<sub>m</sub>L<sub>n-m</sub>, avec 0 ≤ m ≤ n. Le nombre d'atomes d'azote est n (il y a un atome d'azote par résidu acide aminé). Le nombre d'atomes de carbone est 5m + 6(n-m) = 6n - m (5 par résidu Glu et 6 par résidu Leu). Le nombre d'atomes d'oxygène est 4m + 2(n-m) - (n-1) = 1+n+2m (4 par résidu Glu, 2 par résidu Leu, moins un pour chaque liaison peptidique formée).

Donc  $\frac{N}{C} = \frac{n}{6n-m}$ . En résolvant pour n, on trouve  $m = n(6 - \frac{1}{N})$ . D'autre part,  $\frac{O}{C} = \frac{1+n+2m}{6n-m}$ .

D'où l'on tire  $\frac{O}{C} = \frac{N}{C} \frac{1+n+2m}{n} = \frac{N}{C} \frac{1+n}{n} + \frac{N}{C} \frac{2m}{n} = -2 + \frac{N}{C} (12 + \frac{1+n}{n})$ . Sur un graphique portant O/C en fonction de N/C, les points correspondants aux oligopeptides d'une même longueur seront alignés de manière à former une droite (dont la pente sera d'autant plus faible que n est élevé), sur laquelle apparaîtront les différentes stœchiométries E<sub>m</sub>L<sub>n-m</sub> avec des substitutions Glu/Leu croissantes. Toutes les droites convergent vers la même ordonnée à l'origine qui vaut -2. Pour tout autre couple d'acides aminés, on obtiendra des dépendances linéaires, mais avec des pentes différentes selon le système étudié.

Des diagrammes de Van Krevelen dits modifiés ont également été utilisés, notamment en tenant compte des rapports m/z. En effet des ions avec des m/z très différents peuvent avoir le même rapport H/C ou N/C. Ainsi des diagrammes 2D de H/C vs m/z ou N/C vs m/z ont aussi été utilisés pour l'étude de la composition d'aérosols organiques de Titan<sup>252,253</sup> ou de

depsipeptides<sup>144</sup>. Des représentations en trois dimensions selon  $m/z$  peuvent être proposées afin de comparer des échantillons complexes<sup>254,255</sup>. Plusieurs représentations de Van Krevelen ont été testées pour nos systèmes. Le diagramme représentant les rapports O/C en fonction de N/C, montré à la Figure V-11, permet de mettre en évidence la similarité de réactivité entre les deux systèmes. GAVD montre des espèces à des rapports O/C et N/C plus importants que GAVE. Cela était attendu car pour un même taux d'incorporation de l'AA trifonctionnel X, le rapport O/C est plus important dans le cas de Asp car ce dernier possède moins de carbones que Glu pour un même nombre d'oxygènes (à titre d'exemple les coordonnées des tétrapeptides  $[GAE_2-H_2O]H^+$  et  $[GAD_2-H_2O]H^+$  sur le diagramme sont respectivement (0,267 ; 0,533) et (0,222 ; 0,615)). Mais ces observations reflètent aussi que les oligopeptides formés sont plus riches en Asp dans le système GAVD qu'en Glu dans le système GAVE. Une interprétation plus poussée des représentations de Van Krevelen nécessiterait d'établir les liens entre les points d'une même ligne/diagonale, ou d'une même zone par rapport à la composition chimique (taille des oligopeptides ou encore des formes intactes ou déshydratées). L'exemple détaillé plus haut montre qu'il n'est pas évident de définir *a priori* ces corrélations.

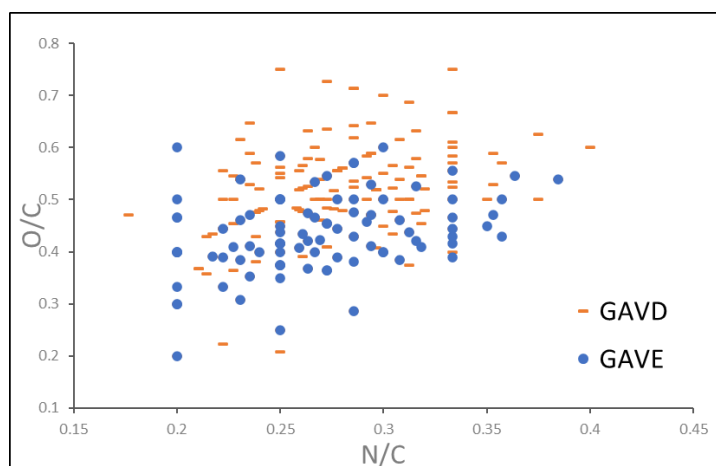


Figure V-11. Diagramme de Van Krevelen des échantillons GAVE et GAVD. Le rapport O/C est porté en fonction de N/C.

Le diagramme portant le rapport N/H en fonction de  $m/z$  (voir Figure V-12), sans doute plus facile à interpréter visuellement, met en évidence des courbes correspondant chacune à une longueur d'oligopeptide ayant un nombre défini d'acides aminés trifonctionnels (Glu ou

Asp). Cette figure est intéressante car elle permet de déduire facilement, à partir de la composition élémentaire, la composition en résidus des oligopeptides en fonction de la courbe sur laquelle ils apparaissent, ce qui facilite leur annotation. On observe que les formes intactes et déshydratées n'apparaissent pas exactement sur la même courbe. Pour une même espèce, la forme déshydratée se trouve décalée d'un vecteur  $(-18,0106 ; 0,0128)$  par rapport à la forme intacte (voir les flèches rouges sur la Figure V-12b). Au sein d'une même série (oligopeptides de même longueur ayant le même nombre de résidus trifonctionnels), deux points situés au même rapport N/H (même position sur l'axe vertical) et séparés horizontalement de  $-3,995$  u correspondent à une substitution d'un résidu Gly par un résidu Ala, accompagnée d'une déshydratation (voir exemple dans le zoom de la Figure V-12b, les espèces numérotées 1 et 2 correspondent respectivement aux espèces  $[D_2VA_2-2H_2O]H^+$  et  $[D_2VAG-H_2O]H^+$ ).

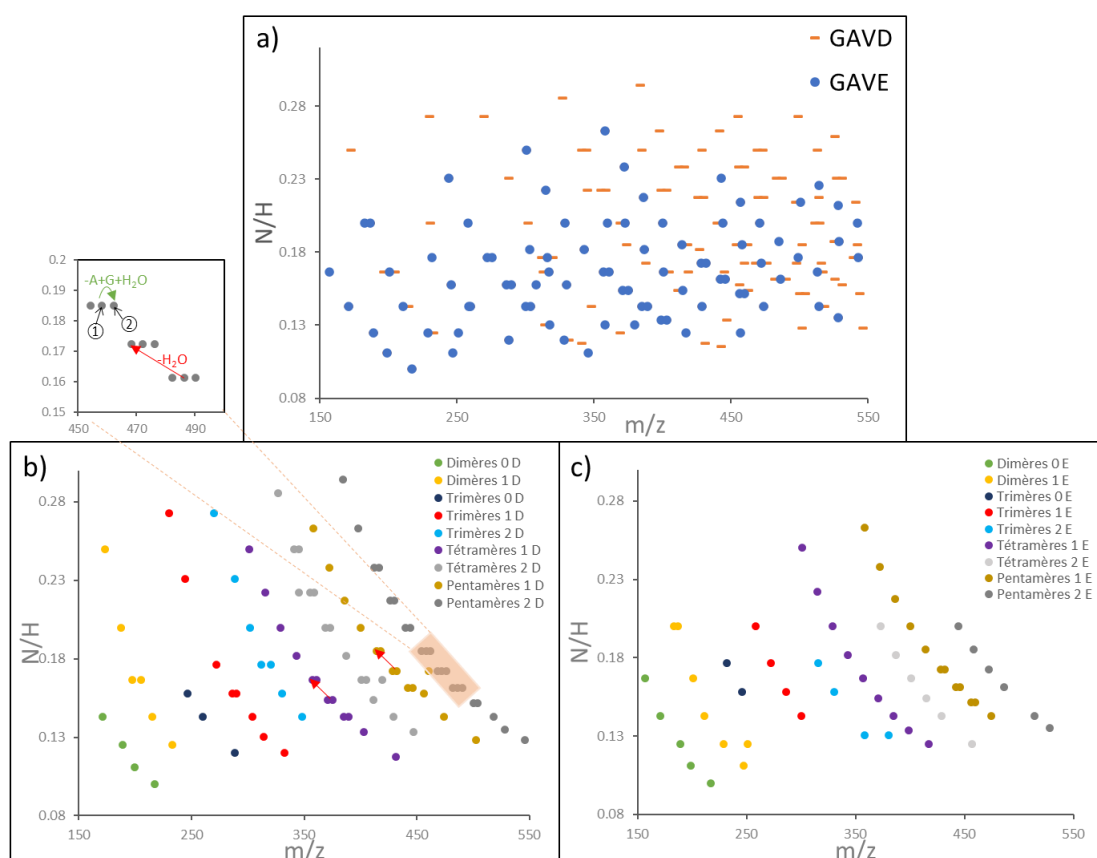


Figure V-12. Diagrammes de Van Krevelen  $f(m/z) = N/H$  pour des espèces allant jusqu'aux pentapeptides a) pour les deux systèmes GAVD et GAVE. Distribution des séries observées pour le système GAVD (b) et pour le système GAVE (c) selon la taille de l'oligopeptide et la teneur en résidu X. Un zoom sur un groupe de points correspondant à des pentapeptides est

*donné. Les espèces numérotées 1 et 2 correspondent respectivement aux espèces  $[D_2VA_2-2H_2O]H^+$  et  $[D_2VAG-H_2O]H^+$ .*

### II.1.1.2. Indices de Kendrick

Le diagramme de Kendrick (aussi appelé carte de Kendrick) est un outil complémentaire au diagramme de Van Krevelen, basé sur les valeurs de défaut de masse<sup>247</sup> (cf. infra). Il correspond à une normalisation du spectre de masse par rapport à un groupement moléculaire défini. Ainsi, dans les conventions de l'IUPAC (où la masse exacte du carbone est de 12,00000), la masse exacte d'un groupement méthylène (CH<sub>2</sub>) est de 14,01565 Da. Dans un diagramme normalisé par rapport à CH<sub>2</sub>, on fixe cette masse à 14,00000 Da en multipliant par le facteur de normalisation 14,00000/14,01565. Ces masses transformées sont appelées masses de Kendrick (ou KM) :

$$KM = M_{\text{mesurée}} \times \frac{14}{14,01565}$$

On définit aussi les défauts de masse de Kendrick (KMD) selon l'équation :

$$KMD = (KM_{\text{nominale}} - KM) \times 1000$$

La représentation de Kendrick permet de mettre en évidence sur une même ligne horizontale des composés ayant des compositions en hétéroatomes communes, permettant ainsi de classer les composés obtenus.

A titre d'exemple, Marshall et Rodgers<sup>256</sup> ont utilisé le diagramme de Kendrick pour visualiser les différentes « classes » d'hydrocarbures présentes dans un échantillon de pétrole (voir Figure V-13). En effet, les composés d'une même classe ayant un nombre différent d'unités CH<sub>2</sub> sont séparés de 14 u sur l'axe des abscisses et ont le même KMD. Ils apparaissent donc sur une même ligne horizontale. D'autre part, l'introduction d'une insaturation se traduit par la perte de 2 H, donc par un déplacement vertical correspondant à deux fois le défaut de masse de l'hydrogène, soit 0,01340 u.

Cette approche est plus ou moins intéressante en fonction de la nature des échantillons. L'analyse de Kendrick a surtout été utilisée en pétrologie où beaucoup de composés ne diffèrent que par l'addition dans le squelette d'un nombre variable d'unités CH<sub>2</sub>. L'utilisation des indices de Kendrick a été appliquée à d'autres échantillons complexes tels que des échantillons de matière organique.

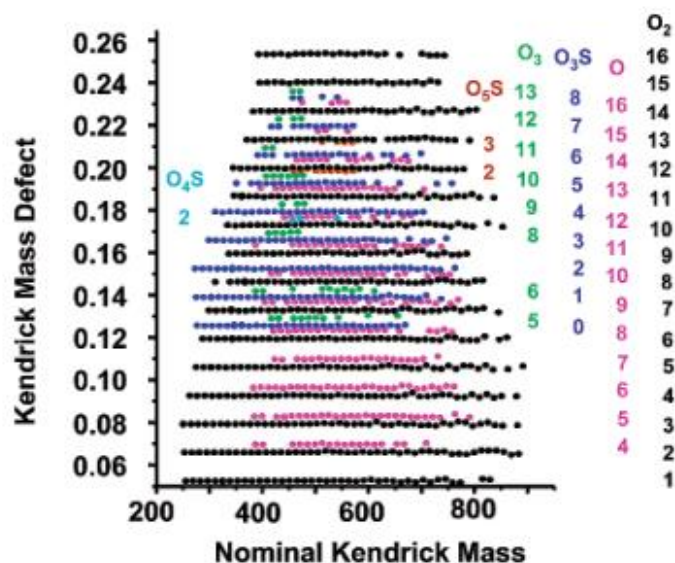


Figure V-13. Diagramme de Kendrick d'un échantillon de pétrole. Figure reproduite d'après Marshall et Rodgers<sup>256</sup>.

La normalisation du spectre de masse peut, dans ce cas, se faire avec une autre unité que  $\text{CH}_2$ , comme  $\text{H}_2$  ou  $\text{O}^{248}$ . Le rapport  $\frac{14}{14,01565}$  est remplacé par le rapport entre la masse nominale et la masse exacte de l'unité considérée. Dans le cas d'analyse de polymères, l'unité de base permettant la normalisation des données peut être l'unité de répétition monomérique<sup>257</sup>.

Une autre représentation graphique possible pour un diagramme de Kendrick est celle du défaut de masse de Kendrick (KMD) en fonction du rapport  $m/z$ .

Dans nos systèmes, divers indices de Kendrick ont été calculés, entre autres basés sur un motif  $\text{CH}_2$  ( $14/14,01565 = 0,99888$ ) ou un atome d'oxygène  $\text{O}$  ( $16/15,99491 = 1,00032$ ). Les diagrammes obtenus sont respectivement reportés Figure V-14 et Figure V-15a. Dans le premier graphique, les différents points sont rassemblés en clusters. Les clusters entourés en vert (a) sur la figure correspondent aux dipeptides comportant 1 X ou aux tripeptides sans X. De même, ceux entourés en rouge (b) correspondent aux tripeptides avec 2 X, tétrapeptides 1 X ou pentapeptides sans X. Dans le second graphique, les mêmes points se regroupent suivant des lignes obliques. On constate que chaque ligne regroupe des oligopeptides ayant la même longueur, et le même nombre de résidus trifonctionnels (Figure V-15b et c). Les formes intactes et déshydratées n'apparaissent pas sur la même ligne. Pour

une même espèce, la forme déshydratée se trouve décalée d'un vecteur (-18,0105 ; -16,4) par rapport à la forme intacte (voir les flèches rouges sur la Figure V-15b).

Il est important de noter que les espèces complexées à des cations ne se situent pas sur la même ligne que les espèces protonées correspondantes. Ainsi les points notés 1, 2 et 3 sur la Figure V-15c correspondent respectivement aux ions  $[\text{EV-H}_2\text{O}]\text{Na}^+$ ,  $[\text{EVG-H}_2\text{O}]\text{Na}^+$  et  $[\text{E}_2\text{V-H}_2\text{O}]\text{Na}^+$  et n'apparaissent pas dans leur série respective.

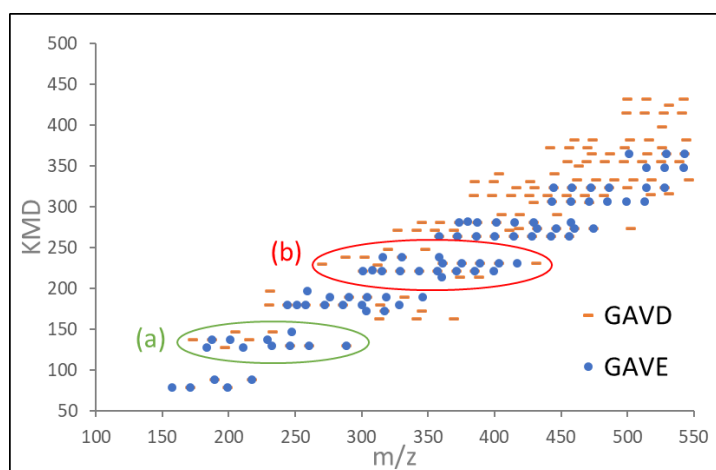


Figure V-14. Diagramme de Kendrick normalisé par le groupement  $\text{CH}_2$  des systèmes GAVD et GAVE.

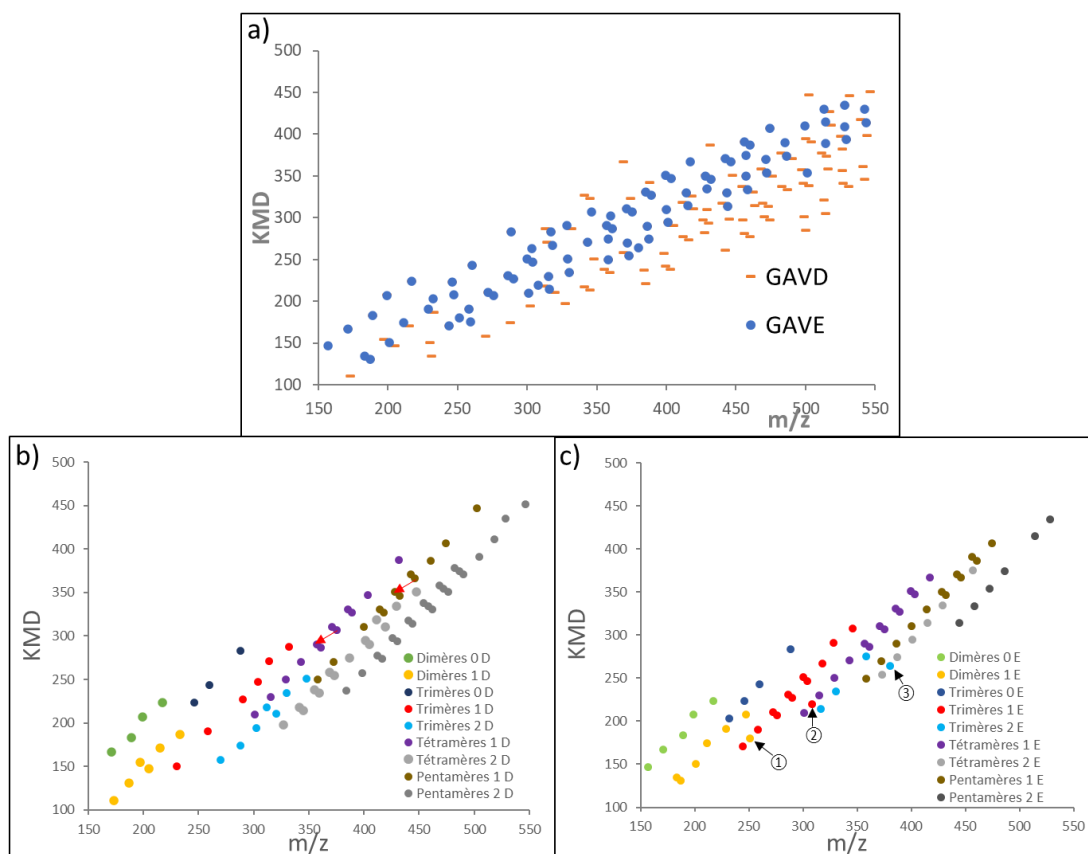


Figure V-15. Diagrammes de Kendrick normalisés par O en fonction de  $m/z$  a) Diagramme pour les deux systèmes GAVD et GAVE. Distribution des séries observées pour le système GAVD (b) et le système GAVE (c) selon la taille de l'oligopeptide et la teneur en résidu X.

Les diagrammes de Van Krevelen et Kendrick permettent de visualiser et comparer les différents oligopeptides formés dans les systèmes, et d'aider à l'identification des oligopeptides. Les diagrammes de Kendrick sont particulièrement intéressants à cet égard, car ils ne nécessitent pas l'identification préalable de la stœchiométrie. Une superposition importante des espèces observées dans les deux échantillons étudiés a été mise en évidence, suggérant une grande similitude des polymérisations des deux systèmes.

Une analyse plus détaillée des différents oligopeptides formés est présentée dans le paragraphe suivant.

### II.1.2. Spécificités des acides aminés trifonctionnels par rapport aux résidus aliphatiques

Les spectres de masse reportés en Figure V-10 montrent une faible proportion d'oligopeptides ne contenant pas d'acide aminé trifonctionnel : Asp et Glu sont davantage

incorporés que les acides aminés aliphatiques. Afin de confirmer cette observation, le taux d'espèces contenant ou non chacun des quatre acides aminés a été calculé. Pour ce faire, pour chaque longueur donnée d'oligopeptide et indépendamment de la structure (forme intacte ou déshydratée) de l'espèce détectée (protonée ou adduit), les intensités des espèces contenant un acide aminé donné ont été sommées pour ensuite accéder au rapport « intensité des espèces contenant l'acide aminé (noté X)/totalité des espèces ». On se souviendra que les intensités utilisées pour ces calculs ne reflètent pas directement les concentrations relatives de chaque espèce, les oligopeptides ayant des efficacités d'ionisation différentes. C'est déjà le cas pour les monomères, pour lesquels nous avons établi l'ordre relatif des affinités protoniques (en Annexe 10) Glu > Leu > Val, et Leu > Asp.

La Figure V-16 donne ce taux calculé pour chaque acide aminé et pour chaque longueur d'oligopeptides. On peut noter que le taux d'espèces contenant l'acide aminé trifonctionnel (Glu ou Asp) est supérieur aux valeurs de ceux des acides aminés aliphatiques pour toutes les longueurs. La seule exception concerne les dipeptides dans le système GAVD, où Val est un peu plus incorporé que Asp. Pour comparaison, si les acides aminés étaient répartis de façon parfaitement aléatoire, la proportion de dipeptides contenant chaque acide aminé serait de  $1 - (3/4)^2 = 43,75$  %. La différence avec le taux observé (de l'ordre de 80 % pour Asp et Glu) est sans doute trop grande pour être uniquement due à la différence d'efficacité d'ionisation des acides aminés trifonctionnels. Le résidu Val est, comme Glu ou Asp, présent en grande quantité dans les dipeptides détectés. On peut donc faire l'hypothèse que la réaction de dimérisation la plus efficace dans le système GAVE est  $E + V \rightarrow EV$  (sans préjuger de la séquence du dipeptide), et est  $D + V \rightarrow DV$  dans le système GAVD.

La glycine, et dans une moindre mesure l'alanine, se comportent différemment. Gly est peu présent dans les dipeptides, et la probabilité de l'incorporer augmente régulièrement avec la longueur de la chaîne. Il semble donc que Gly ne peut réagir qu'avec des oligopeptides préformés (élongation de chaîne).



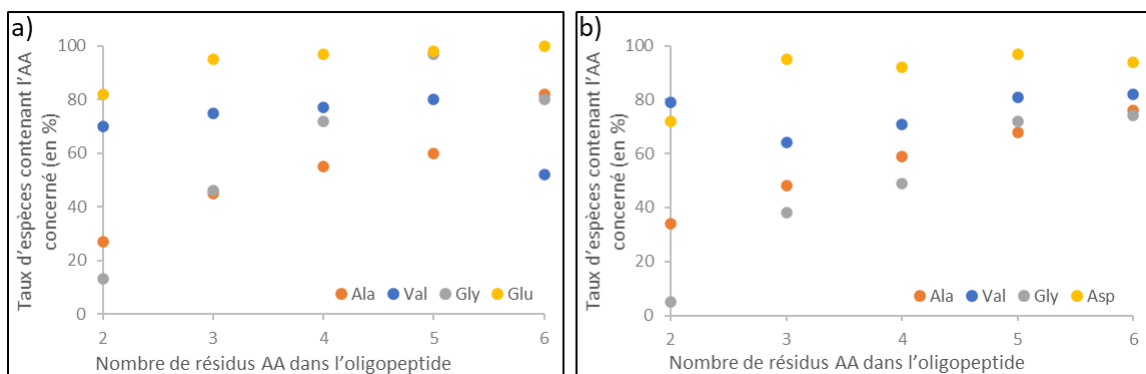


Figure V-16. Taux d'espèces contenant un acide aminé donné en fonction de la taille de l'oligopeptide pour les systèmes a) GAVE et b) GAVD.

On peut aussi estimer un taux d'incorporation de chaque acide aminé X dans les n-mères, ou un taux global, en sommant les abondances relatives de toutes les espèces, pondérées par le nombre de X contenus dans cette espèce.

La répartition des acides aminés au sein des peptides formés est montrée Figure V-17. Pour le système GAVD, le taux d'incorporation dans les dipeptides suit l'ordre Val > Asp > Ala > Gly. Pour les oligopeptides d'ordre supérieur, et globalement, on a Asp > Val > Ala > Gly.

En ce qui concerne le système GAVE, les dipeptides et tripeptides suivent aussi l'ordre Glu > Val > Ala > Gly, mais celui-ci évolue avec l'augmentation de la longueur des oligopeptides et l'ordre global devient Glu > Val > Gly > Ala. Il est difficile de corrélérer l'incorporation des acides aminés à une propriété physico-chimique particulière. Par exemple, d'après la littérature, les températures de sublimation de Gly, Val et Ala sont très proches<sup>258</sup>, et il est donc vain de vouloir expliquer l'ordre d'incorporation par une plus grande facilité de diffusion en phase gazeuse par exemple. On peut penser que l'origine de cet ordre est une cinétique de réaction plus rapide pour certaines paires d'acides aminés, dirigeant le système vers la formation d'oligopeptides particuliers. Dans la majorité des cas, l'ordre entre les acides aminés aliphatiques est Val > Ala > Gly, ce qui suggère une corrélation avec la longueur de la chaîne latérale, mais il existe des exceptions.

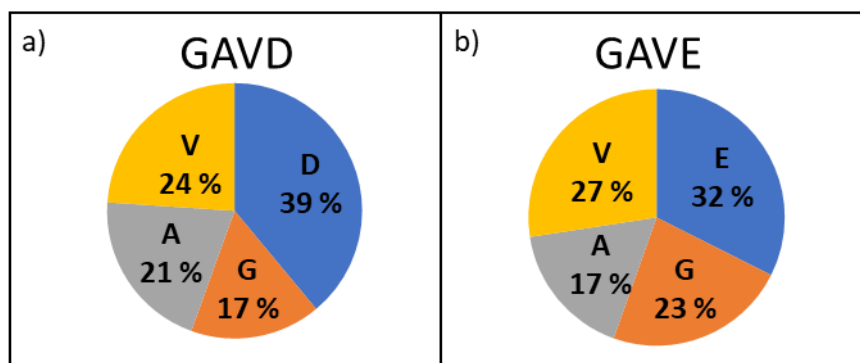


Figure V-17. Taux d'incorporation des acides aminés dans les oligopeptides toutes longueurs confondues pour les systèmes a) GAVD et b) GAVE.

Toutes ces observations indiquent *a minima* que ces systèmes quaternaires montrent une sélectivité stœchiométrique de polymérisation, comme les systèmes binaires de référence. Elle est bien sûr plus difficile à analyser en détail, compte tenu du nombre élevé de combinaisons possibles.

Puisque la présence d'un acide aminé trifonctionnel semble jouer un rôle important dans le scénario de polymérisation, nous avons aussi étudié la polymérisation de GAVL/SiO<sub>2</sub>, système ne possédant pas un tel acide aminé. Le spectre de masse obtenu (en Annexe 12) montre la formation uniquement de dipeptides et tripeptides, traduisant une polymérisation beaucoup moins efficace. L'importance d'un résidu trifonctionnel pour permettre la formation d'oligopeptides relativement longs semble bien confirmée.

## II.2. Différence de réactivité entre les systèmes GAVD et GAVE

De prime abord, Asp et Glu semblent se comporter de manière similaire. En effet une polymérisation efficace est observée pour les deux systèmes, entraînant la formation de nombreuses espèces de stœchiométries différentes, avec des oligopeptides où l'acide aminé trifonctionnel est nettement plus incorporé que les autres. Des formes déshydratées

majoritaires (à 160 °C) sont détectées également dans les deux systèmes. Néanmoins, certaines différences peuvent être discernées entre Asp et Glu.

### II.2.1. Incorporation différente de l'acide aminé trifonctionnel

Au vu des expériences faites sur GAVD versus GAVE, mais également entre les systèmes binaires Leu + Asp et Leu + Glu, les résidus Asp semblent être incorporés en plus grand nombre que les résidus Glu dans les oligopeptides de longueur supérieure à 2.

La Figure V-18a montre l'exemple des tétrapeptides ; des espèces contenant trois résidus Asp sont formées pour le système GAVD alors que pour GAVE, les tétrapeptides contiennent au maximum deux résidus Glu. Si on somme sur toutes les longueurs l'incorporation des résidus Asp et Glu dans les oligopeptides, la même tendance apparaît. Dans la Figure V-18b, le nombre de résidus Asp ou Glu est comparé au nombre total de résidus du peptide : pour le système GAVD, le nombre de résidus Asp est égal ou supérieur à la moitié du nombre de résidus dans 44 % des oligopeptides identifiés, tandis que ce taux est de 33 % pour le système GAVE.

Des résultats similaires sont obtenus pour les deux systèmes binaires Leu + Asp et Leu + Glu. Asp est plus incorporé que Glu dans les oligopeptides pour les mêmes conditions de condensation, comme le montre la Figure V-18c. La proportion d'oligopeptides dans lesquels le nombre de X est supérieur ou égal à la moitié du nombre de résidus est de 76 % pour le système Leu + Asp, contre 49 % pour le système Leu + Glu.

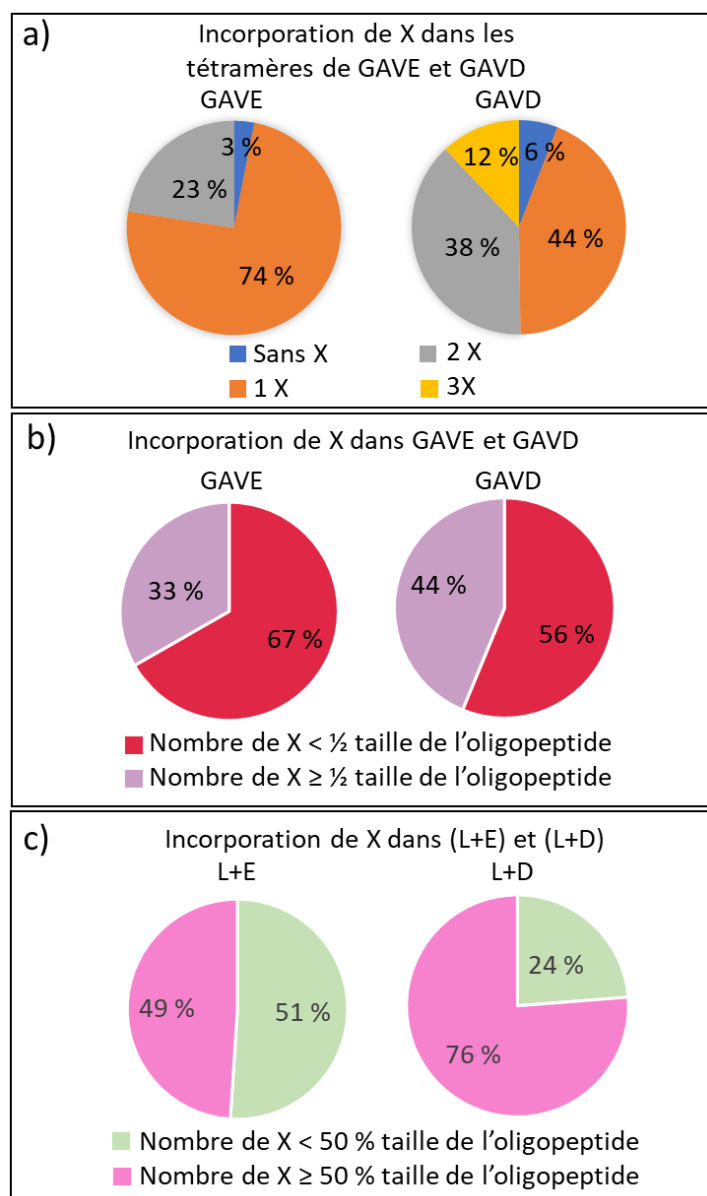


Figure V-18. Etude de l'incorporation de l'acide aminé trifonctionnel ( $X = E$  ou  $D$ ). a) Comparaison du nombre de X incorporés dans les térapeptides de GAVE et GAVD ; b) Comparaison du nombre de résidus X incorporés au sein des oligopeptides par rapport au nombre total de résidus d'acides aminés pour les systèmes GAVE et GAVD ; c) Comparaison du nombre de résidus X et L dans les systèmes Leu + Glu et Leu + Asp.

## II.2.2. Différence de comportement vis-à-vis de la déshydratation

Un autre critère a été étudié dans l'optique de révéler une différence de réactivité entre les deux systèmes : l'apparition des formes déshydratées. D'après l'étude cinétique (voir Chapitre IV), les formes déshydratées sont issues de réactions secondaires des espèces

intactes. Les espèces déshydratées apparaissent en effet surtout à partir de 130 °C dans l'échantillon (L+E)/SiO<sub>2</sub>.

Une activation à 130° C a été menée pour les deux systèmes GAVD et GAVE. Dans les quatre cas (130 °C et 160 °C pour les deux systèmes), pour les peptides d'une longueur donnée, on a calculé le rapport entre la somme des intensités des espèces intactes et la somme des intensités des espèces déshydratées. La Figure V-19 montre l'évolution de ce rapport avec la longueur des oligopeptides. A 130 °C, il y a plus de formes déshydratées pour GAVE que pour GAVD, donc Glu se déshydrate plus facilement que Asp. En revanche, à 160 °C, les deux systèmes se comportent de la même façon, à savoir qu'ils présentent un même taux (élevé) de formes déshydratées. Cette différence de comportement suggère une différence des structures des formes déshydratées entre Asp et Glu. En effet, si la formation d'unités pyroGlu semble probable pour GAVE, une cyclisation interne de type succinimide semble plus favorable pour le système GAVD. Une autre différence notable est l'augmentation de la quantité de formes déshydratées pour une activation à 130 °C avec la longueur de la chaîne dans le système GAVD. A nouveau, cela peut s'expliquer par la différence des structures formées : une cyclisation de type pyroGlu est envisageable pour tous les résidus d'une chaîne peptidique, contrairement à une cyclisation de type succinimide qui ne peut pas avoir lieu en C-terminal. La structure des formes déshydratées dans les différents systèmes sera étudiée plus en détail au Chapitre VI.

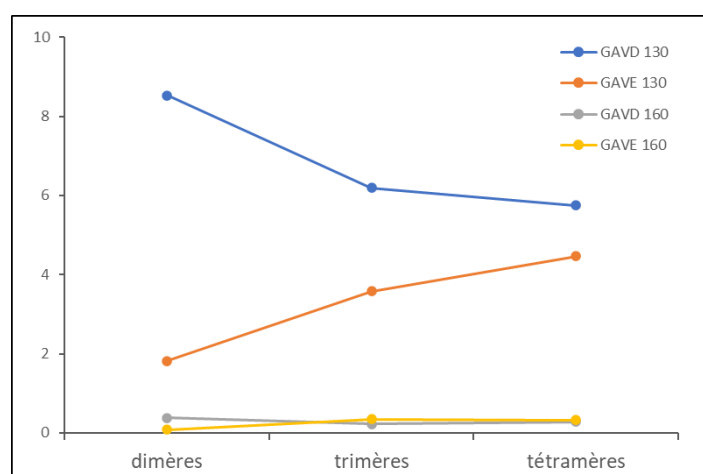


Figure V-19. Comparaison du rapport des oligopeptides intacts/déshydratés en fonction de la longueur de l'oligopeptide pour les échantillons des systèmes GAVD et GAVE activés à 130 °C et 160 °C.

La stéréosélectivité dans la condensation du système complexe GAVE a été également étudiée. Il avait été montré au Chapitre III qu'il n'existait pas de stéréosélectivité pour l'échantillon (L+E)/SiO<sub>2</sub>. La même approche, c'est-à-dire l'utilisation d'énantiomères marqués, a été utilisée : les deux échantillons (G<sub>L</sub>+A<sub>L</sub>+V<sub>L</sub>+E<sub>L</sub>+E<sub>D</sub><sup>\*</sup>)/SiO<sub>2</sub> et (G<sub>L</sub>+A<sub>L</sub>+V<sub>L</sub>+E<sub>L</sub><sup>\*</sup>+E<sub>D</sub>)/SiO<sub>2</sub> ont été comparés. Dans le premier, les acides aminés aliphatiques Gly, Ala et Val sont mélangés avec L-Glu et D-Glu marqué au <sup>13</sup>C en quantité équimolaire, dans le second les acides aminés aliphatiques sont mélangés avec L-Glu<sup>13</sup>C et D-Glu. L'analyse des spectres de masse (en Annexe 13) pour chaque stœchiométrie indique qu'aucune stéréosélectivité significative n'est observable pour les oligopeptides allant jusqu'aux hexapeptides.

### III. Etude de la polymérisation de nucléotides

Des exemples de polymérisation de nucléotides en présence de surfaces ont été rapportés dans la littérature (voir Chapitre I), mais ils concernaient des nucléotides activés<sup>47</sup> ou une polymérisation assistée par lipides<sup>50</sup>. A notre connaissance, aucune étude n'a démontré la possibilité de polymériser des nucléotides non activés sur surface minérale. Nous avons donc étudié la possibilité d'une polymérisation de nucléotides dans des conditions comparables à celles où nous avons observé la polymérisation d'acides aminés, c'est-à-dire par activation thermique de monomères déposés sur une surface minérale à partir d'une solution aqueuse (les monomères d'acides aminés étant remplacés par des nucléotides).

La procédure de dépôt est donc similaire à celle utilisée pour les acides aminés (voir Chapitre II). Plusieurs couples de nucléotides monophosphatés ou triphosphatés, ont été testés : AMP+GMP (adénosine monophosphate + guanosine monophosphate), ATP+GTP (adénosine triphosphate + guanosine triphosphate) et ATP+UTP (adénosine triphosphate + uridine triphosphate). En ce qui concerne les bases portées par les nucléotides, les deux premières expériences impliquent des couples purine/purine, la troisième un couple purine/pyrimidine. Les analyses ont été faites sur FTICR en injection directe en mode négatif. Pour le couple de nucléotides monophosphatés, aucune oligomérisation n'a eu lieu.

Avec les nucléotides triphosphates, en revanche, des pics correspondant à des dimères sont détectés. Le spectre obtenu pour l'échantillon (ATP+UTP)/SiO<sub>2</sub> est montré à la Figure V-20.

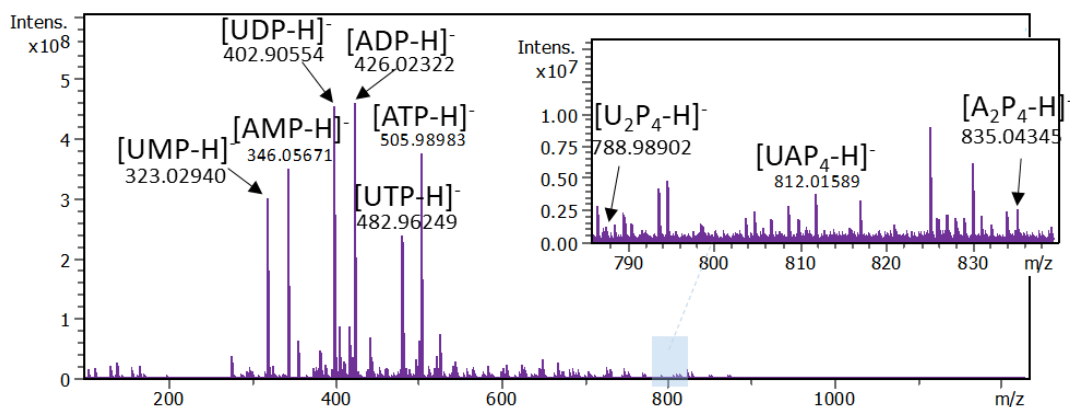


Figure V-20. Spectre de masse de l'échantillon (ATP+UTP)/SiO<sub>2</sub> activé à 160 °C (ESI négatif, instrument FTICR). Un zoom sur les différents dimères obtenus est montré.

Les monomères eux-mêmes sont modifiés. Les pics majoritaires correspondent aux monomères diphosphatés, et les formes monophosphatées sont aussi présentes avec des intensités importantes. Il a été vérifié que les conditions d'analyses ne causent pas de déphosphorylations.

Stœchiométriquement, la perte d'un groupement phosphate implique une hydrolyse d'une liaison anhydride phosphorique P-O-P. Or, dans nos conditions d'activation, il ne reste pas d'eau sur la surface de silice aux températures (> 100°C) auxquelles on observe l'apparition d'espèces déphosphorylées. Dès lors, on peut faire l'hypothèse que la première étape de ces réactions de modification des monomères implique des cyclisations, comme montré à la Figure V-21. Ces réactions ne font pas intervenir de molécules d'eau. Elles aboutissent à des espèces de stœchiométrie (AMP-H<sub>2</sub>O) et (ADP-H<sub>2</sub>O), respectivement (UMP-H<sub>2</sub>O) et (UDP-H<sub>2</sub>O), qui sont très réactives et probablement réhydratées lors de la désorption ; les ions détectés sur le spectre de masse sont alors les ions [AMP-H]<sup>-</sup>, [ADP-H]<sup>-</sup>, [UMP-H]<sup>-</sup> et [UDP-H]<sup>-</sup>.

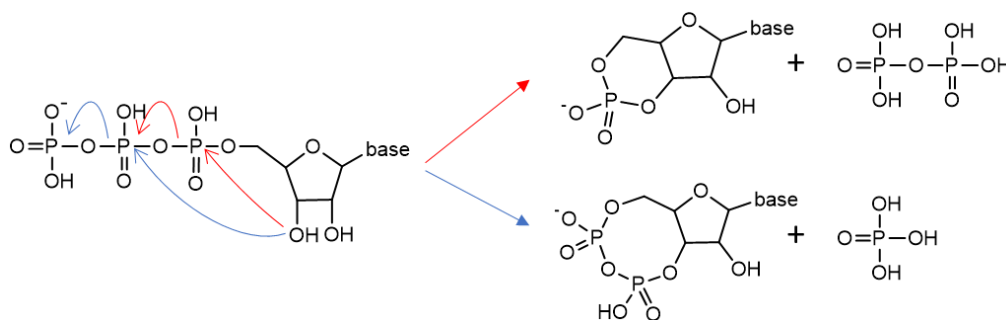


Figure V-21. Schéma proposé pour l'étape de déphosphorylation, responsable de la formation de dimères partiellement déphosphorylés.

Des dimères tétra-phosphatés sont aussi détectés, bien qu'avec des intensités faibles (voir zoom Figure V-20). Ils sont notés en fonction de la stœchiométrie globale : l'espèce notée  $[UAP_4-H]^-$  peut donc avoir les séquences suivantes : PPPUPA ou PPPAPU (voire PPUPPA, etc). La polymérisation ne va pas plus loin (pas de trimères...) et aucun pic correspondant aux motifs des nucléobases ou ribose, ni à des déshydratations n'est observé.

Il est probable que la condensation a lieu entre un des groupements hydroxyle d'un nucléotide et un phosphate d'un autre pour former une liaison phosphodiester par perte d'un groupement pyrophosphate<sup>259</sup> comme schématisé dans la Figure V-22. La figure montre la formation d'une liaison entre les carbones 5' et 3' de deux riboses successifs, mais il faut souligner que la technique utilisée ne permet pas de déterminer la régiosélectivité.

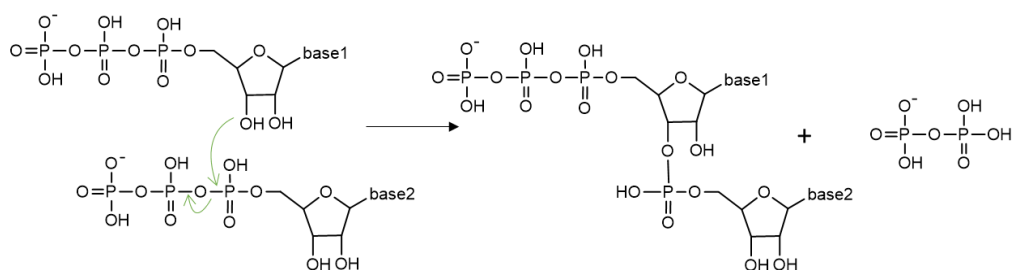


Figure V-22. Schéma de condensation de deux nucléotides et formation d'un dinucléotide de stœchiométrie ( $X_2P_4$ ).

Ce même échantillon a été activé à 200 °C. Le spectre de masse obtenu ne montre pas la formation d'espèces plus longues. En revanche, parmi les monomères, on peut observer davantage de formes mono et diphosphatées, prouvant que ces formes correspondent à des produits de dégradation thermique.

L'échantillon (ATP+GTP)/SiO<sub>2</sub> (cette fois les deux nucléobases sont des purines) activé à 160 °C analysé dans les mêmes conditions montre les mêmes caractéristiques de dimérisation : les monomères diphosphatés sont majoritaires et des dimères bi, tri ou tétra-phosphatés sont aussi détectés. En revanche, les intensités des pics correspondant aux dimères sont plus faibles de deux ordres de grandeur que celles des monomères, alors que ce rapport est d'un ordre de grandeur dans le cas ATP+UTP, indiquant une dimérisation moins efficace.



En résumé, seule une dimérisation est possible pour une activation de 160 °C sur surface de silice et il est nécessaire de partir de deux nucléotides triphosphatés. Bien que ces espèces soient de tailles limitées, l'intérêt n'est pas négligeable car certaines coenzymes naturelles sont des dinucléotides comme la nicotinamide adénine dinucléotide NAD<sup>260</sup>. Afin d'obtenir une polymérisation plus efficace menant à des espèces plus longues, une perspective possible serait d'augmenter la charge en nucléotides sur silice ou d'effectuer une réhumidification pour permettre une mobilité des nucléotides (voir étude des cycles hydratation/déshydratation au Chapitre IV). Néanmoins, la dimérisation des nucléotides sur silice ne se déroule pas comme la polymérisation des acides aminés. Elle résulte en effet d'une recombinaison entre deux nucléotides triphosphates, qui sont déjà des molécules chimiquement activées : le prix en enthalpie libre de la formation d'une liaison est payé par la rupture d'une autre liaison.

Dans ce chapitre, le scénario de polymérisation sur silice a été généralisé à d'autres systèmes. Tout d'abord, plusieurs systèmes binaires, se différenciant par la nature de l'acide aminé trifonctionnel (Asp,  $\beta$ Glu ou Gln), ensuite deux systèmes quaternaires, plus réalistes du point de vue prébiotique (GAVD et GAVE). Dans toutes les expériences une polymérisation efficace se produit, avec formation de polypeptides de longueurs comparables.

Des différences dans la stœchiométrie des espèces formées ont été constatées entre les différents systèmes, notamment en raison des réactivités respectives des acides aminés trifonctionnels étudiés. Dans les systèmes binaires comme quaternaires, on peut observer une réactivité supérieure de Asp, en ce sens que les résidus Asp sont davantage incorporés, et donnent aussi lieu à la formation d'oligopeptides plus longs.

Enfin le protocole de polymérisation sur surface de silice a été élargi à l'étude des nucléotides. La polymérisation n'est pas très efficace. Une faible quantité de dinucléotides est observée, et ce, seulement en partant de nucléotides triphosphates.



Chapitre VI. Etude structurale des  
oligopeptides formés par condensation  
sur surface de silice

Dans les chapitres précédents, notamment au Chapitre III, l'existence d'une sélectivité de polymérisation au niveau de la stœchiométrie a été mise en évidence. Nous avons remis à plus tard la question de la formation de séquences spécifiques pour des peptides de même stœchiométrie, une forme de sélectivité cruciale pour l'obtention de peptides fonctionnels. Pour identifier des séquences peptidiques différentes, la méthode la plus appropriée est la spectrométrie de masse en tandem (MS/MS). Ainsi, des peptides d'une même stœchiométrie mais de séquences différentes ne sont pas distinguables en spectrométrie de masse dite à une dimension. En revanche, la MS/MS donne accès à la séquence en détectant majoritairement des ions fragments dits de séries, *b<sub>i</sub>* ou *y<sub>i</sub>* en mode CID (voir Chapitre I, § II.4).

## I. Présentation des techniques utilisées

### I.1. Détermination des séquences peptidiques par MS/MS

La spectrométrie de masse en tandem est très utilisée pour déterminer les séquences peptidiques en protéomique<sup>134</sup>. Bien que notre étude se rapproche de ce domaine, les méthodes classiquement employées, comme l'utilisation de banques de données, ne sont pas applicables. Par ailleurs, les voies de fragmentation étant « séquences dépendantes »<sup>134</sup>, elles ne sont pas prédictibles et l'attribution sans ambiguïté nécessite donc une comparaison avec les fragmentations de composés de référence. En effet, si en théorie, l'obtention d'ions fragments permet d'accéder directement à la séquence d'une molécule, en pratique le séquençage peut être incomplet et est composé-dépendant en termes d'ions fragments et d'intensités relatives de ces ions fragments. Il a néanmoins été montré des tendances générales de voies de fragmentation dans une cellule de collisions<sup>261-263</sup> en fonction de la présence de certains résidus. Par exemple, si un résidu proline est présent dans la séquence, le pic fragment le plus intense dans le spectre CID est généralement dû à la fragmentation du côté N-terminal de la proline<sup>264,265</sup>. De même, une comparaison des voies de fragmentation des peptides contenant un résidu Asp ou Glu a montré qu'en raison de la capacité de la chaîne latérale de Asp à former un cycle succinimide avec la liaison amide adjacente, le départ du fragment carboxyl est favorisé, et que les peptides contenant Asp

avaient ainsi tendance à se fragmenter au niveau de la liaison amide adjacente à Asp du côté C-terminal (c.a.d. Asp-Xxx). En revanche ce comportement n'est pas observé dans le cas d'une liaison Glu-Xxx, la chaîne latérale formant un cycle à six carbones<sup>266</sup>. Par ailleurs, des observations empiriques montrent que des fragmentations internes (pertes de petites masses) se produisent souvent en présence de certains résidus tels que la proline<sup>264,265</sup> ou l'acide aspartique<sup>266</sup>, et encore plus au niveau de liaison Asp-Pro<sup>267</sup>, soulignant ainsi que toutes les liaisons peptidiques n'ont pas la même capacité à se fragmenter sous CID.

Afin d'identifier sans ambiguïté les composés présents dans le mélange de désorption de l'échantillon (L+E)/SiO<sub>2</sub>, il est nécessaire de disposer de peptides de référence de tous les oligopeptides observés (intacts et déshydratés). Certains dipeptides et tripeptides commerciaux ont été achetés, d'autres (certains tétrapeptides, pentapeptides et hexapeptides) ont été synthétisés (voir Annexe 14). Il faut noter que les peptides de référence sont des peptides formés par liaison peptidique  $\alpha$ . Or la formation de régioisomères peut être envisagée dans nos conditions de condensation, ce qui apporte une ambiguïté dans l'attribution des séquences. Nous avons systématiquement effectué des expériences MS/MS en infusion directe sur l'instrument FTICR. Bien que l'analyse de ces spectres MS/MS soit complexe sans séparation préalable (*a priori* en présence d'un mélange de plusieurs séquences), c'est un moyen direct et rapide qui constitue une première approche instructive. Nous avons étudié en parallèle les voies de fragmentation des peptides du système Leu + Glu et des autres systèmes activés (Leu + Gln, GAVD, GAVE etc). Un tel processus permet de mettre en évidence les voies de fragmentation caractéristiques des oligopeptides formés, en fonction de l'acide aminé trifonctionnel impliqué, en vue d'établir des rapprochements et d'avoir une vision globale de leur réactivité. Par la suite, nous avons également utilisé la MS/MS en couplage, après une étape de séparation par chromatographie liquide (LC/MS/MS), ou par la mobilité ionique (IM-MS/MS) sur l'instrument TIMS-TOF (voir Chapitre II). Bien que le même mode d'activation CID soit appliqué dans les deux instruments utilisés, les spectres MS/MS (obtenus avant ou sans séparation au préalable) diffèrent entre les appareils utilisés. Ainsi, les fragments de bas m/z, notamment les ions  $\gamma_1$ , sont rarement observés dans les spectres MS/MS opérés sur l'appareil TIMS-TOF, sans doute en raison de la perte des ions des basses masses lors du transfert vers le temps de vol.

## I.2. Couplage de la spectrométrie de masse avec des techniques de séparation

La chromatographie liquide couplée à la spectrométrie de masse (LC/MS) est une technique d'analyse exhaustive que nous avons choisi d'utiliser pour caractériser les oligopeptides produits par condensation. A chaque pic du chromatogramme correspond un spectre de masse. De plus, la combinaison de la chromatographie liquide et de la MS/MS (mode LC/MS/MS) est également possible et permet la distinction des pics correspondant aux différentes séquences. Ce couplage a été utilisé pour étudier les échantillons (L+E)/SiO<sub>2</sub>, (GAVE)/SiO<sub>2</sub> et (GAVD)/SiO<sub>2</sub> activés à 160 °C. Le chromatogramme total obtenu pour l'échantillon (L+E)/SiO<sub>2</sub> est montré dans la Figure VI-1. Les pics majoritaires correspondent aux espèces [EL-H<sub>2</sub>O]H<sup>+</sup>, [EL<sub>2</sub>-H<sub>2</sub>O]H<sup>+</sup> et [L<sub>2</sub>-H<sub>2</sub>O]H<sup>+</sup>; plusieurs pics correspondent à une même stœchiométrie, ce qui signifie que plusieurs structures sont présentes. Par ailleurs, à partir de ce chromatogramme total, le logiciel de traitement des données Data Analysis permet d'extraire les « *Extract Ion Chromatograms* » (EICs), qui correspondent aux temps de rétention auxquels apparaît un rapport m/z donné, afin d'identifier plus clairement les temps de rétention des différentes espèces.

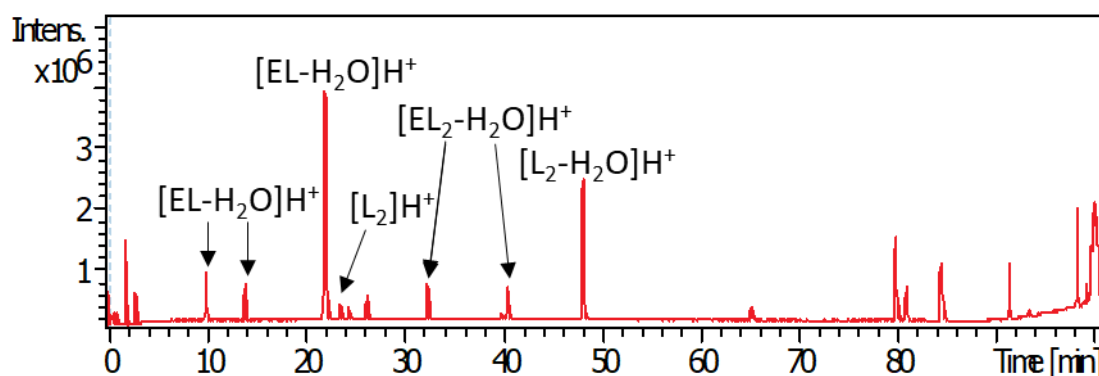


Figure VI-1. Chromatogramme de l'échantillon (L+E)/SiO<sub>2</sub>, activé à 160 °C (analyse LC/MS, détection en mode ESI positif, instrument TIMS-TOF).

La seconde technique séparative utilisée est la mobilité ionique (voir les descriptions théorique et expérimentale de ces deux techniques dans les Chapitres I et II). Le couplage IM-MS/MS, qui permet d'accéder aux fragmentations après séparation par mobilité des oligopeptides protonés, a été utilisé pour étudier les échantillons (L+E)/SiO<sub>2</sub>, (L+βE)/SiO<sub>2</sub>

(L+D)/SiO<sub>2</sub>, (L+Q)/SiO<sub>2</sub>, (GAVE)/SiO<sub>2</sub> et (GAVD)/SiO<sub>2</sub> produits par condensation dans des conditions prébiotiques et activés à 130 °C et 160 °C. Le mobilogramme obtenu pour l'échantillon (L+E)/SiO<sub>2</sub> activé à 160°C est montré dans la Figure VI-2. Conformément aux résultats de LC/MS, les pics majoritaires correspondent aux espèces [EL-H<sub>2</sub>O]Na<sup>+</sup> et [EL<sub>2</sub>-H<sub>2</sub>O]H<sup>+</sup>, au détail près que la complexation de l'espèce (EL-H<sub>2</sub>O) est différente<sup>††</sup>. Tout comme pour les chromatogrammes LC/MS, il est possible de tracer les « *Extract Ion Mobilograms* » (EIMs) pour suivre le mobilogramme pour un rapport m/z donné. L'intégration de chaque pic de mobilité pour un rapport m/z donné permet aussi d'obtenir spectre MS/MS lui correspondant (en mode IM-MS/MS).

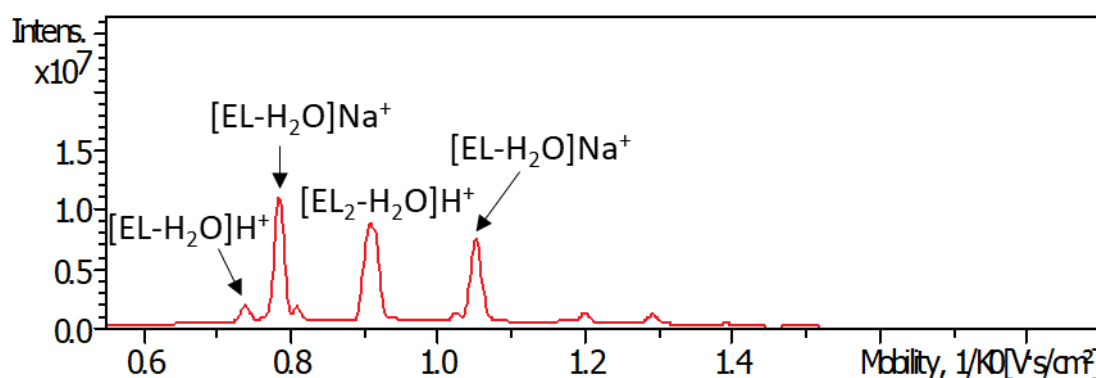


Figure VI-2. Mobilogramme obtenu pour l'analyse IM-MS de l'échantillon (L+E)/SiO<sub>2</sub> activé à 160 °C (ESI positif, instrument TIMS-TOF).

Globalement, la séparation par mobilité ionique s'est révélée moins performante que par chromatographie liquide (en tout cas pour les systèmes étudiés par les deux approches). En revanche c'est une méthode plus rapide et qui permet un choix manuel des expériences MS/MS, contrairement à la LC/MS/MS pour lesquelles l'acquisition des spectres MS/MS est automatique (mode Data Dependent Acquisition, DDA), entraînant une fragmentation des espèces majoritaires uniquement. Néanmoins, de la même manière que pour des acquisitions en infusion directe, cette méthode de séparation post-ionisation est dépendante de l'intensité des ions issus de l'ESI et donc peut être limitée par des effets matrice (voir Chapitre I).

<sup>††</sup> Il est classique d'avoir plus de formes complexées pour des analyses en injection directe. De plus en LC/MS, ces adduits sortent dans le volume mort.



## II. Elucidation structurale des peptides formés par condensation peptidique dans le système Leu + Glu

La première étude a concerné le système Leu + Glu (certains de ces résultats sont présentés dans *ACS Earth Space Chem.* 2020, 4, 1802-1812, voir Annexe 1) pour lequel on dispose de peptides de référence commerciaux ou synthétisés (soit les séquences LLL, EEE, ELL, LEL, LLE, EEL, ELE, LEE, ELLL, LELL, ELEL, ELLE, LELE, EELL, LEEL). Il a donc été possible de comparer les spectres MS/MS de certains des oligopeptides produits par activation thermique de l'échantillon (L+E)/SiO<sub>2</sub>.

### II.1. Etude structurale des oligopeptides intacts

Dans un premier temps nous nous sommes intéressés à l'élucidation structurale des oligopeptides intacts. Cependant, les formes intactes sont minoritaires dans l'échantillon (L+E)/SiO<sub>2</sub> activé à 160 °C et n'ont pas pu être sélectionnées lors de l'acquisition automatique des expériences LC/MS/MS en mode DDA. Celle-ci n'a permis de fragmenter que les espèces déshydratées majoritaires. En revanche, des expériences MS/MS sur FTICR et/ou par couplage IM-MS/MS sur l'appareil TIMS-TOF ont été effectuées pour les échantillons (L+E)/SiO<sub>2</sub> activés à 160 et 130 °C, le dernier échantillon permettant d'étudier une plus grande proportion de formes intactes.

Afin de clarifier la lecture des spectres MS/MS, les flèches correspondant à des pertes de neutres caractéristiques seront annotées en différentes couleurs. Le code couleur, ainsi que les masses exactes des molécules sont indiqués sur la Figure VI-3. Si nécessaire et par souci de clarté, les pertes seront précisées sur les spectres MS/MS, arrondies à l'entier le plus proche. De même, les masses nominales seront utilisées dans le corps du manuscrit pour ce chapitre pour tous les rapports m/z des ions précurseurs et fragments détectés. Toutes les attributions ont été faites avec une précision inférieure à 3 ppm (voir spectres de masse et MS/MS).







Notation	Pertes	Masse exacte
	NH <sub>3</sub>	-17,12655
	H <sub>2</sub> O	-18,01056
	CO	-27,99491
	CH <sub>2</sub> CO	-42,01056
	HCONH <sub>2</sub>	-45,02155
	HCOOH	-46,00556

Figure VI-3. Code couleur utilisé pour annoter les voies de fragmentation dues à des pertes de petites masses, dans spectres MS/MS dans ce chapitre.

Les voies de fragmentation de l'espèce [EL<sub>2</sub>]<sup>H+</sup> de rapport m/z 374 sont montrées Figure VI-4d et comparées avec celles des tripeptides de référence de séquences ELL, LLE et LEL (Figure VI-4a, b et c). Outre les ions fragments de séries b et y théoriquement attendus, des fragmentations consécutives sont observées, notamment la perte de groupements CO et H<sub>2</sub>O. Ces pertes de faibles masses proviennent aussi bien des groupements terminaux du peptide que des chaînes latérales. Au vu des différents spectres MS/MS, on peut établir que la perte de 18 u (H<sub>2</sub>O) à partir de l'ion précurseur est particulièrement favorable pour des peptides où le résidu Glu est en position N-terminale (visible dans le spectre MS/MS du peptide de référence ELL de la Figure VI-4 et aussi EEL et ELE en Annexe 15). On peut également noter qu'une différence de fragmentation consécutive est observée entre les peptides pour lesquels le résidu Glu est localisé en position N-terminale et ceux pour lequel le résidu Glu est au sein de la chaîne peptidique. Pour les premiers, une perte consécutive de 18 u est visible à partir du fragment b<sub>2</sub>, tandis que quand Glu n'est pas en N-terminal, c'est une perte de 28 u qui est observée à partir de ce même fragment b<sub>2</sub>, ce qui peut constituer une perte diagnostique d'un oligopeptide avec le résidu Glu en position N-terminale.

Les voies de fragmentation de l'ion correspondant à la stœchiométrie [EL<sub>2</sub>]<sup>H+</sup> dans le mélange semblent majoritairement correspondre à celles de la séquence ELL (voir Figure VI-4). Néanmoins, le spectre MS/MS n'est pas rigoureusement identique à celui du peptide de référence ELL, bien que les expériences aient été menées dans les mêmes conditions d'activation et le même jour. La variation des intensités relatives des ions fragments peut

signifier un mélange de plusieurs séquences, potentiellement présentes à des abondances différentes (ici ELL est sous doute majoritaire), ou plus probablement la présence de régioisomères (comme discuté au chapitre précédent, soit les peptides classiques dits  $\alpha$ , soit les peptides  $\gamma$  impliquant la chaîne latérale de résidus acides dans la condensation (voir discussion plus bas)).

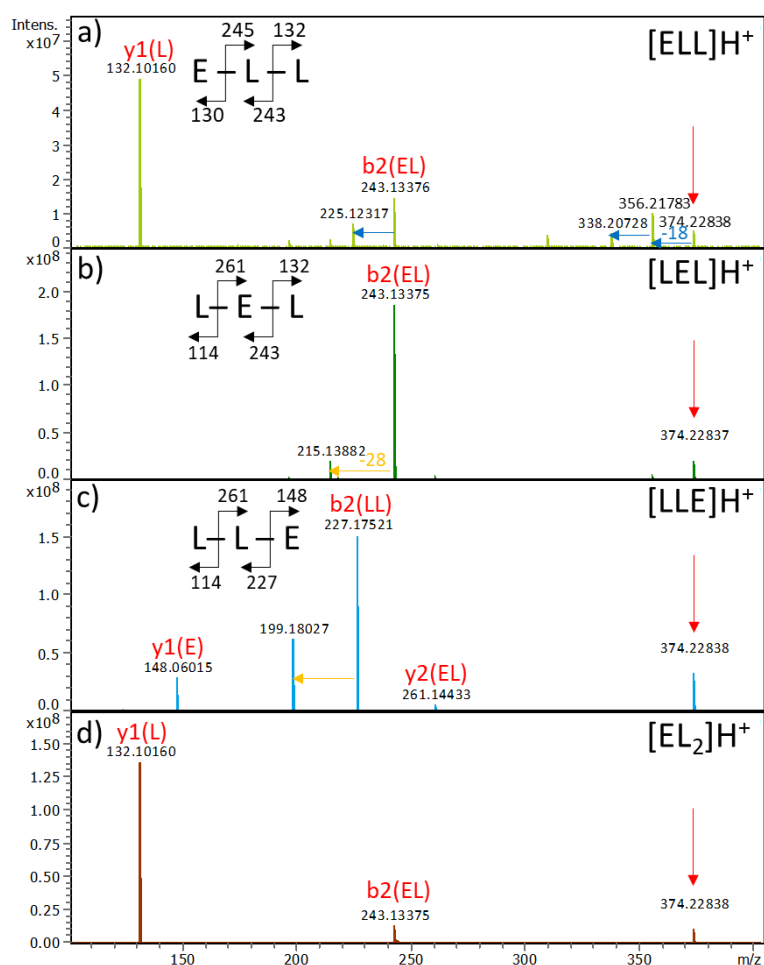


Figure VI-4. Spectres MS/MS des tripeptides de référence de séquences a) ELL (Elab = 8 eV), b) LEL (Elab = 8 eV), c) LLE (Elab = 8 eV) et d) de l'espèce de stœchiométrie  $[EL_2]H^+$  dans le produit de condensation de l'échantillon (L+E)/SiO<sub>2</sub> activé à 160 °C (Elab = 6 eV) (ESI positif, instrument FTICR).

Les expériences IM-MS/MS menées sur le même système Leu + Glu et sur les tripeptides de références n'ont pas permis d'arriver à davantage de certitude sur la ou les séquence(s) de l'espèce de stœchiométrie EL<sub>2</sub>. En effet, les valeurs de mobilité réduite 1/K<sub>0</sub> sont très proches pour les séquences ELL, LEL et LLE. L'écart (RSD) entre ces valeurs de référence et celle de l'espèce de notre mélange est respectivement de 2,08 %, 0,62 % et 2,51 %. Comme

la reproductibilité des mesures et la tolérance généralement admise pour attribuer les valeurs de CCS est de 1-2 %, il nous est difficile d'exclure avec certitude une séquence. De plus, comme pour le spectre MS/MS obtenu sur FTICR, aucune fragmentation des composés de référence ne correspond strictement à celle obtenue pour l'échantillon obtenu par condensation en conditions prébiotiques. Toutefois, comme vu plus haut, la séquence ELL semble majoritaire.

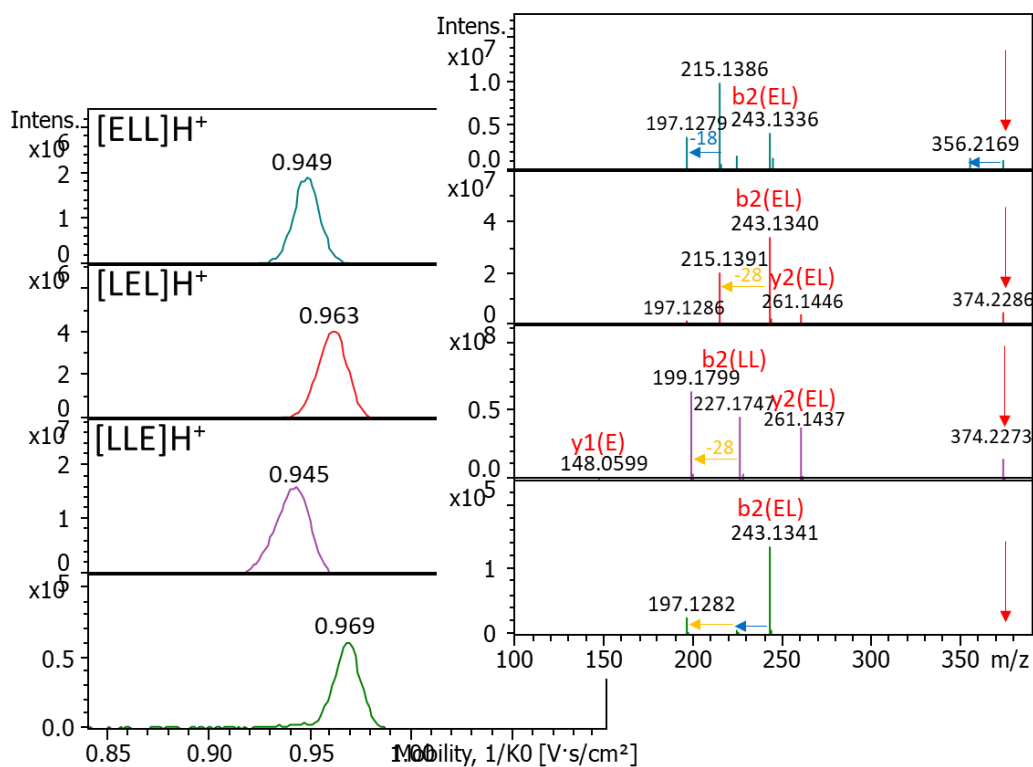


Figure VI-5. Comparaison des mobilogrammes extraits et des spectres MS/MS de l'espèce de stœchiométrie  $EL_2$  dans l'échantillon (L+E)/SiO<sub>2</sub> activé à 160 °C avec ceux des tripeptides de référence de séquence ELL, LEL et LLE (analyse IM-MS/MS, en mode ESI positif, CID  $E_{lab} = 20$  eV, instrument TIMS-TOF).

En ce qui concerne le tripeptide de stœchiométrie  $[E_2L]H^+$ , l'étude IM-MS/MS est montrée Figure VI-6. Au vu des valeurs de RSD, qui sont respectivement de 0,75 %, 3,38 % et 0,11 % pour les séquences ELE, EEL et LEE, la séquence EEL peut être écartée, mais il n'est pas possible de choisir sans ambiguïté entre les deux autres séquences. De la même manière, bien que le spectre MS/MS de l'espèce dans le mélange de désorption semble similaire à celui de la séquence de référence LEE, la présence de la séquence ELE ne peut être exclue (avec une abondance relative sans doute moindre). Par ailleurs, le spectre MS/MS obtenu sur le FTICR (en Annexe 15) montre des voies de fragmentation ne correspondant à aucun

des peptides de référence testés (ou combinaison de ceux-ci). La prédominance du fragment  $y_1(E)$  à  $m/z$  148 suggère la présence de la séquence ELE, mais le fragment  $y_1(L)$  est également présent, tout comme le fragment  $y_2(EL)$ . A l'aune de ces observations, il est difficile de déterminer la séquence de cette espèce, qui est par ailleurs une espèce très minoritaire dans l'échantillon.

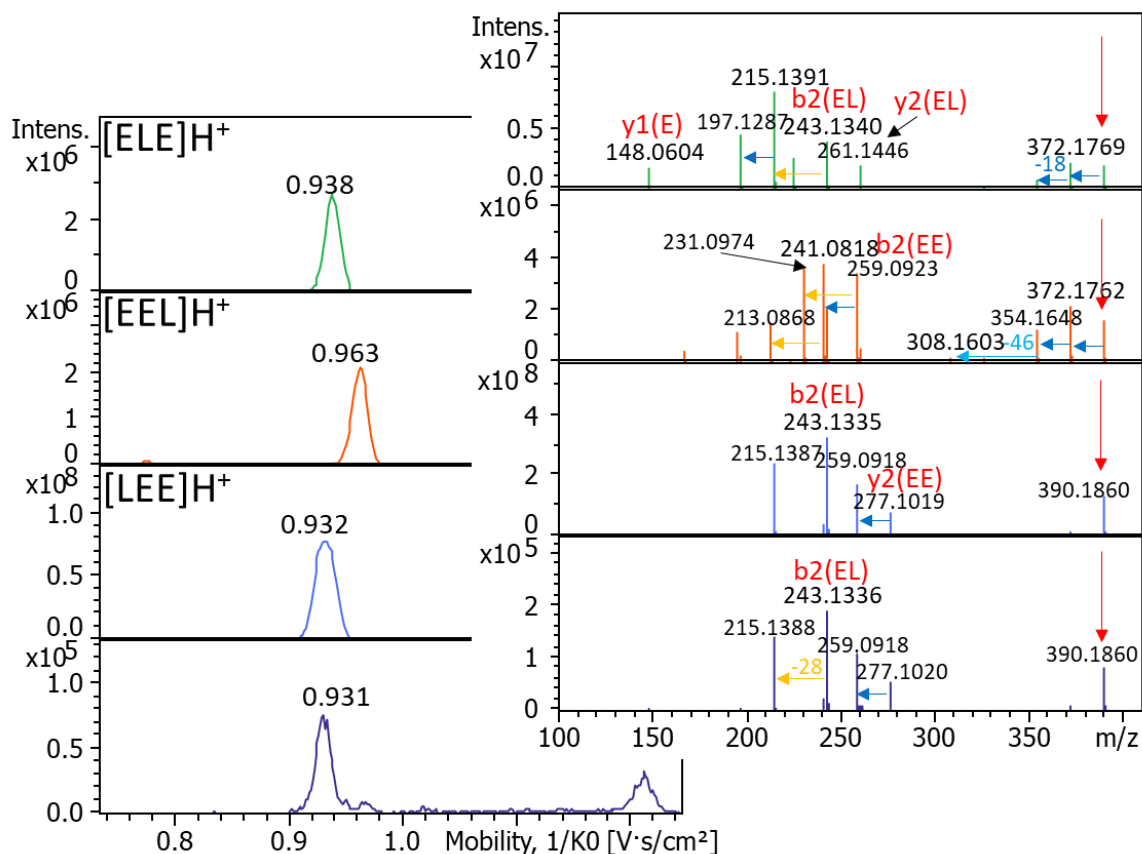


Figure VI-6. Comparaison des mobilogrammes extraits et des spectres MS/MS de l'espèce de stœchiométrie  $E_2L$  dans l'échantillon (L+E)/SiO<sub>2</sub> activé à 160 °C avec ceux des tripeptides de référence de séquence ELE, EEL et LEE (ESI positif, CID Elab = 20 eV, instrument TIMS-TOF).

Pour la stœchiométrie  $EL_3$ , on dispose uniquement des tétrapeptides de référence de séquences ELLL et LELL. L'étude IM-MS/MS est reportée en Annexe 16. D'après les fragmentations, l'espèce dans le mélange correspond à une de ces deux séquences mais il n'est pas possible de conclure pour l'une ou l'autre, les voies de fragmentation des séquences LELL et ELLL étant identiques. Par ailleurs, les différences d'abondances relatives des ions fragments suggèrent, comme auparavant, la présence de régioisomères (voir discussion plus bas).

Une analyse intéressante a pu être faite pour la stœchiométrie  $E_2L_2$ . Ainsi, la Figure VI-7 compare le mobilogramme de l'ion  $m/z$  503 correspondant au térapeptide  $[E_2L_2]H^+$  avec les mobilogrammes extraits de chaque peptide de référence disponible dans des conditions de meilleure séparation (mode ultra du TIMS-TOF, voir Chapitre II). Malgré l'utilisation du mode ultra, les cinq peptides de référence ne sont pas totalement séparés. Néanmoins, au vu de l'EIM à  $m/z$  503, l'espèce présente dans l'échantillon  $(L+E)/SiO_2$  semble correspondre à une combinaison de plusieurs séquences, avec un pic majoritaire correspondant à ELEL (et/ou LELE).

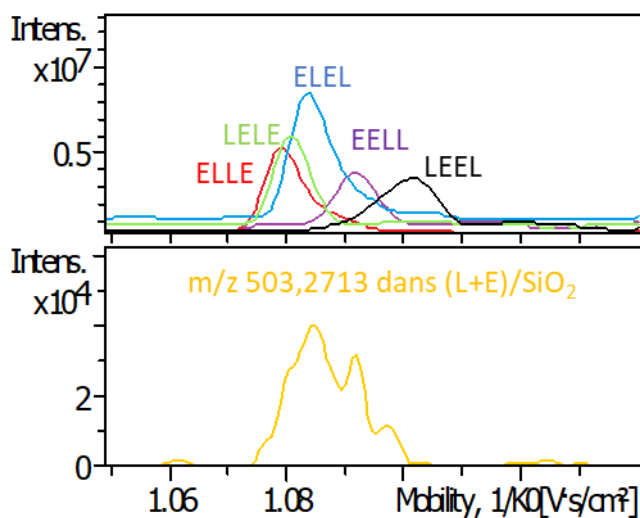


Figure VI-7. Comparaison des mobilogrammes extraits des térapeptides de référence et de l'ion correspondant  $[E_2L_2]H^+$  dans l'échantillon  $(L+E)/SiO_2$  activé à 160 °C (ESI positif, instrument TIMS-TOF). Le mode ultra (mobilité de plus haute résolution) a été utilisé avec une fenêtre de mobilité réduite  $1/K_0$  1-1,15 V·s/cm<sup>2</sup>.

En tout état de cause, même s'il est souvent difficile d'attribuer une séquence unique pour une stœchiométrie donnée, les expériences MS/MS sur le système Leu + Glu présentent des indices de sélectivité pour la plupart des cas étudiés. Concernant la position que préfère adopter le(s) résidu(s) Glu au sein d'une chaîne peptidique intacte, pour les stœchiométries  $EL_2$ ,  $E_2L_2$  et peut être  $EL_3$ , il(s) semble(nt) avoir tendance à se situer en N-terminal. En outre, si les séquences (au moins partielles) ont pu être déterminées par MS/MS, les différences constatées entre les composés de référence et les oligopeptides formés par polymérisation doivent sans doute s'expliquer par la coexistence de régioisomères (voir discussion ci-dessous).

## II.2. Sélectivité de séquence des formes déshydratées

Dans nos conditions d'activation standard, les peptides déshydratés, qui sont *a priori* issus de réactions secondaires (voir Chapitre IV), sont les espèces majoritaires. La détermination de leurs séquences est alors un point capital pour obtenir des informations sur une sélectivité de séquence de la condensation. Comme évoqué précédemment, les formes déshydratées, et notamment les dipeptides DKP, peuvent être impliquées dans les processus de condensation d'oligopeptides plus longs (voir Chapitre IV).

L'étude MS/MS de ces formes déshydratées s'est révélée complexe. En effet, d'une part plusieurs structures déshydratées sont possibles pour une séquence donnée, d'autre part les structures envisagées diffèrent en fonction de la composition en acides aminés. Si pour Glu on a fait l'hypothèse d'une cyclisation interne en pyroGlu, d'autres formes doivent être envisagées pour les oligopeptides formés à partir d'autres acides aminés tels que Asp pour lequel une cyclisation interne en succinimide est plus probable (voir Chapitre V). En outre, d'autres structures peuvent être proposées, indépendamment de la composition de l'oligopeptide. D'après une étude antérieure de Liu et al<sup>268</sup> sur les produits d'activation thermique d'un peptide modèle, l'angiotensine II (un octapeptide contenant un résidu Asp en N-terminal), plusieurs mécanismes de déshydratation sont possibles : une cyclisation totale du peptide, une perte d'eau sur la chaîne latérale (dans notre cas, très plausible avec des résidus Asp ou Glu, sous forme pyro ou imide), ou encore des mécanismes de cyclisation interne *via* la formation d'un cycle à 6, soit des formes DKP ou oxazolone (ces structures sont illustrées dans la Figure VI-8). Cette étude a surtout prouvé la difficulté d'élucider ces structures. En s'appuyant sur l'analyse LC/MS/MS des produits d'activation thermique et de composés de référence cycliques, les auteurs ont déterminé que les peptides issus de décompositions thermiques étaient principalement linéaires (les formes cycliques étant très minoritaires) et par ailleurs ils ont supposé que la perte d'eau provenait majoritairement de la chaîne latérale du résidu Asp, par formation d'un succinimide.

A titre d'exemple, les différentes structures des formes déshydratées possibles pour un tripeptide de séquence ELL sont montrées en Figure VI-8.

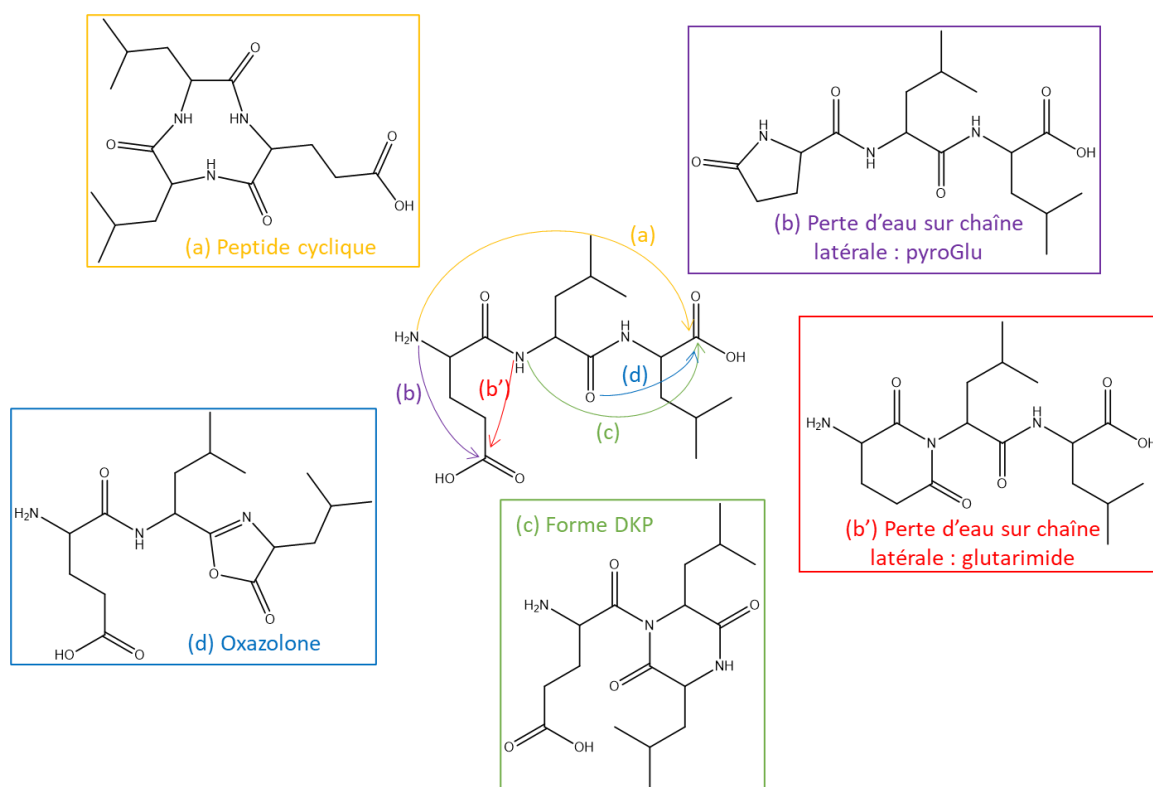


Figure VI-8. Structures des formes déshydratées envisagées à partir d'une séquence ELL : une cyclisation totale du peptide (a), une perte d'eau sur la chaîne latérale du résidu Glu (b et b'), ou encore des mécanismes de cyclisation interne via la formation d'un cycle à 6 et formation de DKP (c) ou d'oxazolone (d).

Dans cette partie, nous nous intéresserons uniquement aux informations sur la détermination de séquences. L'étude structurale des différentes formes déshydratées sera traitée ultérieurement (voir § IV).

## II.2.1. Comparaison avec les peptides de référence déshydratés pour le système Leu + Glu

### II.2.1.1. Synthèse de peptides de référence déshydratés

Les peptides déshydratés n'étant pas disponibles commercialement, nous avons tenté d'utiliser la procédure de dépôt/activation sur silice pour les synthétiser. Pour cela, un peptide intact de référence est déposé sur silice puis activé à 160 °C ; les produits sont ensuite désorbés selon le protocole décrit précédemment (Chapitre II). La vérification de la formation de l'espèce déshydratée a été effectuée par spectrométrie de masse FTICR. Le Tableau VI-1 regroupe les séquences de peptides testées et indique quelle(s) forme(s)



(intacte ou déshydratée) sont obtenues ; si les deux formes coexistent, il est précisé quelle est la forme majoritaire détectée. Afin d'établir s'il existe d'importantes différences d'efficacité d'ionisation entre ces deux formes, un mélange des formes intactes et déshydratées du peptide ELEL a été préparé en quantité équimolaire. Il en résulte que les efficacités d'ionisation des deux formes sont similaires dans ce cas (voir Annexe 17). Nous ferons l'hypothèse que cette conclusion est valable pour l'ensemble des systèmes testés.

*Tableau VI-1. Produits de l'activation thermique des peptides de référence. La présence de chacune des deux formes (intacte (de départ) et déshydratée) selon l'analyse par FTICR est indiquée. Les abréviations « maj » et « min » indiquent quelle forme est majoritaire ou minoritaire sur le spectre de masse.*

Séquence du peptide	Forme intacte	Forme déshydratée
EEL	-	oui
ELE	-	oui
LEE	maj	min
ELL	-	oui
LEL	maj	min
LLE	oui	-
ELEL	min	maj
ELLL	min	maj

On peut constater que l'activation thermique d'un peptide peut former dans nos conditions, un produit déshydraté de façon quantitative à la condition que le résidu Glu soit en position en N-terminale (séquences surlignées dans le Tableau VI-1). Les formes déshydratées sont formées plus ou moins efficacement selon les cas (comparer les séquences LEL et LLE dans le Tableau VI-1). La formation quantitative de formes déshydratées dans le cas de peptides contenant le résidu Glu en position N-terminale corrobore l'hypothèse selon laquelle la déshydratation implique la formation de résidus pyroGlu par cyclisation interne.

#### II.2.1.2. Comparaison des peptides de référence avec les espèces dans le mélange

Les spectres MS/MS, acquis sur l'instrument FTICR, des formes déshydratées obtenues par activation thermique des peptides de référence ELL et LEL (LLE n'ayant pas produit de forme

déshydratée) sont comparés à celui de l'espèce  $[EL_2-H_2O]H^+$  détectée dans le mélange de condensation du système Leu + Glu dans la Figure VI-9. Les voies de fragmentation de ces peptides de référence activés correspondent à celles couramment observées pour les peptides : les ions de séquences  $b_i$ ,  $y_i$  (voir ci-dessous) ainsi que la perte de 46 u qui correspond à une perte de  $(H_2O+CO)^{269}$ .

Pour l'espèce  $(ELL-H_2O)$ , les fragments détectés peuvent être expliqués par les structures obtenues par cyclisation interne du résidu Glu, illustrées à la Figure VI-8. Elles correspondent à une cyclisation de la chaîne latérale du résidu Glu sur l'amine primaire précédente ou l'amine secondaire suivante. Ces deux formes produisent les mêmes ions fragments  $y_1$  et  $b_2$  (Figure VI-9a). Bien que la déshydratation du tripeptide de référence de séquence LEL soit moins favorable (l'espèce intacte reste majoritaire, voir Tableau VI-1), l'espèce déshydratée obtenue a pu être fragmentée, et les fragments détectés suggèrent un mélange des deux oligopeptides linéaires (formes pyroGlu et glutarimide) illustrées sur la Figure VI-9b. Il doit être noté qu'il y a une ambiguïté sur l'attribution du fragment  $b_2$  à  $m/z$  225. En effet il pourrait aussi correspondre à une perte d'eau à partir de  $m/z$  243, les masses exactes d'un fragment  $b_2$  et d'un  $y_2-H_2O$  étant identiques pour un même motif (même composition en acides aminés). Bien qu'il soit difficile d'attribuer sans ambiguïté une structure aux peptides de référence déshydratés, on peut établir que le spectre MS/MS de l'espèce de stœchiométrie  $[EL_2-H_2O]H^+$  dans le mélange de désorption (Figure VI-9c) ne résulte pas uniquement des produits de déshydratation des deux peptides de référence. En effet, si certains ions fragments attestent de la présence des deux séquences ELL et LEL (probablement sous forme de pyroGlu et/ou glutarimide), la détection d'un ion fragment à  $m/z$  227 (non observé pour les peptides de référence activés) suggère la présence d'une forme supplémentaire. Cet ion fragment peut s'expliquer de deux manières ; soit un ion fragment  $b_2(LL)$ , soit son ion isomère  $y_2(LL-H_2O)$ . Afin de distinguer ces deux ions on peut s'appuyer sur l'observation des pertes consécutives : on détecte une perte de 45 u ( $HCONH_2$ ) et aucune perte de 28 u (CO). Une perte de 45 u étant caractéristique d'une structure DKP (voir plus loin, § III.1, Figure VI-22), on déduit que l'ion  $m/z$  227 est un ion fragment  $y_2(LL-H_2O)$ .

Ces observations laissent supposer l'existence d'une forme déshydratée de séquence ELL de structure DKP (type (c) de la Figure VI-8), différente des produits d'activation thermique des peptides de référence. La présence de cette espèce est très instructive. En effet, une

structure de ce type n'étant pas formée par activation thermique du tripeptide de référence ELL, sa présence dans les produits de polymérisation pourrait suggérer que la stœchiométrie ( $EL_2-H_2O$ ) n'est pas forcément un produit de décomposition thermique de l'espèce  $EL_2$  intacte (contrairement aux formes pyroGlu et glutarimide), mais correspondrait à un oligopeptide issu de la polymérisation, par exemple par réaction d'un dipeptide DKP(LL) avec un monomère de Glu.

En conclusion, l'espèce  $[EL_2-H_2O]H^+$  dans le mélange du système Leu + Glu s'explique donc par la combinaison de plusieurs structures, provenant de deux séquences différentes (ELL et LEL), ce qui est confirmé par des expériences LC/MS/MS (voir ci-dessous).

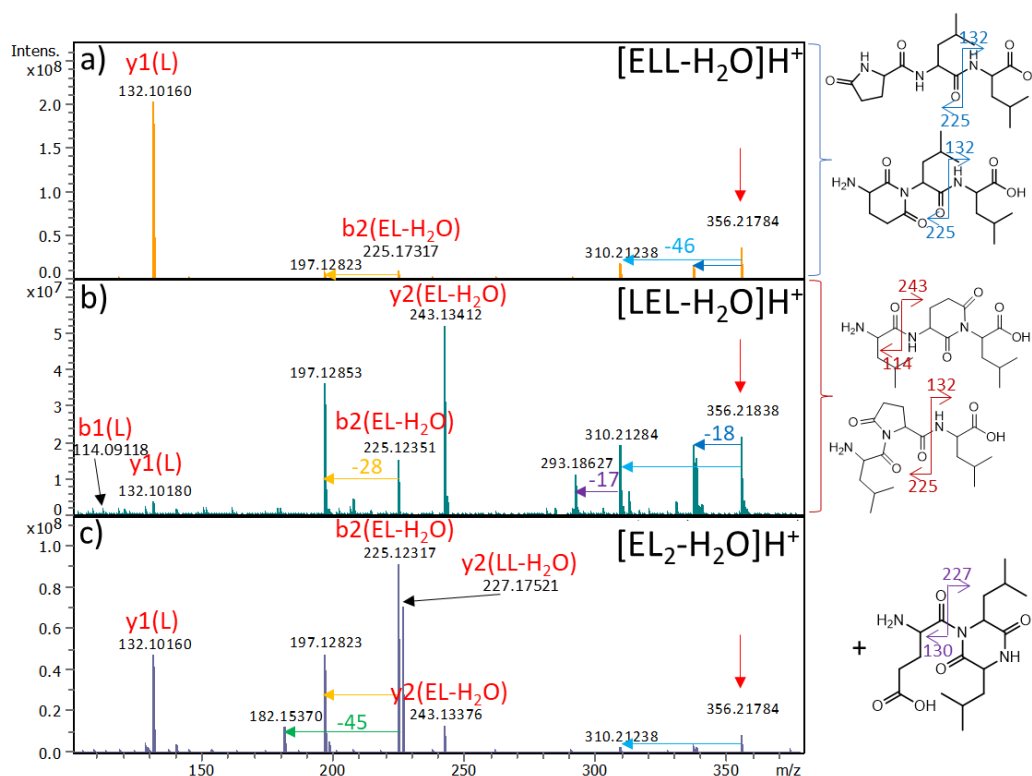


Figure VI-9. Spectres MS/MS des peptides de références activés a) ELL (Elab = 6 eV) et b) LEL (Elab = 15 eV) et c) de l'espèce  $[EL_2-H_2O]H^+$  dans le mélange de condensation du système Leu + Glu (Elab = 14 eV) (ESI positif, instrument FTICR). Les structures correspondant aux voies de fragmentation observées sont montrées à droite.

De la même manière, des expériences MS/MS sur d'autres espèces ont été effectuées puis comparées aux spectres MS/MS des produits d'activation des peptides de référence. Ainsi, pour le tripeptide déshydraté  $[E_2L-H_2O]H^+$ , les fragments observés ne correspondent à aucun des produits d'activation des peptides de références EEL, ELE et LEE (voir Annexe 18). Les spectres MS/MS des térapeptides  $[EL_3-H_2O]H^+$  et  $[E_2L_2-H_2O]H^+$  ont aussi été comparés à

ceux des produits d'activation des peptides de références ELLL et ELEL respectivement (voir Annexe 19). Dans les deux cas, les voies de fragmentation des peptides de références ne correspondent pas à celles de l'ion présent dans le mélange de condensation de l'échantillon (L+E)/SiO<sub>2</sub>.

On peut conclure de façon générale que plusieurs formes en termes de séquences (avec une localisation de Glu plutôt du côté N-terminal), mais aussi plusieurs structures déshydratées doivent être présentes en mélange dans la solution de désorption.

## II.2.2. Apport des techniques de séparation en couplage

Nous avons effectué l'analyse par LC/MS du système Leu + Glu (Figure VI-1). L'EIC de l'ion de rapport m/z 356, correspondant à l'espèce [EL<sub>2</sub>-H<sub>2</sub>O]H<sup>+</sup> est montré Figure VI-10. Plusieurs pics sont observables à différents temps de rétention, attestant de la présence de plusieurs formes et/ou séquences, déjà suggérée par les résultats MS/MS obtenus sur FTICR (Figure VI-9). Les spectres MS/MS des pics chromatographiques les plus intenses sont montrés sur la Figure VI-10.

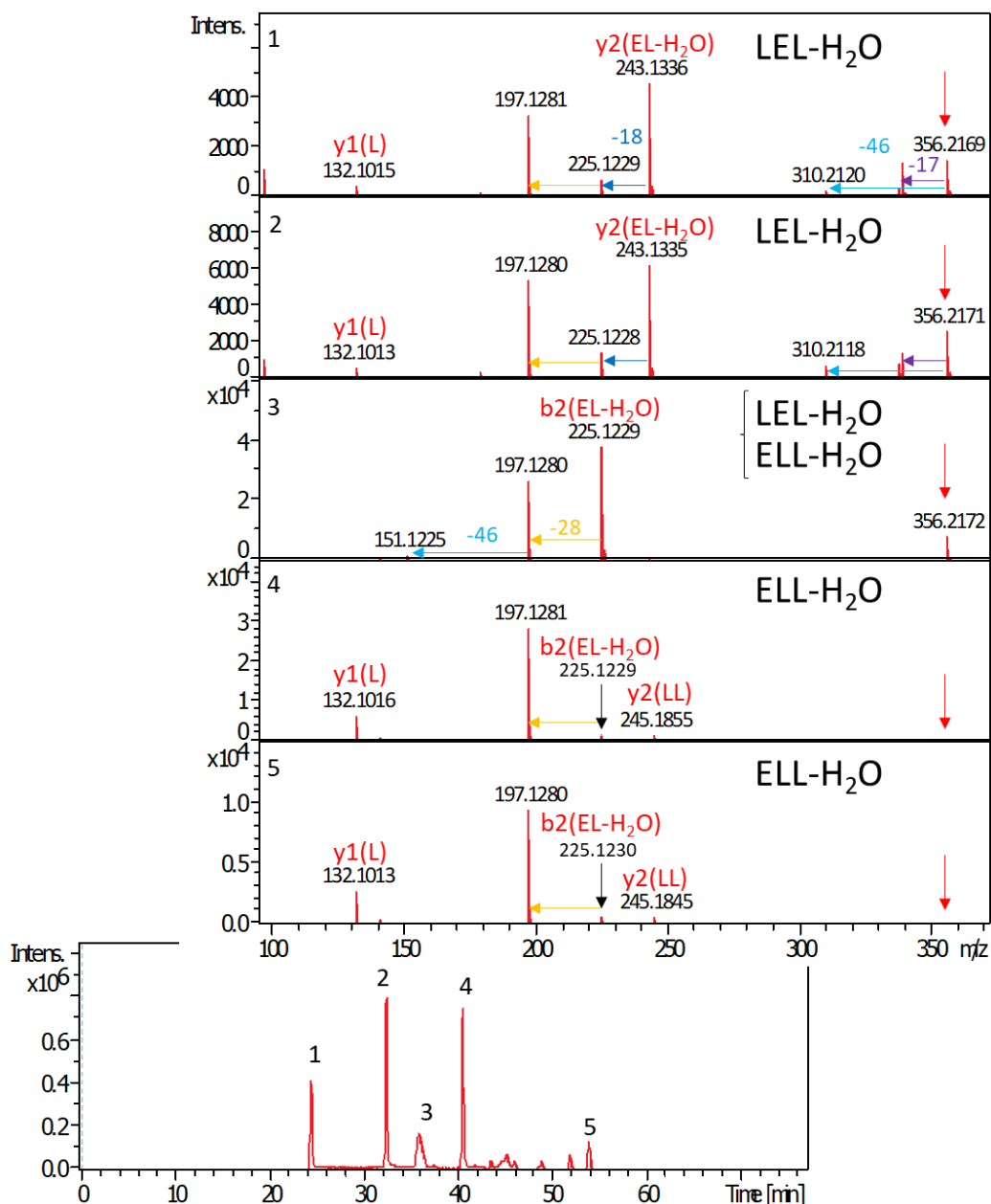


Figure VI-10. Chromatogramme extrait (EIC) de  $m/z$  356  $[EL_2-H_2O]H^+$  et spectres MS/MS des différents pics obtenus à partir de l'expérience LC/MS/MS de (L+E)/SiO<sub>2</sub> activé à 160 °C (ESI positif, 26,4 eV, instrument TIMS-TOF). Les séquences probables sont indiquées sur chaque spectre.

Pour les cinq pics majoritaires, les spectres MS/MS ont été acquis. D'après les ions fragments de série y et b détectés, différentes séquences peuvent être envisagées. Ainsi, les spectres MS/MS notés 1 et 2 (identiques) peuvent correspondre à deux formes déshydratées de structures différentes de la séquence LEL. De la même manière, les spectres 4 et 5 sont conformes à des formes déshydratées d'une séquence ELL. Quant au spectre 3, en l'absence de détection de fragment y2, il est impossible de déterminer si la forme déshydratée

provient d'une séquence LEL et/ou ELL. La présence pour une séquence donnée de plusieurs formes déshydratées est ainsi démontrée. En outre, la similitude des spectres 1 et 2, et 4 et 5, suggère la présence de structures proches (voir discussion plus loin). Dans le cas de cette espèce, la mobilité ionique n'apporte pas d'informations supplémentaires en raison d'une séparation partielle des formes détectées par LC/MS.

### III. Etude de la sélectivité de séquence des produits de condensation

Des expériences MS/MS ont été également menées sur les autres systèmes activés (Leu + Gln, GAVD, GAVE, etc ...voir Chapitre V). Cependant en l'absence de peptides de référence, seules des séquences hypothétiques peuvent être proposées, au vu des voies de fragmentation.

#### III.1. Séquençage par MS/MS et position de l'acide aminé trifonctionnel

##### III.1.1. Comparaison des systèmes contenant Glu ou Asp

Une analyse systématique des ions fragments pour les différents oligopeptides formés à partir du système GAVE permet d'établir que, pour ce système, le résidu acide glutamique a tendance à être localisé en N-terminal. A titre d'exemple, pour le tripeptide de stœchiométrie  $[EVA]H^+$  au moins deux séquences sont présentes selon les ions fragments détectés, et ce sont des séquences dans lesquelles le résidu E est plutôt en position N-terminale (Figure VI-11a)\*. Néanmoins seul un séquençage partiel est obtenu et il n'est pas possible de confirmer sans ambiguïté la présence d'une séquence particulière. Pour d'autres oligopeptides avec le résidu Glu, notamment pour  $[EG]H^+$ ,  $[EVG]H^+$  et  $[EGA]H^+$ , il semble également que les séquences avec Glu en N-terminal sont privilégiées, sans que l'on puisse attribuer une séquence particulière avec certitude (voir en Annexe 20).

---

\*Ces oligopeptides n'ont pas pu être étudiés en LC/MS car ils sont peu ou pas retenus du fait de leur polarité. De même, la mobilité ionique n'a pas permis de séparer différentes séquences.

Globalement, on a observé que le résidu Asp était plus incorporé au sein des oligopeptides que le résidu Glu (voir Chapitre V). On peut se demander si cette différence s'accompagne d'un positionnement préférentiel des acides aminés trifonctionnels au sein de la chaîne peptidique. Dans le cas de polymérisations à l'état massique, les travaux précurseurs de Fox et Harada<sup>237,238</sup> ont montré, en utilisant la méthode de Sanger pour connaître la composition des peptides, que résidu Glu a tendance à être localisé en N-terminal. En revanche, ce n'est pas le cas pour Asp : la répartition de celui-ci doit être plus aléatoire au sein des oligopeptides.

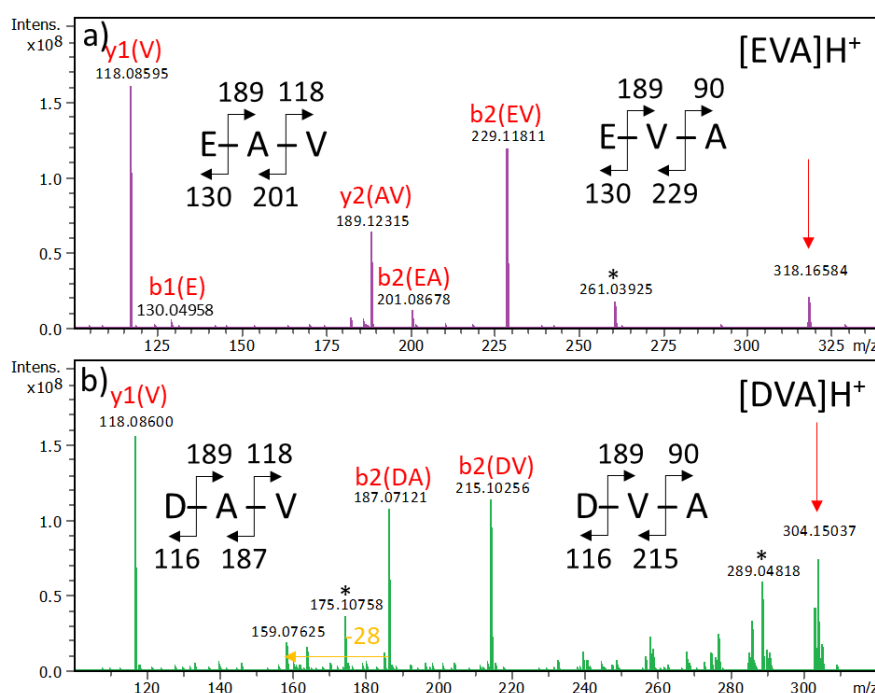


Figure VI-11. Spectres MS/MS des tripeptides a) [EVA]H<sup>+</sup>, b) [DVA]H<sup>+</sup> (ESI positif, 6 eV, instrument FTICR). Parmi toutes les séquences possibles, celles pouvant expliquer les fragments observés sont montrées. Les pics notés d'un astérisque n'ont pas pu être attribués.

L'EIM de l'ion m/z 348, correspondant au tripeptide [D<sub>2</sub>V]H<sup>+</sup> ainsi que le spectre MS/MS correspondant sont reportés sur la Figure VI-12. Un seul pic de mobilité est observé mais les spectres MS/MS montrent la présence de plusieurs séquences. Les mêmes observations sont faites pour les espèces [D<sub>2</sub>A]H<sup>+</sup> et [D<sub>2</sub>G]H<sup>+</sup> (voir Annexe 21) : les fragments majoritaires correspondent à la séquence DDY (Y étant l'acide aminé aliphatique). Une séquence YDD est présente, mais en moindre abondance. Sur la base de ces résultats, les espèces obtenues

sont des copolymères à blocs plutôt que des copolymères alternés. Les résidus Asp sont successifs. D'autre part, ce bloc D-D préfère être localisé du côté N-terminal.

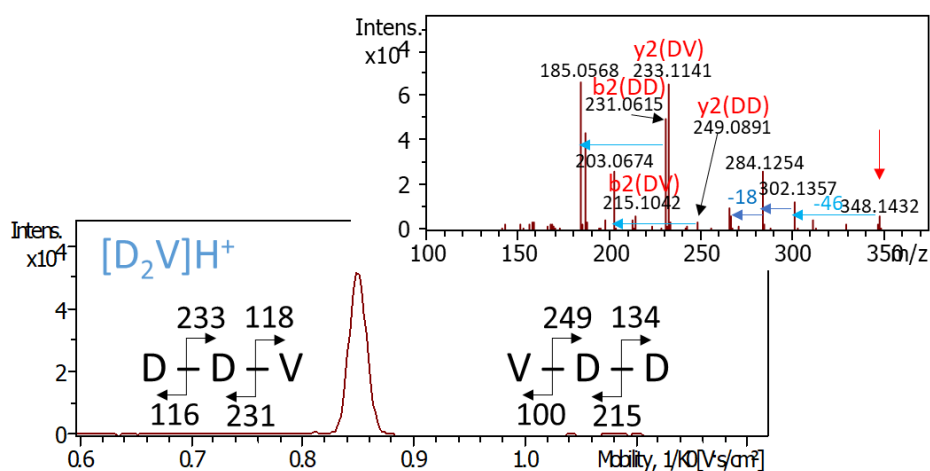


Figure VI-12. Mobilogramme extrait et spectre MS/MS de l'ion  $m/z$  348 correspondant à l'espèce  $[D_2V]H^+$ , obtenus pour l'analyse IM-MS/MS du système GAVD activé à 130 °C (ESI positif, CID Elab = 25 eV, instrument TIMS-TOF).

En ce qui concerne les systèmes contenant le résidu Glu, peu d'espèces comportant plus d'un résidu Glu sont formées (voir Chapitre V). Nous avons montré auparavant qu'il est difficile d'attribuer une séquence au tripeptide  $E_2L$ . Concernant, le pentapeptide  $[E_2L_3]H^+$ , la Figure VI-13 reporte les voies de fragmentation obtenues par analyse IM-MS/MS. Les fragments détectés sont compatibles avec une séquence LELEL, correspondant à une polymérisation alternée (la séquence ELLEL serait également possible).

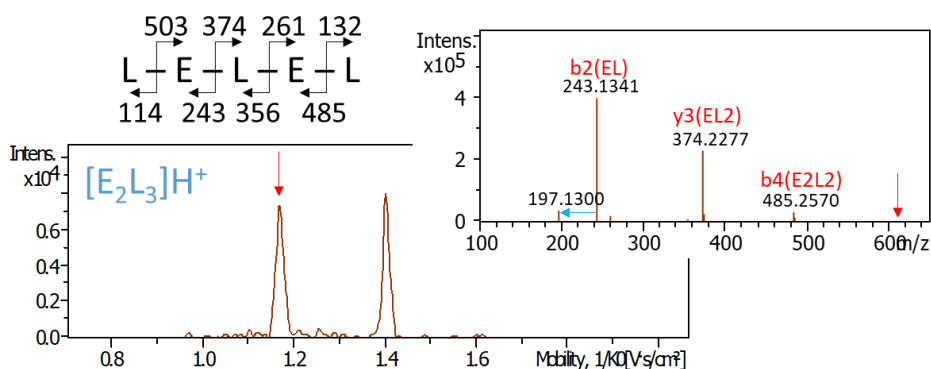


Figure VI-13. Mobilogramme extrait et spectre MS/MS de l'ion  $m/z$  616 correspondant à l'espèce  $[E_2L_3]H^+$ , obtenus pour l'analyse IM-MS/MS de  $(L+E)/SiO_2$  activé à 160 °C (ESI positif, CID Elab = 30 eV, instrument TIMS-TOF). Seul le premier pic a pu être fragmenté.



Il apparaît donc une différence de comportement entre les systèmes contenant le résidu Asp et ceux contenant le résidu Glu. Dans les deux cas, le résidu acide trifonctionnel semble se localiser préférentiellement en N-terminal mais le mode de polymérisation n'est pas le même. Il est plutôt alterné dans les systèmes contenant Glu, et plutôt à blocs pour ceux contenant Asp. On peut sans doute imaginer une justification en termes de sélectivité thermodynamique : Asp ayant un groupement méthyle de moins sur la chaîne latérale, la gêne stérique moindre s'oppose moins à une polymérisation par blocs, dans laquelle deux résidus Asp sont accolés. Mais à ce stade, rien ne permet d'exclure qu'il s'agisse d'un effet cinétique. Des informations plus poussées sur le mécanisme de polymérisation seraient nécessaires.

Le système Leu + Gln produisant des espèces similaires à celles du système Leu + Glu, son étude n'est pas discutée ici.

### III.1.2. Le cas de l'acide $\beta$ -glutamique

Des expériences MS/MS ont été effectuées sur l'échantillon (L+ $\beta$ E)/SiO<sub>2</sub> activé à 160 °C en infusion directe sans séparation sur l'instrument FTICR et après séparation par mobilité ionique sur l'instrument TIMS-TOF.

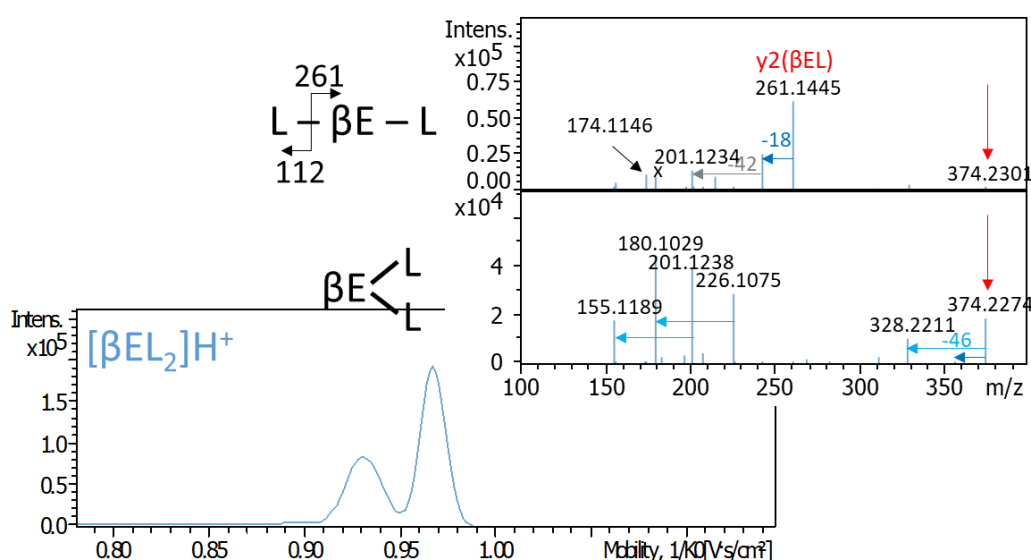


Figure VI-14. Mobilogramme extrait et spectres MS/MS de l'ion  $m/z$  374 correspondant à l'espèce  $[\beta EL_2]H^+$ , obtenus pour l'analyse IM-MS de (L+ $\beta$ E)/SiO<sub>2</sub> activé à 160 °C (ESI positif, CID avec  $E_{lab} = 30$  eV, instrument TIMS-TOF).

Le mobilogramme extrait pour l'ion  $m/z$  374 montre deux pics de mobilité bien distincts qu'il est possible de fragmenter (Figure VI-14). Les ions fragments obtenus pour chacun permettent de déterminer deux séquences différentes. Concernant le premier pic, la détection de l'ion fragment  $\gamma_2$  et d'ions fragments à  $m/z$  201 et  $m/z$  174, qui sont des fragmentations consécutives de  $\gamma_2$ , suggèrent la présence de la séquence L- $\beta$ E-L<sup>§§</sup> (les voies de fragmentation correspondantes sont expliquées en Annexe 22, par exemple la perte de 87 u correspondant au groupement  $C_3H_5O_2N$ , localisé sur le résidu  $\beta$ Glu). Les fragments obtenus pour la deuxième forme ne correspondent pas à des ions de séquences et peuvent traduire la présence d'une structure « ramifiée » permise par la structure particulière de  $\beta$ Glu. Les voies de fragmentation envisagées menant aux fragments  $m/z$  226 et  $m/z$  201 sont montrées en Annexe 22 également. Le spectre MS/MS obtenu sur FTICR sans séparation préalable est un mélange des deux spectres correspondant aux deux pics du mobilogramme. L'acide aminé trifonctionnel  $\beta$ Glu a donc un comportement différent. Bien que toutes les séquences possibles ne soient pas présentes, la présence des deux explicitées au-dessus montre un manque de sélectivité pour une séquence particulière, au contraire des acides aminés Asp, et Glu. Ces mêmes observations sont faites pour le tripeptide  $[\beta E_2 L]H^+$  pour lequel le spectre MS/MS montre un mélange des deux séquences  $\beta E-\beta E-L$  et  $L-\beta E-\beta E$  (voir Annexe 23).

Notons que les espèces décalées de +0,98 u (voir Chapitre V, § I.2.3) n'ont pas pu être étudiées, car il n'a pas été possible de les distinguer du pic isotopique  $^{13}C$  des espèces peptidiques classiques.

### III.2. La question de la régiosélectivité

Dans le cas des acides aminés trifonctionnels protéinogènes Asp, Glu et Gln, différents régioisomères peuvent être produits selon le groupement acide carboxylique (amide dans le cas de Gln) participant à la formation de la liaison peptidique (voir Chapitre V). Ainsi, Pivcová<sup>270</sup> a mis en évidence que les polymères d'acide aspartique obtenus par chauffage de Asp dissous dans une solution d'acide phosphorique (polyAsp thermiques) contenaient des liaisons  $\alpha$  aussi bien que  $\beta$ . En spectrométrie de masse, il est difficile de mettre en évidence

---

<sup>§§</sup> Dans ce cas, pour la notation des séquences, les résidus sont séparés d'un tiret pour faciliter la lecture.

des régioisomères sans disposer de composés de référence pour comparaison. Néanmoins, des méthodes d'activation comme l'ECD (electron capture dissociation) et l'ETD (electron transfer dissociation) ont permis de différencier ces résidus isomériques dans les peptides ou protéines<sup>271</sup>. Cette méthode a notamment été utilisée pour différencier Asp et iso-Asp dans des  $\beta$ -peptides<sup>272,273</sup> et pour la discrimination de formes  $\alpha$ Glu et  $\gamma$ Glu<sup>274</sup>. En revanche l'ECD nécessite d'avoir des ions multichargés, qui ne sont pas produits dans notre cas. Cependant, des études de fragmentation en mode CID ont montré une différenciation possible des peptides contenant  $\alpha$ Asp et  $\beta$ Asp<sup>275,276</sup>, dans le cas où Asp est en N-terminal. Ainsi, dans le cas des  $\alpha$ -peptides, on observe une perte préférentielle de  $\text{H}_2\text{O}$  suivie d'une perte de  $\text{NH}_3$  tandis que dans le cas de  $\beta$ Asp, c'est une perte de  $(\text{H}_2\text{O}+\text{CO})$  qui est majoritairement observée (voir Figure VI-15 a et b). De même, Harrison<sup>277</sup> a montré la possibilité de distinguer par MS/MS les dipeptides contenant des résidus  $\alpha$ Glu et  $\gamma$ Glu en N-terminal. Les premiers éliminent préférentiellement  $\text{H}_2\text{O}$  tandis que les seconds éliminent  $\text{NH}_3$  (voir Figure VI-15). Le même comportement a été rapporté pour les tripeptides avec un Glu en N-terminal. Ces voies de fragmentation théoriques ont été vérifiées expérimentalement avec l'analyse d'un  $\gamma$ -peptide de référence, le glutathion.

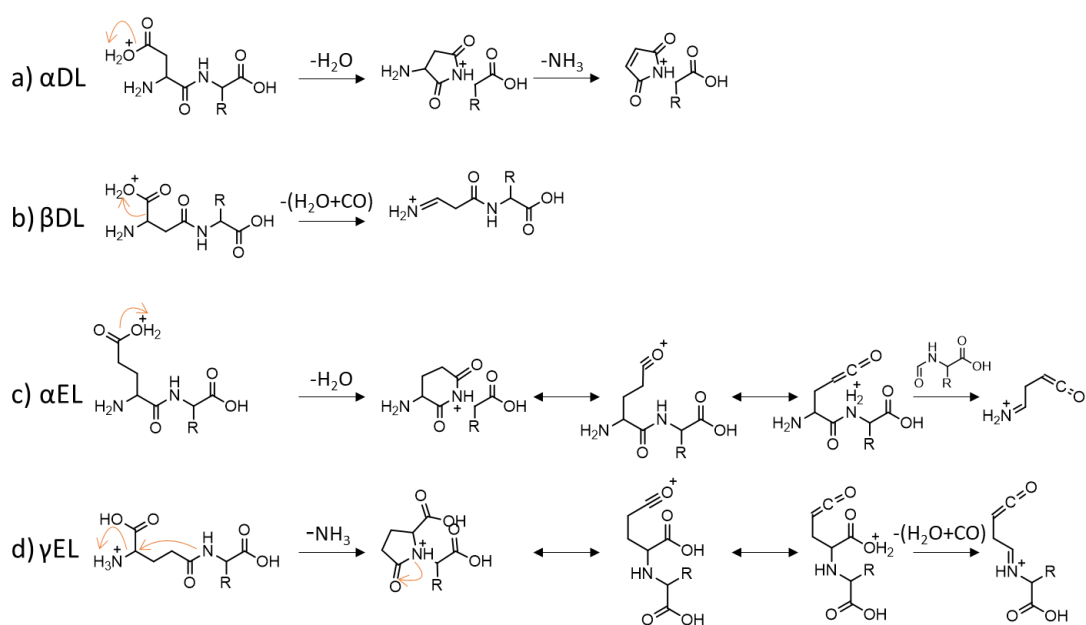


Figure VI-15. Voies de fragmentation attendues à partir des dipeptides régioisomères a)  $\alpha$ DL, b)  $\beta$ DL, c)  $\alpha$ EL et d)  $\gamma$ EL (pour R = C<sub>4</sub>H<sub>10</sub>), adapté des travaux de Harrison<sup>275,277</sup>.

### III.2.1. Oligopeptides comportant le résidu Asp

Une comparaison avec la bibliographie est possible mais doit être effectuée avec prudence car les fragmentations observées dépendent de l'instrument utilisé. La comparaison avec des peptides de référence sur le même instrument doit toujours être privilégiée quand cela est possible. De plus, bien que la perte de 46 u semble « diagnostique » d'une forme  $\beta$ , une perte de 46 à partir de l'ion moléculaire peut également avoir lieu en C-terminal de toute chaîne peptidique.

Les dipeptides intacts  $[DX]H^+$ , avec  $X = G, A, V$  ou  $L$ , obtenus dans les échantillons (GAVD)/SiO<sub>2</sub> et (L+D)/SiO<sub>2</sub>, ont été étudiés. Les spectres MS/MS obtenus sur FTICR pour les trois premiers montrent pour chacun un mélange des formes  $\alpha$  et  $\beta$  (perte de CO<sub>2</sub>H<sub>2</sub> et succession des pertes neutres de 18 et 17 u) (voir Annexe 24). En revanche, pour  $[DL]H^+$ , seules les pertes de CO<sub>2</sub>H<sub>2</sub> et H<sub>2</sub>O sont visibles sur le spectre MS/MS, suggérant la présence majoritaire, si ce n'est unique, d'une forme  $\beta$ . Afin de confirmer ces hypothèses, les dipeptides  $[DV]H^+$  et  $[DL]H^+$  ont été étudiés plus en détails, respectivement, par LC/MS/MS et/ou IM-MS/MS.

La Figure VI-17 montre l'EIC correspondant à  $[DV]H^+$  : trois pics chromatographiques sont observés, indiquant la présence de trois formes différentes. L'interprétation des spectres MS/MS de chacun des pics permet d'accéder à leurs séquences respectives : le pic 1 correspond à une forme  $\alpha DV$ . Tout d'abord la présence de l'ion  $y_1$  à  $m/z$  118 traduit une séquence DV. Ensuite, des pertes successives de 18 u puis 17 u à partir de l'ion moléculaire sont observées. De plus, les pertes consécutives de 18 u et 42 u (-H<sub>2</sub>O suivie de -CH<sub>2</sub>CO, voir Figure VI-16) ne sont envisageables que dans le cas d'une forme  $\alpha$  (chaîne latérale de Asp).

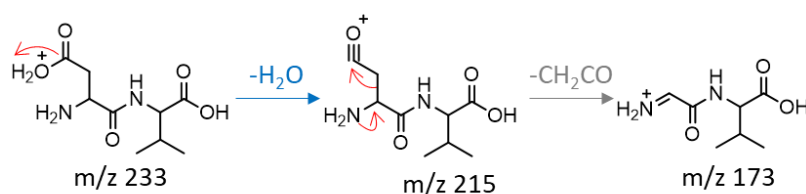


Figure VI-16. Mécanismes des pertes consécutives de 18 et 42 u, caractéristiques d'une forme  $\alpha$ .

Les pics numérotés 2 et 3 ont des voies de fragmentation identiques, suggérant des structures similaires. La fragmentation majoritaire correspond à une perte de (H<sub>2</sub>O+CO), caractéristique de la forme βDV (voir Figure VI-15b).

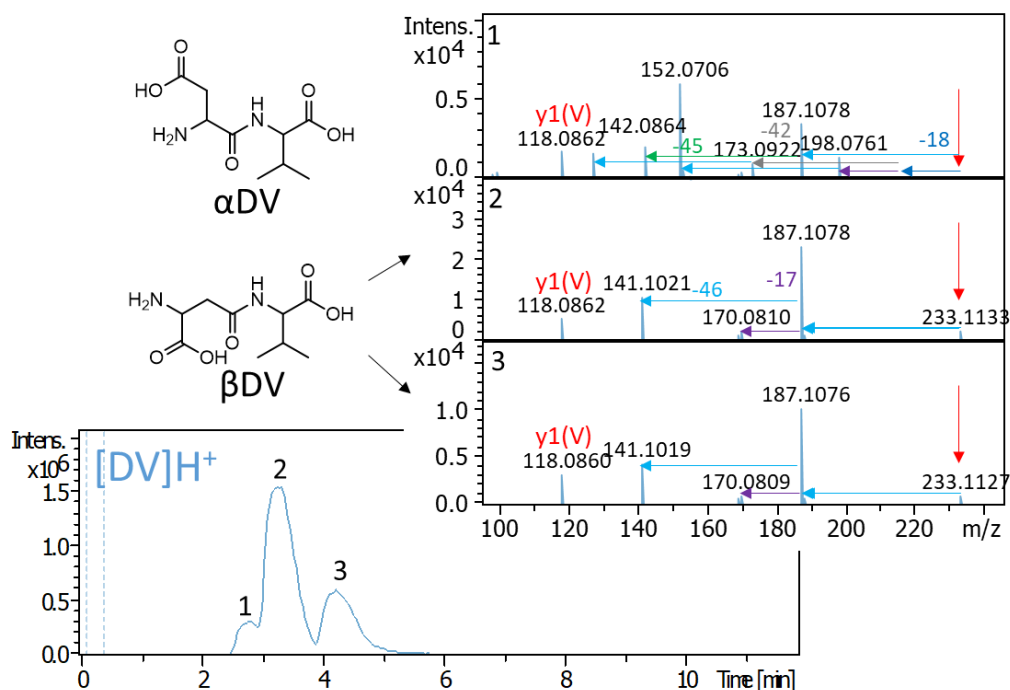


Figure VI-17. Chromatogramme extrait et spectres MS/MS de l'ion  $m/z$  233, correspondant au dipeptide  $[DV]H^+$ , obtenus pour l'analyse LC/MS/MS de GAVD/SiO<sub>2</sub> activé à 160 °C (ESI positif, CID avec  $E_{lab} = 23,9$  eV, instrument TIMS-TOF).

De la même manière, l'analyse du dipeptide  $[DL]H^+$  par IM-MS/MS permet de confirmer les observations faites par MS/MS sur le FTICR (Annexe 24). La Figure VI-18 montre l'EIM de l'ion  $m/z$  247, correspondant à  $[DL]H^+$ . Un seul pic, dont la fragmentation montre uniquement une perte de CO<sub>2</sub>H<sub>2</sub>, est observé. Cela confirme la présence (majoritaire) de la forme βDL.

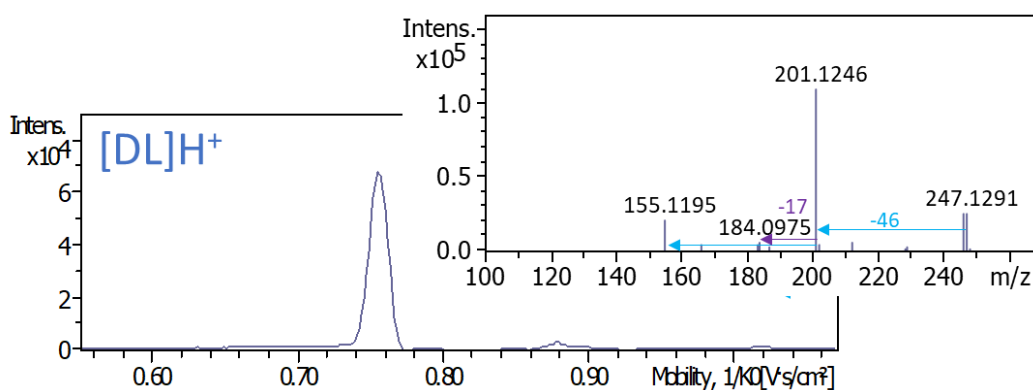


Figure VI-18. Mobilogramme extrait et spectre MS/MS de l'ion  $m/z$  247, correspondant au dipeptide  $[DL]H^+$ , obtenus pour l'analyse IM-MS/MS de  $(L+D)/SiO_2$  activé à 160 °C (ESI positif, CID = 20 eV, instrument TIMS-TOF).

Par ailleurs, ici l'absence de fragments  $y_1$  ne permet *a priori* pas de discriminer les séquences  $\beta DL$  et  $\beta LD$ . En revanche le fragment  $y_1(L)$  est présent sur le spectre MS/MS obtenu sur FTICR et il a été montré que la prédominance de la perte de  $HCOOH$  était caractéristique d'une séquence  $\beta DL$ <sup>275</sup>. Une différence de comportement se dessine en fonction de l'acide aminé aliphatique condensé avec Asp : pour  $[DA]H^+$ ,  $[DG]H^+$  et  $[DV]H^+$  les formes  $\alpha$  et  $\beta$  sont présentes, tandis que pour  $[DL]H^+$  il semblerait que la forme  $\beta$  soit la seule produite, ou au moins très majoritairement. Une explication possible est l'influence de la taille de l'acide aminé aliphatique : l'acide aminé Leu étant plus volumineux que les acides aminés Val, Ala et Gly, il a tendance à se condenser préférentiellement avec le groupement acide le moins encombré, donc celui de la chaîne latérale de Asp, entraînant une forme  $\beta DL$ . De même la forme majoritaire pour  $[DV]H^+$  est la forme  $\beta$ . En revanche les acides aminés aliphatiques plus petits peuvent se condenser indifféremment avec les deux groupements acides de Glu. Il convient de noter ici que pour les quatre dipeptides étudiés, le résidu Asp se situe en N-terminal.

Nous avons tenté de généraliser cette étude à des peptides plus longs. La Figure VI-19 montre l'EIM de  $m/z$  360, correspondant au tripeptide  $[DL_2]H^+$ , et les spectres MS/MS des trois pics de mobilité observés. Les pics étant mal séparés, l'intégration de chacun pour aboutir à leurs spectres MS/MS respectifs n'est pas aisée. En revanche, il est possible de tracer les EIMs de chaque ion fragment pour vérifier leurs appartenances. Les trois formes proposées sont indiquées à côté des spectres MS/MS dans la Figure VI-19. Le pic 1 correspond à une séquence  $\alpha D-L-L$ , le pic 2 à  $L-D-L$  (les formes  $L-\alpha D-L$  et  $L-\beta D-L$  peuvent être

envisagées) et le pic 3 à  $\beta$ D-L-L. En effet, la perte de ( $\text{H}_2\text{O}+\text{CO}$ ), abondante pour le pic 3 à l'inverse des deux premiers composés, est caractéristique d'un  $\beta$ Asp lorsque le résidu Asp est en N-terminal (voir Figure VI-15). L'exemple de ce tripeptide montre une nouvelle fois que le couplage IM-MS/MS peut distinguer efficacement des régioisomères  $\alpha$  et  $\beta$ . En revanche, dans cet exemple, le résidu Asp n'est pas forcément localisé en N-terminal. Cette observation est également faite pour le tétramère  $[\text{D}_2\text{L}_2]\text{H}^+$  pour lequel les séquences DLLD et LDLD semblent formées. Tout comme dans les travaux de Fox et Harada<sup>237,238</sup>, la répartition de Asp au sein des oligopeptides ne favorise pas fortement la position N-terminale.

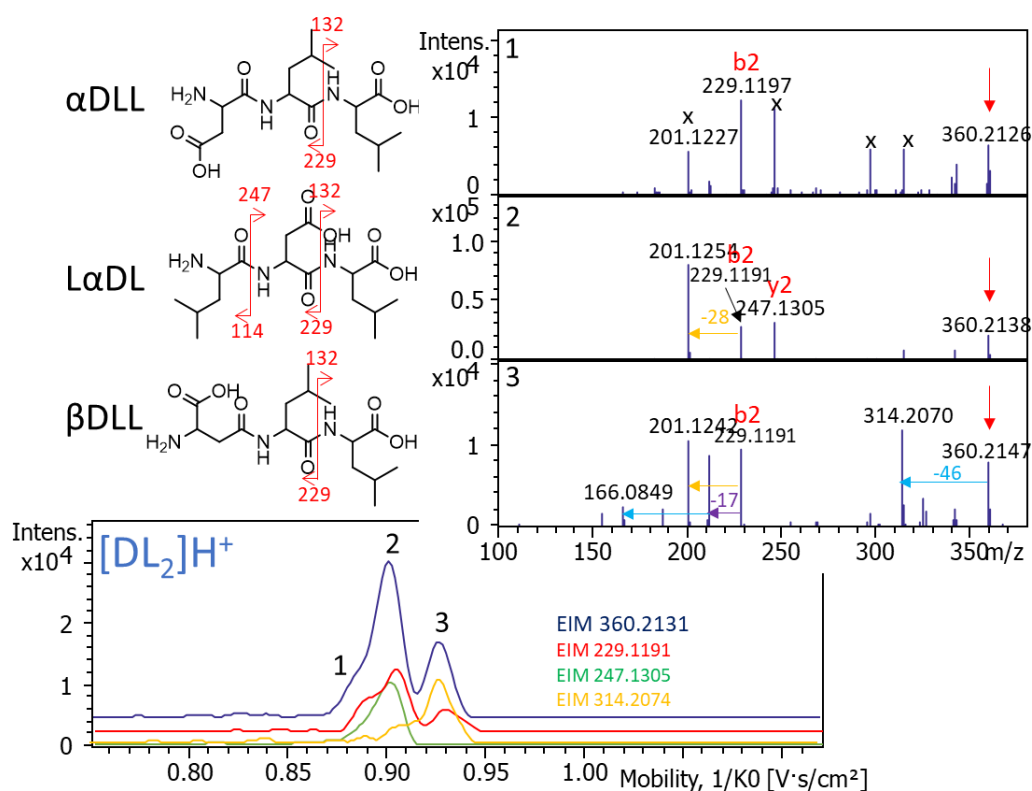


Figure VI-19. Mobilogramme extrait et spectres MS/MS de l'ion  $m/z$  360 correspondant au tripeptide  $[\text{DL}_2]\text{H}^+$ , obtenus pour l'analyse IM-MS/MS de (L+D)/ $\text{SiO}_2$  activé à 160 °C (ESI positif, CID Elab = 20 eV, instrument TIMS-TOF). Les EIMs des ions fragments ont également été extraits. Pour le pic 2, la séquence L $\beta$ DL (non illustrée) est également possible. Les pics notés d'une croix ne sont pas des ions fragments appartenant au spectre MS/MS.

### III.2.2. Comparaison des systèmes Leu + Gln et Leu + Glu du point de vue de la régiosélectivité

De la même manière, nous avons cherché à déterminer si en présence de Glu, des oligopeptides de type  $\gamma$  sont formés lors de la polymérisation. Dans le cas des dipeptides, Harrison<sup>277</sup> a rapporté une perte de 17 u diagnostique d'une forme  $\gamma$  (voir Figure VI-15). En revanche, pour les oligopeptides plus longs, il n'existe pas de données affirmant que cette perte diagnostique reste prédominante.

Une idée pour déterminer si des formes  $\gamma$  sont produites dans l'échantillon (L+E)/SiO<sub>2</sub> est de comparer les oligopeptides produits dans celui-ci avec ceux détectés dans l'échantillon (L+Q)/SiO<sub>2</sub>. Pour rappel, les oligopeptides obtenus dans le système Leu + Gln ont une stœchiométrie ExLy (et non QxLy) et sont donc *a priori* des formes  $\gamma$  (voir discussion du Chapitre V). Dans cette hypothèse, la comparaison des mobilités et des voies de fragmentation de ce système avec le système Leu + Glu renseigne sur les peptides obtenus dans ce dernier système. A titre d'exemple, la Figure VI-20 montre les spectres MS/MS du térapeptide de stœchiométrie [EL<sub>3</sub>]<sup>H+</sup> analysé par IM-MS/MS pour les échantillons (L+Q)/SiO<sub>2</sub> et (L+E)/SiO<sub>2</sub>.

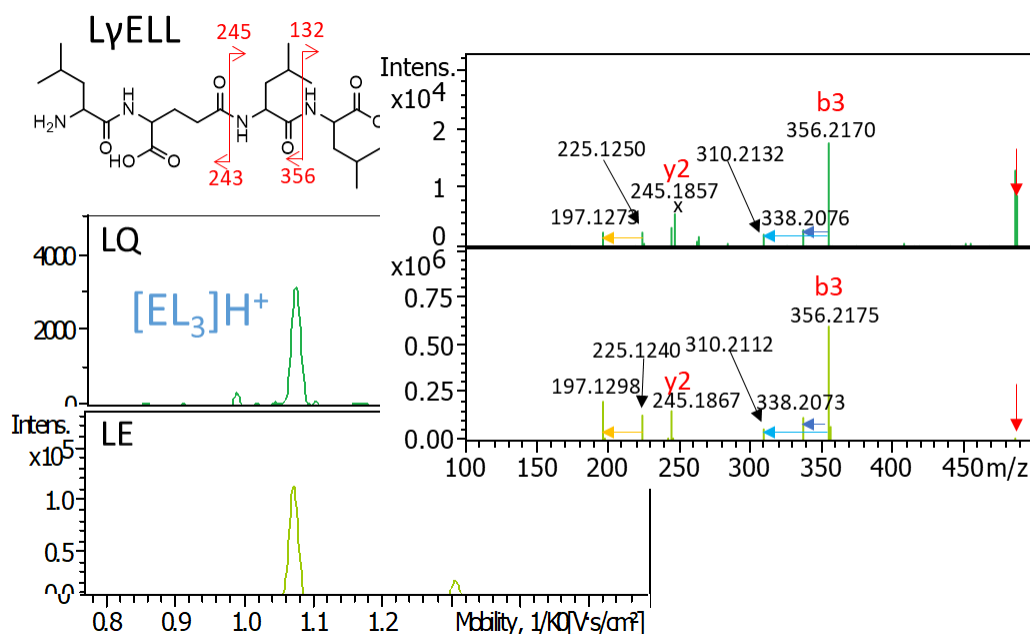


Figure VI-20. Mobilogrammes extraits et spectres MS/MS de l'ion  $m/z$  487 correspondant au térapeptide [EL<sub>3</sub>]<sup>H+</sup>, obtenus pour l'analyse IM-MS/MS des échantillons (L+Q)/SiO<sub>2</sub> et (L+E)/SiO<sub>2</sub> activés à 160 °C (ESI positif, CID Elab = 25 eV, instrument TIMS-TOF). L'ion marqué d'une croix résulte sans doute d'un isolement non suffisant de l'ion précurseur.



La comparaison des mobilités et voies de fragmentation de l'ion  $[EL_3]H^+$  dans les deux systèmes permet d'établir que les espèces formées semblent identiques, suggérant la présence des formes  $\gamma$  majoritaires dans le système Leu + Glu. En effet, tout d'abord les mobilités obtenues pour les deux systèmes sont identiques :  $1/K_0 = 1,067 \text{ V.s.cm}^{-2}$ . Cette valeur est légèrement différente de celles des deux peptides de références de séquence  $L\alpha ELL$  et  $\alpha ELLL$  dont on dispose, qui sont respectivement de 1,076 et 1,097  $\text{V.s.cm}^{-2}$  (voir Annexe 16). De plus les fragmentations des références sont similaires, ce qui laisse présager que la forme dans le système Leu + Glu est une forme  $\gamma$ . Une étude détaillée des ions fragments permet également de valider cette hypothèse : l'absence de l'ion  $b_2$   $m/z$  243 mais la présence de l'ion  $m/z$  225 correspondant à une perte de  $H_2O$  à partir de 243 indiquent que l'ion fragment  $b_2$  est instable. En effet, théoriquement, un ion  $b_2$  est stabilisé par son réarrangement en oxazolone (voir Chapitre I). Or, dans le cas des ions  $b_2(Xxx-\gamma Glu)$ , la formation de l'oxazolone est défavorable ; le cycle hypothétiquement formé serait un cycle à 7, donc moins stable. En revanche l'ion  $b_2$  non détecté subit des pertes consécutives, notamment la perte de  $H_2O$  pour former l'ion fragment  $m/z$  225 (voir Figure VI-21). La séquence majoritaire est donc  $L-\gamma E-L-L$  pour les deux systèmes Leu + Gln et Leu + Glu.

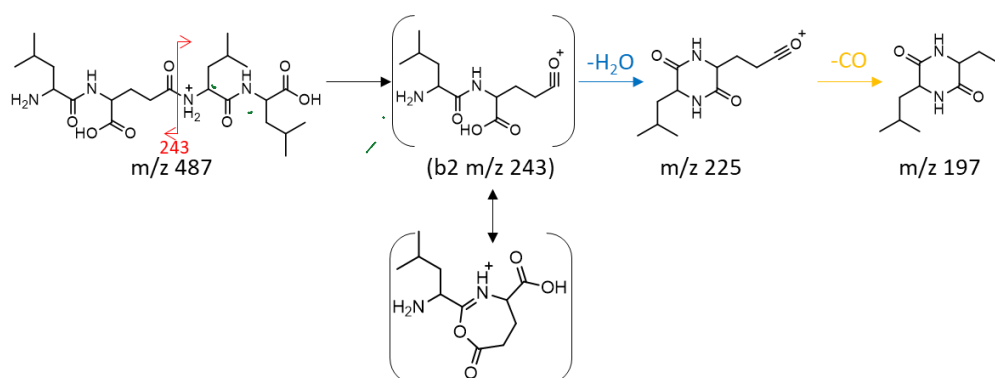


Figure VI-21. Fragmentations successives proposées pour l'ion fragment  $m/z$  243 (absent du spectre MS/MS). Le réarrangement de l'ion  $b_2$  en oxazolone ne semble pas favorable, au contraire des  $\alpha$ -peptides.

Concernant la stœchiométrie  $EL_3$  l'espèce formée est donc une espèce  $\gamma$ . En revanche, pour  $EL_2$  les voies de fragmentation obtenues pour les espèces dans les deux systèmes sont légèrement différentes (voir Annexe 25), suggérant la présence probable des deux régioisomères.

## IV. Etude structurale des formes déshydratées

L'étude des formes déshydratées est également importante pour déterminer l'existence d'une sélectivité de séquence. En effet, il a été montré au Chapitre IV que les formes déshydratées étaient des produits de dégradation des oligopeptides intacts. Les séquences des formes déshydratées doivent donc être prises en compte pour déterminer s'il existe une sélectivité de séquence.

### IV.1. Dipeptides déshydratés

Nous nous sommes d'abord intéressés aux dipeptides déshydratés. Plusieurs études ont affirmé que le dipeptide cyclique de forme DKP (dicétopipérazine) était initialement produit dans des processus de polymérisation sur surface et constituait un intermédiaire de condensation peptidique<sup>56,57,228</sup>. Notre travail a également démontré que le DKP n'était pas une impasse réactionnelle. Il apparaît donc nécessaire de déterminer la structure des dipeptides déshydratés présents dans nos mélanges. Or, si la formation de dipeptides déshydratés est systématiquement détectée dans nos conditions, leur formation peut être due soit à une cyclisation totale du dipeptide (forme DKP), soit à une déshydratation impliquant la condensation d'un groupement acide carboxylique d'une chaîne latérale (cf Chapitre V, § I.1 et I.2).

On peut envisager de discriminer ces structures par MS/MS. Les voies de fragmentation caractéristiques des dipeptides DKP ont été répertoriées<sup>278</sup>, indiquant notamment une perte caractéristique de neutre de 45 u (-HCONH<sub>2</sub>) et/ou de 28 u (CO) par ouverture du cycle (voir Figure VI-22).

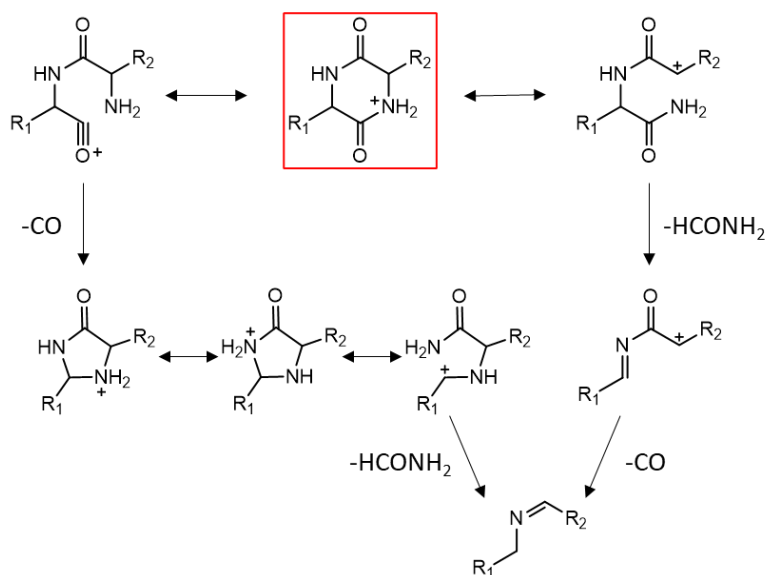


Figure VI-22. Voies de fragmentation d'un dipeptide cyclique, adapté de Guo et al.<sup>278</sup>.

Conformément à la littérature, nous avons observé que, dans le cas des dipeptides de stœchiométrie (XY-H<sub>2</sub>O) sans acide aminé trifonctionnel, une perte de 45 u est la voie de fragmentation majoritaire (comme illustré par l'exemple du composé de référence [cycloAA]H<sup>+</sup> reporté Figure VI-23a et du dipeptide [AV-H<sub>2</sub>O]H<sup>+</sup> présent dans le système GAVD, reporté Figure VI-23c). A l'inverse, lorsqu'un dipeptide déshydraté contient un résidu trifonctionnel tel que Glu, ses voies de fragmentation obtenues par MS/MS ne donnent aucune indication sur son attribution à une forme DKP ou à une forme obtenue par cyclisation interne (par exemple contenant un pyroGlu). En effet, le spectre MS/MS du dipeptide de référence protoné [cycloEE]H<sup>+</sup> montre une perte préférentielle du groupement HCOOH de la chaîne latérale (de 46 u), masquant la perte de 45 u attendue au niveau du cycle (Figure VI-23b). L'absence de perte de 45 u dans les spectres MS/MS des dipeptides déshydratés obtenus dans les systèmes GAVD et GAVE (voir l'exemple des espèces [DV-H<sub>2</sub>O]H<sup>+</sup> et [EV-H<sub>2</sub>O]H<sup>+</sup> Figure VI-23d et e) ne permet donc pas d'exclure la possibilité d'une structure DKP.

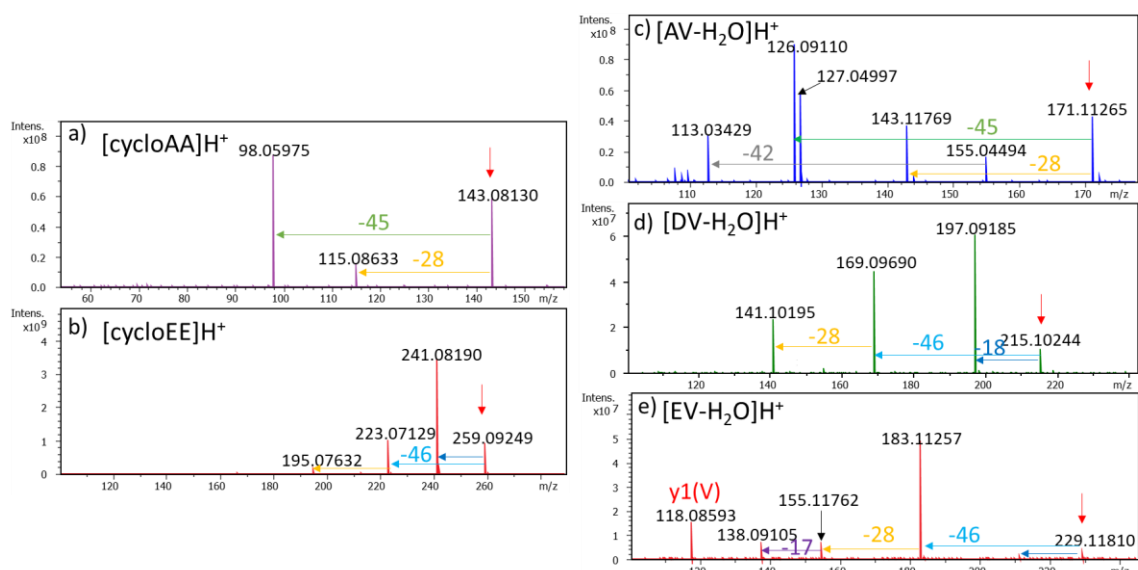


Figure VI-23. Spectres MS/MS des dipeptides déshydratés. Les spectres a et b sont des dipeptides cycliques de référence : a) [cycloAA]H<sup>+</sup> (Elab = 10 eV), b) dipeptide cyclique de référence [cycloEE]H<sup>+</sup> (Elab = 6 eV). Les spectres MS/MS c, d et e sont des dipeptides déshydratés obtenus dans les systèmes GAVD et GAVE activés à 160 °C : c) [AV-H<sub>2</sub>O]H<sup>+</sup> (Elab = 12 eV) (les fragments à 155 et 127 correspondent respectivement aux pertes de chaînes latérales de Ala et Val), d) [DV-H<sub>2</sub>O]H<sup>+</sup> (Elab = 6 eV) et e) [EV-H<sub>2</sub>O]H<sup>+</sup> (Elab = 10 V) (ESI positif, instrument FTICR).

Par ailleurs, une différence non négligeable est observée entre les voies de fragmentation des espèces contenant le résidu Glu et celles contenant le résidu Asp. En effet avec Asp, aucun fragment y ni b n'est observé. Seules des pertes de 18, 28 et 46 u (correspondant respectivement à une perte de H<sub>2</sub>O, CO et HCOOH) pouvant provenir de la chaîne latérale ou de la terminaison sont détectées (voir le spectre MS/MS de [DV-H<sub>2</sub>O]H<sup>+</sup> Figure VI-23 d), ne permettant pas de conclure quant à la structure, cyclique ou linéaire, du dipeptide déshydraté.

Néanmoins, l'espèce [DV-H<sub>2</sub>O]H<sup>+</sup> a été aussi étudiée par LC/MS/MS. Son analyse, reportée en Figure VI-24 montre trois formes dont deux (majoritaires) sont fragmentées sous CID. Le premier pic peut correspondre à une forme (succinimide Asp)-Val et le deuxième à une forme DKP, comme le suggèrent les pertes de neutres caractéristiques.

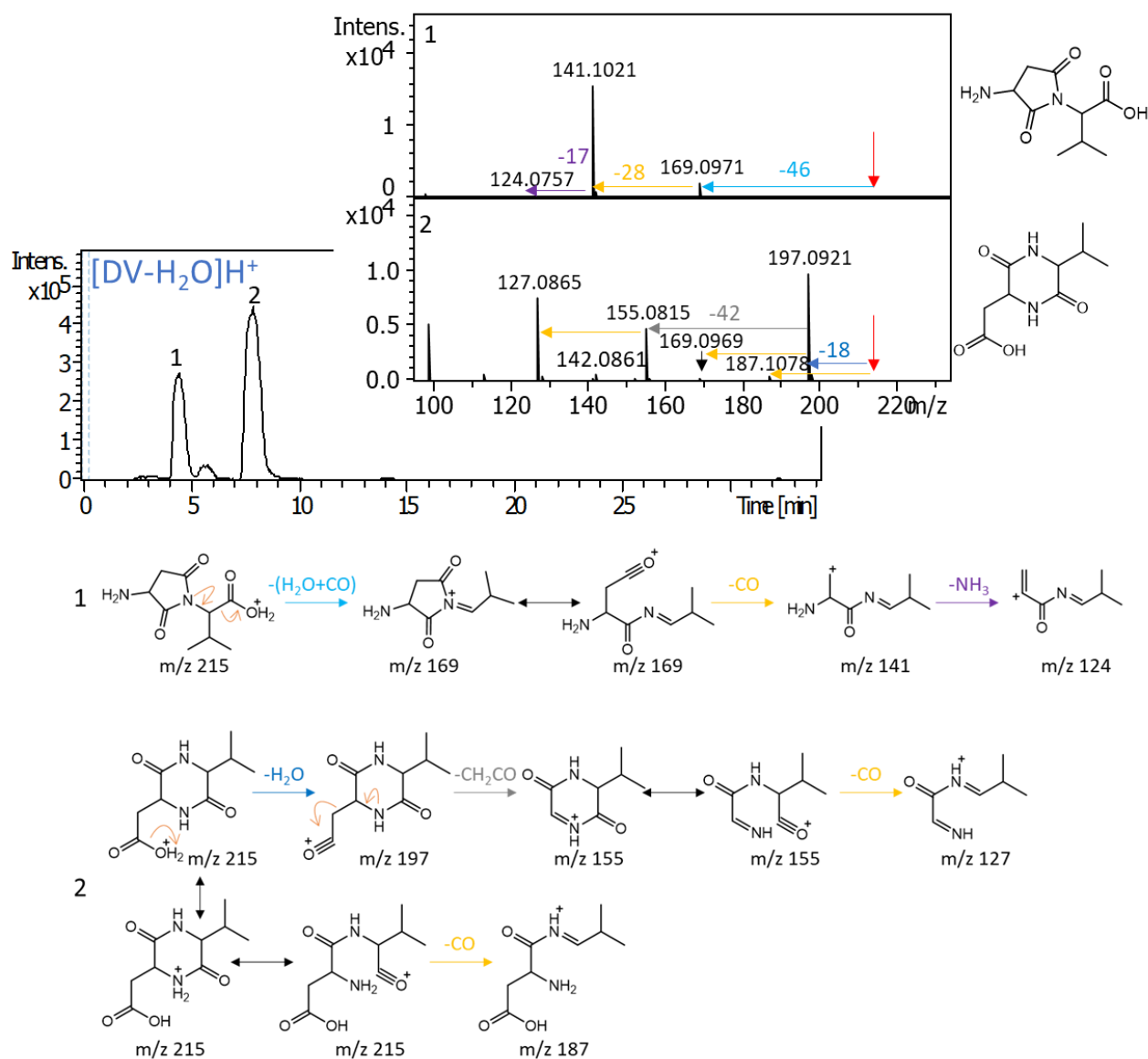


Figure VI-24. Chromatogramme extrait et spectres MS/MS de l'ion  $m/z$  215 correspondant au dipeptide  $[DV-H_2O]H^+$ , obtenus pour l'analyse LC/MS/MS de GAVD/SiO<sub>2</sub> activé à 160 °C (ESI positif, CID avec Elab = 23,5 eV, instrument TIMS-TOF). Des propositions de voies de fragmentation cohérentes avec les ions fragments observés sont données.

Pour les systèmes contenant Glu, on avait supposé jusqu'à présent que la forme déshydratée la plus favorable était pyroGlu. L'analyse LC/MS/MS de l'espèce  $[EL-H_2O]H^+$  de l'échantillon (L+E)/SiO<sub>2</sub> activé à 160 °C montre six pics chromatographiques, indiquant la présence de différentes structures de la forme déshydratée. L'EIC correspondant et les spectres MS/MS sont reportés en Figure VI-25.

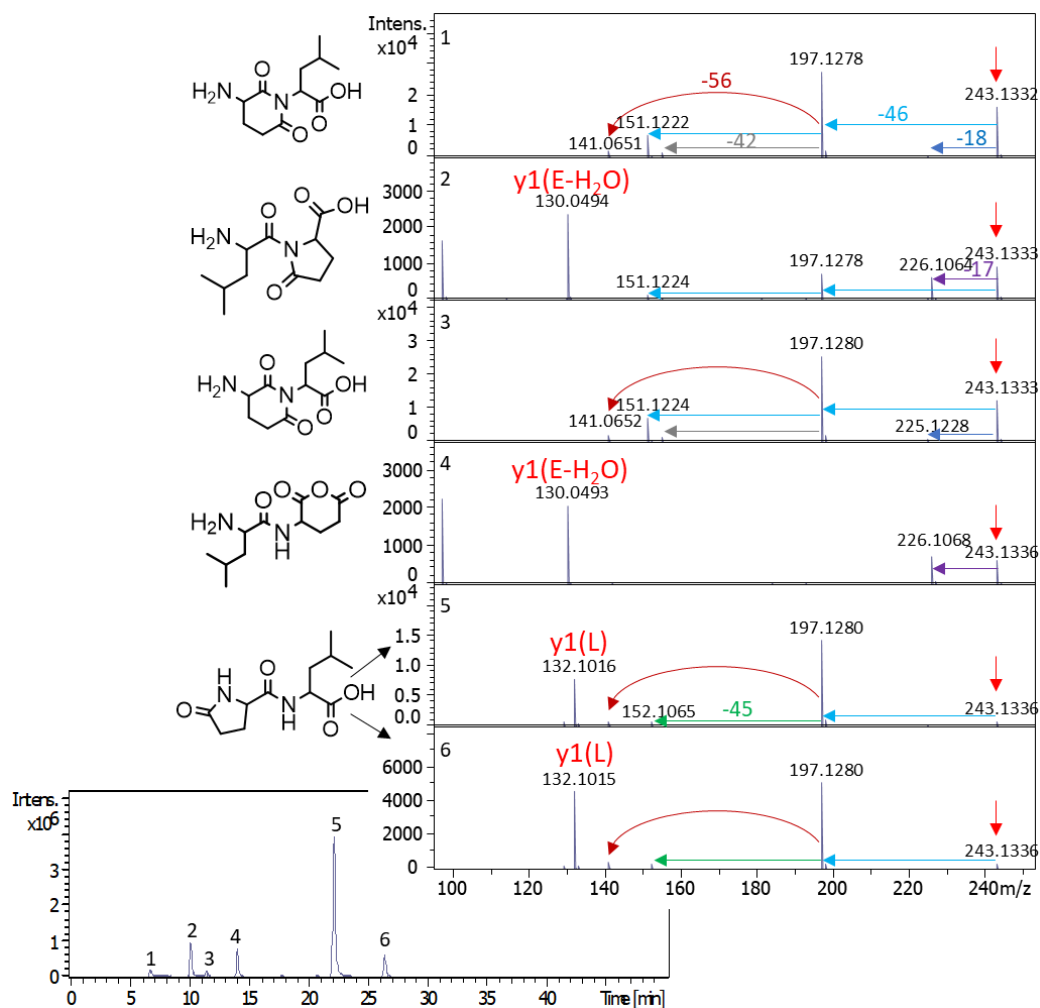


Figure VI-25. Chromatogramme extrait et spectres MS/MS de l'ion  $m/z$  243 correspondant au dipeptide  $[EL-H_2O]H^+$ , obtenus pour l'analyse LC/MS/MS de (L+E)/SiO<sub>2</sub> activé à 160 °C (ESI positif, CID Elab = 24,1 eV, instrument TIMS-TOF). La flèche courbe rouge correspond à une perte de C<sub>4</sub>H<sub>8</sub> (-56,06260 u).

Ainsi, d'après les voies de fragmentation observées sur les spectres MS/MS des pics 1 et 3, ces pics correspondent à une forme glutarimide. Les structures attribuables aux pics 2 et 4 sont illustrées sur la Figure VI-25. Dans ces deux cas le résidu Glu est localisé en position C-terminale, comme suggéré par la détection de l'ion fragment  $m/z$  130, soit  $y1(E-H_2O)$ . Au vu des intensités des pics, ces formes semblent être minoritaires (à condition de supposer une réponse MS analogue entre les différentes structures déshydratées). Concernant les pics chromatographiques 5 et 6 (majoritaires), ils semblent correspondre à une structure de type pyroEL, avec notamment la présence de l'ion  $y1(L)$  à  $m/z$  132 et les pertes consécutives de 46 u et 45 u. Ces voies de fragmentation sont illustrées sur la Figure VI-26.

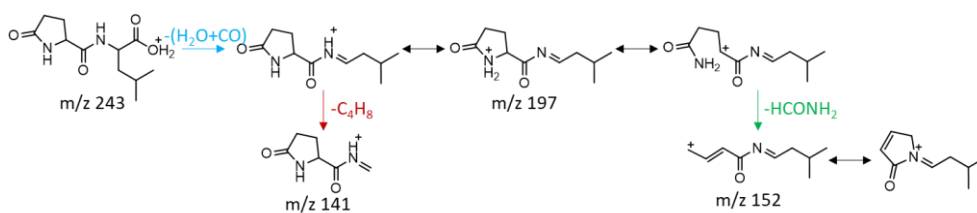


Figure VI-26. Illustration des voies de fragmentation de l'espèce pyroEL.

Les spectres MS/MS des pics chromatographiques 1 et 3, et 4 et 5, sont identiques deux à deux et seules des structures glutarimide et pyroGlu semblent correspondre aux fragmentations observées. On peut expliquer cela par la présence de stéréoisomères cis/trans. Il est possible de faire un parallèle entre ces stéréoisomères et ceux de la proline. En effet, l'isomérisation cis-trans de la proline est un phénomène bien connu, fréquemment observé dans les peptides et protéines<sup>279</sup>. Or, il a été montré dans la littérature qu'il est possible de séparer la proline cis de la proline trans par HPLC en phase inverse<sup>280,281</sup>.

Le dipeptide  $[EL-H_2O]H^+$ , a été également étudié par IM-MS/MS (pour les deux échantillons (L+E)/SiO<sub>2</sub> et (L+Q)/SiO<sub>2</sub>) (voir Figure VI-27). Les deux espèces majoritaires de chaque mobilogramme ont pu être fragmentées.

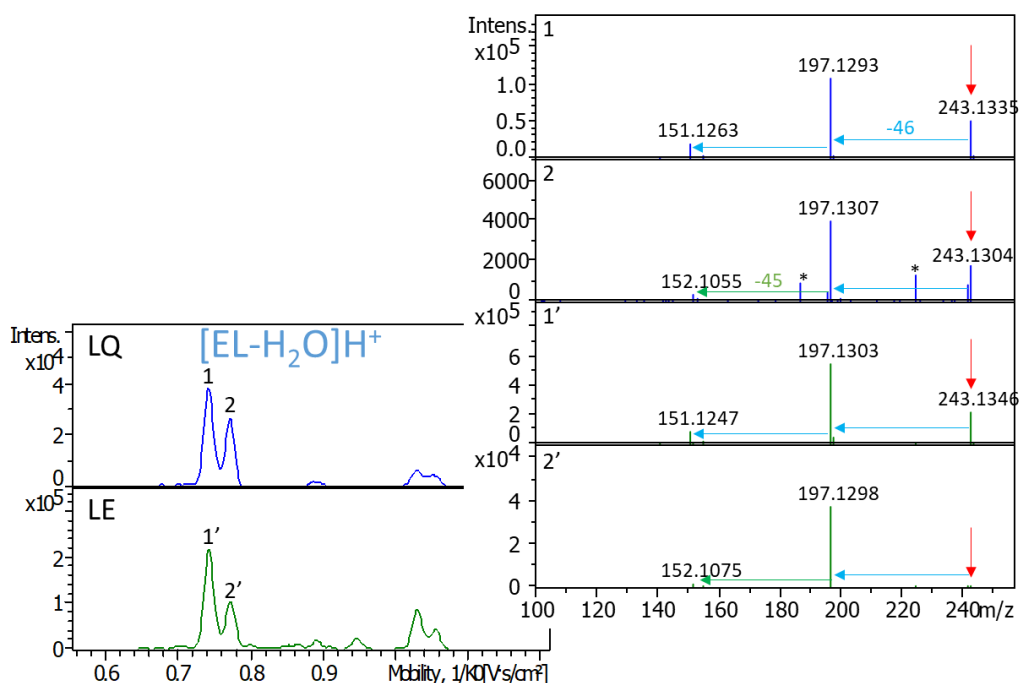


Figure VI-27. Mobilogrammes extraits et spectres MS/MS de l'ion  $m/z$  243 correspondant au dipeptide  $[EL-H_2O]H^+$ , obtenus pour les analyses IM-MS/MS de (L+Q)/SiO<sub>2</sub> et (L+E)/SiO<sub>2</sub> activés à 160 °C (ESI positif, CID Elab = 25 eV, instrument TIMS-TOF). Dans l'échantillon (L+Q)/SiO<sub>2</sub>, cette espèce peut aussi être annotée  $[QL-NH_3]H^+$ .

Plusieurs informations intéressantes peuvent être extraites de ces expériences. Tout d'abord, les mobilités observées des deux pics majoritaires sont identiques. Ensuite, les deux pics de mobilité d'un même système montrent des voies de fragmentation différentes, suggérant la séparation de deux structures distinctes pour un système donné. Enfin les voies de fragmentation des pics 1 et 1' et 2 et 2' sont similaires deux à deux. Or il convient de souligner que les deux systèmes ne peuvent pas former exactement les mêmes espèces. En effet, en considérant que les produits d'activation de l'échantillon (L+Q)/SiO<sub>2</sub> sont uniquement de stœchiométries  $\gamma$ ExLy (voir Chapitre V), les structures déshydratées correspondantes ne peuvent sans doute pas adopter de structures DKP (le cycle formé serait un cycle à 8, défavorable), ni faire une cyclisation interne de type pyroGlu. Par ailleurs, dans ce système, le résidu Glu semble uniquement se localiser en position N-terminale. Ainsi, les seules structures déshydratées envisageables pour une séquence  $\gamma$ EL sont les structures glutarimide et oxazolone (les dipeptides correspondant aux structures b' et d proposées Figure VI-8).

Les spectres MS/MS 1 et 1' de la Figure VI-27 semblent correspondre aux spectres MS/MS des pics 1 ou 3 de l'expérience LC/MS/MS de la Figure VI-25, et donc rendre compte d'une structure glutarimide dans les deux systèmes. Ici, les deux formes observées en LC/MS, que l'on a précédemment imputées à deux structures cis et trans, ne sont pas séparées.

En revanche il est plus difficile d'attribuer les pics 2 et 2'. En effet, bien qu'ils semblent correspondre aux spectres MS/MS des pics 5 ou 6 de l'expérience LC/MS/MS de la Figure VI-25, une forme de type pyroGlu, envisagée pour le système Leu + Glu n'est pas envisageable pour le système Leu + Gln (voir au-dessus). Une explication possible pour le pic 2 est alors une forme oxazolone, la perte de 45 u étant imputée à des pertes successives de 17 et 28 u. Quant au pic chromatographique 2', il peut soit correspondre à la structure pyroGlu-Leu (dans ce cas les deux formes auraient des mobilités proches), soit à une forme oxazolone également. En tout état de cause, l'absence des fragments  $\gamma$ 1 dans les analyses IM-MS/MS ne permet pas de se prononcer avec certitude sur la présence ou non d'une structure définie.

En conclusion, les oligopeptides déshydratés comportant un résidu Glu présentent plusieurs structures différentes. D'après les analyses LC/MS/MS la structure présentant pyroGlu en N-terminal est majoritaire. Mais d'autres formes telles que la forme glutarimide sont également présentes.



## IV.2. Généralisation aux oligopeptides déshydratés

Concernant les oligopeptides plus longs, les structures sont plus difficiles à démontrer car les ions fragments de séries b et y dominent les spectres MS/MS. Néanmoins, les observations expérimentales faites précédemment pour l'ion  $[EL_2-H_2O]H^+$  (voir Figure VI-10) sont cohérentes avec des structures déshydratées de formes pyroGlu et glutarimide, illustrées à la Figure VI-28, qui peuvent expliquer l'ensemble des pics détectés en LC/MS/MS pour cette stœchiométrie.

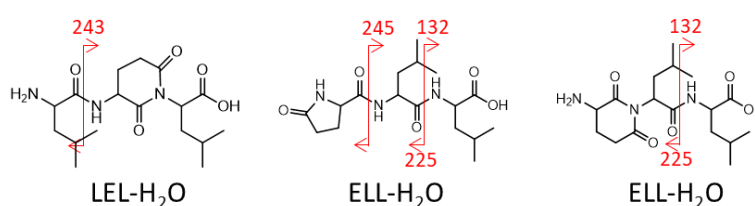


Figure VI-28. Structures correspondant aux séquences déterminées en Figure VI-10 pour les espèces  $[EL_2-H_2O]H^+$ .

De la même manière, les EIC des tripeptides  $[DVA-H_2O]H^+$  et  $[EVA-H_2O]H^+$  (Figure VI-29 et Figure VI-30) montrent la présence de plusieurs formes déshydratées pour chaque espèce. Les spectres MS/MS obtenus en infusion directe sur l'instrument FTICR démontrent un mélange de ces formes (voir Annexe 26). La caractérisation par IM-MS/MS du tripeptide  $[DVA-H_2O]H^+$  est reportée en Figure VI-29. Trois formes montrant des fragmentations différentes sont séparées. L'une semble avoir une structure DKP (perte de 45 u caractéristique d'un cycle DKP en compétition avec la perte de 46 u due à la chaîne latérale), et les deux autres des structures linéaires (présence d'ions fragments y et b), probablement les formes succinimide des séquences DAV et DVA. Ces trois formes sont compatibles avec les deux séquences déterminées pour l'espèce intacte correspondante (voir Figure VI-11), confirmant la sélectivité de séquence de la polymérisation. De même, les trois séquences attribuées au tripeptide déshydraté  $[DL_2-H_2O]H^+$  correspondent à celles déterminées pour  $[DL_2]H^+$  (voir Annexe 27). Pour le tripeptide  $[EVA-H_2O]H^+$ , seul le pic chromatographique majoritaire à 14 min a pu être fragmenté. Les ions fragments obtenus correspondent à une séquence majoritaire pE-A-V. Les autres formes sont des formes minoritaires, elles n'ont pas pu être sélectionnées lors de l'acquisition automatique des expériences LC/MS/MS en raison

de leurs trop faibles intensités. Comme pour  $[DVA-H_2O]H^+$ , la séquence déterminée est compatible avec la séquence majoritaire déterminée pour la forme intacte (voir Figure VI-11).

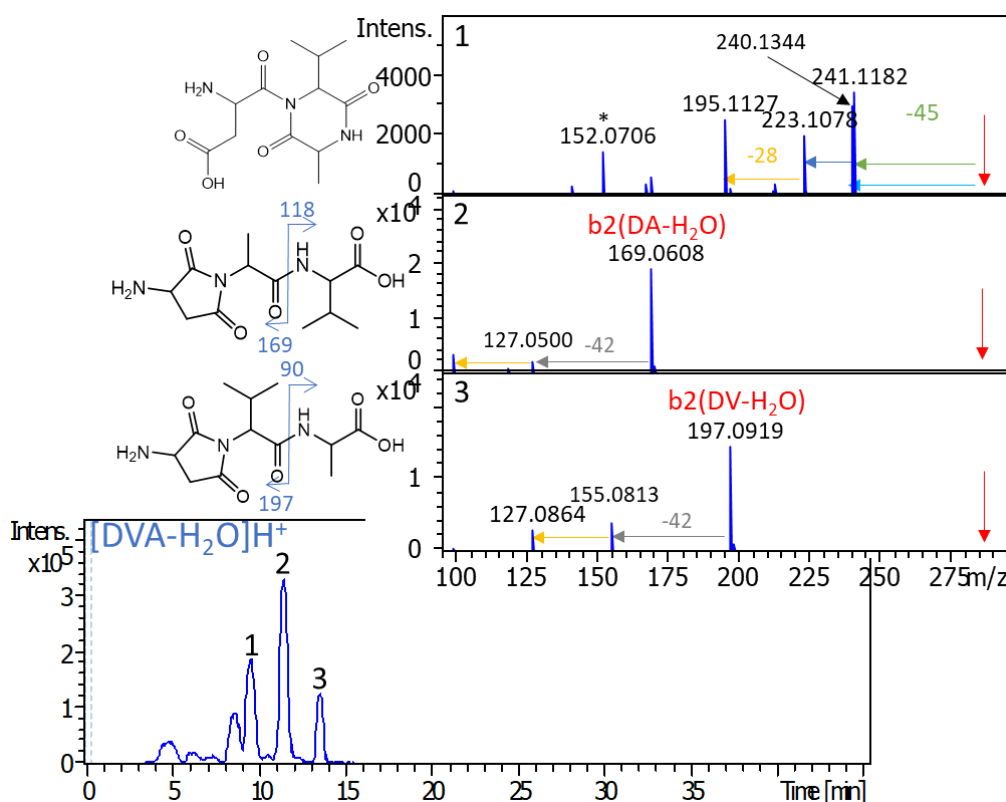


Figure VI-29. Chromatogramme extrait et spectres MS/MS de l'ion  $m/z$  286 correspondant au tripeptide  $[DVA-H_2O]H^+$ , obtenus pour l'analyse LC/MS/MS de GAVD/SiO<sub>2</sub> activé à 160 °C (ESI positif, CID Elab = 25 eV, instrument TIMS-TOF). Il faut noter que d'autres structures DKP peuvent être proposées.

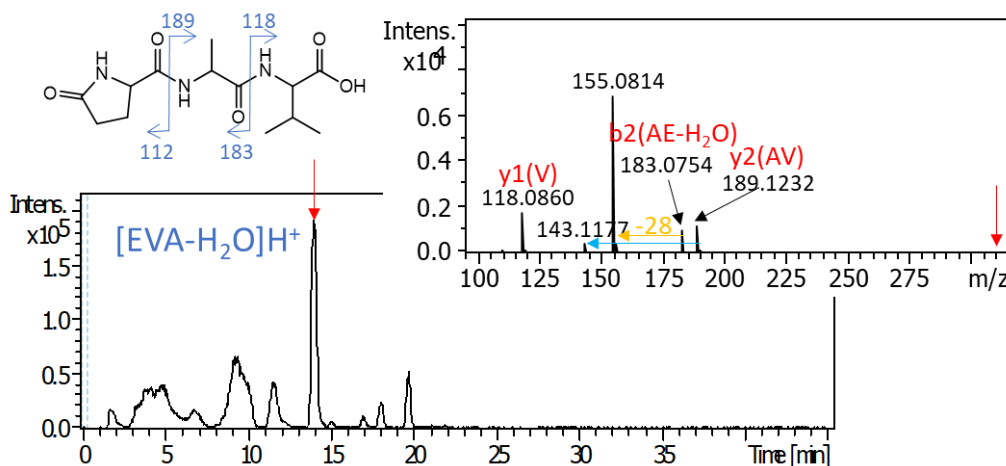


Figure VI-30. Chromatogramme extrait et spectre MS/MS de l'ion  $m/z$  300 correspondant au tripeptide  $[EVA-H_2O]H^+$ , obtenus pour l'analyse LC/MS/MS de GAVE/SiO<sub>2</sub> activé à 160 °C (ESI positif, CID Elab = 25,3 eV, instrument TIMS-TOF).

Dans le cas du système Leu +  $\beta$ Glu, les formes déshydratées sont très minoritaires, donc leurs séquences n'ont pas pu être déterminées. Dans le système Leu + Gln, les formes (QxLy-H<sub>2</sub>O) sont largement minoritaires ; les principales espèces détectées sont de stœchiométrie (ExLy-H<sub>2</sub>O) et montrent le même comportement que celles du système Leu + Glu.

En conclusion, les formes déshydratées des espèces contenant Glu semblent être majoritairement des formes de type pyroGlu, mais d'autres espèces de type glutarimide, DKP et oxazolone sont également présentes. Les espèces contenant Asp se déshydratent en formant des structures DKP ou succinimide. On peut alors poser sur de nouvelles bases la question initiale : les formes déshydratées sont-elles des impasses, ou peuvent-elles poursuivre la polymérisation ? Dans le cas des formes DKP, nous avons montré au Chapitre IV qu'elles en sont capables. C'est également le cas des formes pyroGlu (la polymérisation peut continuer par le côté C-terminal) et éventuellement des formes glutarimide et succinimide pour lesquelles les deux groupements  $\text{NH}_2$  et  $\text{COOH}$  restent disponibles.

Plus important, indépendamment de la structure des formes déshydratées, leur stœchiométrie et leurs séquences semblent correspondre à celles des formes intactes, confirmant que leur formation est une réaction secondaire à la polymérisation, et donc que la déshydratation conserve la sélectivité de séquence. Ainsi, on a pu observer dans plusieurs cas une localisation préférentielle de l'acide aminé trifonctionnel en position N-terminale. Les acides aminés aliphatiques sont donc rarement présents à cette position mais ne montrent pas de préférence marquée pour une autre position particulière, comme on a pu le constater dans l'étude des systèmes GAVD et GAVE. Un point plus étonnant concerne l'alternance des résidus Asp ou Glu : une polymérisation par blocs semble être privilégiée pour les oligopeptides avec Asp (sauf quand ils copolymérisent avec Leu) alors que les produits contenant Glu montrent une polymérisation alternée. Cette observation expérimentale a son importance car pour de plus grandes longueurs, les polymérisations par blocs ou alternées induiraient des structures secondaires différentes.

Concernant la régiosélectivité, diverses preuves expérimentales suggèrent la formation de peptides  $\beta$  et  $\gamma$ , à partir de acides aminés trifonctionnels Asp et Glu, respectivement. Néanmoins, on ne peut pas faire d'estimation de leurs quantités respectives, d'autant plus que les formes intactes sont minoritaires dans nos conditions d'activation. On peut dire en tout cas que notre scénario de polymérisation n'est pas régiosélectif, dans le sens où il ne sélectionne pas les liaisons  $\alpha$  typiques des protéines actuelles.

Si la correspondance entre les formes intactes et déshydratées est établie en termes de séquences, il est plus difficile de prouver que la déshydratation conserve aussi la régiosélectivité. En effet, on ignore si les formes  $\alpha$  et  $\gamma$  se déshydratent avec la même efficacité.



## Conclusion générale

Dans ce travail de thèse, des approches analytiques basées essentiellement sur la spectrométrie de masse à haute résolution ont été appliquées à l'analyse globale de mélanges de désorption issus de l'activation thermique d'acides aminés sur surface minérale, dans des conditions compatibles avec un scénario prébiotique de synthèse d'oligopeptides.

Une étude préliminaire s'est portée sur l'analyse d'un échantillon modèle, à savoir l'activation thermique d'un mélange d'acide glutamique et de leucine déposés sur une surface de silice. Des techniques d'analyse *in situ* ainsi que la spectrométrie de masse à haute résolution après désorption en solution des produits de polymérisation ont mis en évidence une polymérisation efficace pour une activation thermique à 160 °C.

Notre approche permet d'apporter une meilleure compréhension de la polymérisation des acides aminés sur un support minéral.

Tout d'abord, notre étude a démontré une certaine chimiosélectivité au niveau de la stœchiométrie des oligopeptides formés ; il apparaît notamment que la formation d'hétéropeptides est privilégiée par rapport à une polymérisation aléatoire.

Par la suite, nous avons entrepris une étude mécanistique dans l'optique d'étudier l'ensemble des étapes menant à la polymérisation des acides aminés sur les surfaces minérales. L'apparition successive d'oligopeptides de plus en plus longs a été observée pour des températures ou des temps de réaction croissants, suggérant des processus de polymérisation par étapes ; les oligopeptides peuvent être formés par condensation d'un résidu acide aminé sur un oligopeptide déjà présent et/ou par condensation de deux oligopeptides de taille intermédiaire.

Concernant d'autres aspects de la sélectivité de polymérisation, les conclusions sont plus mitigées : l'utilisation de la détection à haute résolution et de la spectrométrie de masse en tandem, en couplage avec des méthodes séparatives, ont permis de conclure que dans nos conditions de polymérisation, aucune régiosélectivité ou énantiosélectivité significatives ne sont présentes. Ces conclusions ont été confirmées par l'étude d'autres systèmes.

En effet, notre scénario de polymérisation a ensuite été étendu à différents autres mélanges binaires d'acides aminés, ainsi qu'à des systèmes plus complexes contenant quatre acides aminés (GAVD et GAVE), plus réalistes du point de vue prébiotique, ayant comme point

commun la présence d'un acide trifonctionnel. Une comparaison des systèmes complexes GAVD et GAVE a été réalisée en utilisant les analyses de Kendrick et de Van Krevelen. Dans tous les cas, une polymérisation donnant lieu à la production d'oligopeptides a été mise en évidence à condition qu'un acide aminé trifonctionnel soit présent dans le mélange. Néanmoins, les résultats obtenus pour les différents acides aminés trifonctionnels ont montré certaines variations dans la réactivité de ces résidus. Par exemple, une différence a été relevée au niveau de l'incorporation plus importante du résidu Asp par rapport au résidu Glu.

Une étude structurale des oligopeptides formés a été également menée par spectrométrie de masse en tandem, éventuellement couplée à la chromatographie liquide ou à la mobilité ionique. Elle a permis de mettre en évidence une sélectivité de séquence : dans la plupart des cas qui ont pu être résolus, une ou des séquences particulières sont favorisées, et en particulier il apparaît que le résidu trifonctionnel (en tout cas pour Asp et Glu) a tendance à se localiser en position N-terminale. Cette observation a été confirmée tant au niveau de la structure des formes déshydratées que des espèces intactes, démontrant que la sélectivité de séquence s'établit lors de la polymérisation initiale. Ces données permettent d'interroger plus précisément le rôle des acides aminés trifonctionnels, qui semble être essentiellement celui d'un initiateur de polymérisation en N-terminal. Cela expliquerait qu'ils soient efficaces même en faibles proportions.

Le mode de polymérisation diffère en fonction des acides aminés considérés : l'acide aminé Asp semble privilégier une polymérisation par bloc à l'inverse de Glu, ce qui permet d'envisager des structures secondaires différentes (hélices alpha et feuillet bêta) pour des oligopeptides de plus grandes longueurs.

En outre, la formation de régioisomères (des peptides  $\beta$  et  $\gamma$ ) a été démontrée à partir des acides aminés trifonctionnels Asp et Glu, respectivement, confirmant que dans nos conditions de polymérisation le scénario ne manifeste pas de régiosélectivité pour les liaisons  $\alpha$  prédominantes dans les protéines. On peut alors se demander comment cette sélectivité, presque exclusive, des protéines pour les liaisons  $\alpha$ , a pu s'instaurer ultérieurement.

La formation d'espèces déshydratées a été presque systématiquement observée pour tous nos systèmes. Leur nature et leur rôle dans la polymérisation posent question et ont été



examinés en détail. L'étude cinétique (en fonction du temps, ou de la température de réaction) a montré que les espèces déshydratées sont en majorité formées ultérieurement aux peptides intacts. Néanmoins, ces espèces peuvent aussi participer au processus de polymérisation, comme on l'a observé pour la réaction « DKP + monomère », bien que la polymérisation apparaisse moins efficace que pour la réaction « dipeptide linéaire + monomère ». La caractérisation avancée de ces formes déshydratées a montré une grande variété structurale. Ainsi des formes pyroGlu, mais également glutarimide, DKP et oxazolone sont observées pour les espèces contenant le résidu Glu. Les espèces contenant Asp se déshydratent quant à elles en formant des structures DKP ou succinimide.

En tout état de cause, la polymérisation d'acides aminés sur silice dans des conditions cohérentes avec les scénarios prébiotiques conduit à la formation d'oligopeptides non aléatoires et présentant des caractéristiques intéressantes pour l'émergence de la Vie sur Terre. Quels que soient les systèmes étudiés, des oligopeptides de taille modérée (parfois jusqu'aux octapeptides) ont été produits dans nos conditions d'activation thermique en une seule étape. Ces conclusions obtenues sur une surface de silice ne sont pas généralisables à d'autres supports minéraux (alumine et montmorillonite), attestant de la complexité du monde minéral et de son interaction avec les premières molécules du vivant.

La question de la formation d'oligopeptides plus longs, permettant de se rapprocher des biomolécules du vivant, se pose. Des voies d'amélioration sont envisageables, par exemple une polymérisation à plus faible température sur des échelles de temps plus importantes. La direction sans doute la plus prometteuse est l'application de cycles hydratation/déshydratation, mimant les variations environnementales.

## Références

- (1) Arrhenius, S. Die Verbreitung Des Lebens Im Weltenraum. *Umschau* **1903**, *7*, 481–485.
- (2) Kvenvolden, K.; Lawless, J.; Pering, K.; Peterson, E.; Flores, J.; Ponnampereuma, C.; Kaplan, I. R.; Moore, C. Evidence for Extraterrestrial Amino-Acids and Hydrocarbons in the Murchison Meteorite. *Nature* **1970**, *228* (5275), 923–926. <https://doi.org/10.1038/228923a0>.
- (3) Stoks, P. G.; Schwartz, A. W. Uracil in Carbonaceous Meteorites. *Nature* **1979**, *282* (5740), 709–710. <https://doi.org/10.1038/282709a0>.
- (4) Whittet, D. C. B. INTERSTELLAR ICES STUDIED WITH THE INFRARED SPACE OBSERVATORY. *Orig. Life Evol. Biosph.* **1997**, *27* (1), 101–113. <https://doi.org/10.1023/A:1006570113567>.
- (5) Miller, S. L. A Production of Amino Acids under Possible Primitive Earth Conditions. *Sci. New Ser.* **1953**, *117* (3046), 528–529.
- (6) Parker, E. T.; Cleaves, H. J.; Dworkin, J. P.; Glavin, D. P.; Callahan, M.; Aubrey, A.; Lazcano, A.; Bada, J. L. Primordial Synthesis of Amines and Amino Acids in a 1958 Miller H<sub>2</sub>S-Rich Spark Discharge Experiment. *Proc. Natl. Acad. Sci.* **2011**, *108* (14), 5526–5531. <https://doi.org/10.1073/pnas.1019191108>.
- (7) Forsythe, J. G.; Yu, S.-S.; Mamajanov, I.; Grover, M. A.; Krishnamurthy, R.; Fernández, F. M.; Hud, N. V. Ester-Mediated Amide Bond Formation Driven by Wet-Dry Cycles: A Possible Path to Polypeptides on the Prebiotic Earth. *Angew. Chem. Int. Ed.* **2015**, *54* (34), 9871–9875. <https://doi.org/10.1002/anie.201503792>.
- (8) Bernstein, M. P.; Dworkin, J. P.; Sandford, S. A.; Cooper, G. W.; Allamandola, L. J. Racemic Amino Acids from the Ultraviolet Photolysis of Interstellar Ice Analogues. *Nature* **2002**, *416* (6879), 401–403. <https://doi.org/10.1038/416401a>.
- (9) Peltzer, E. T.; Bada, J. L.  $\alpha$ -Hydroxycarboxylic Acids in the Murchison Meteorite. *Nature* **1978**, *272* (5652), 443–444. <https://doi.org/10.1038/272443a0>.
- (10) Bada, J. L. New Insights into Prebiotic Chemistry from Stanley Miller's Spark Discharge Experiments. *Chem. Soc. Rev.* **2013**, *42* (5), 2186. <https://doi.org/10.1039/c3cs35433d>.
- (11) Oró, J. Synthesis of Adenine from Ammonium Cyanide. *Biochem. Biophys. Res. Commun.* **1960**, *2* (6), 407–412. [https://doi.org/10.1016/0006-291X\(60\)90138-8](https://doi.org/10.1016/0006-291X(60)90138-8).
- (12) Sutherland, J. D.; Whitfield, J. N. Studies on a Potentially Prebiotic Synthesis of RNA. *Tetrahedron* **1997**, *53* (34), 11595–11626. [https://doi.org/10.1016/S0040-4020\(97\)00466-3](https://doi.org/10.1016/S0040-4020(97)00466-3).
- (13) Powner, M. W.; Gerland, B.; Sutherland, J. D. Synthesis of Activated Pyrimidine Ribonucleotides in Prebiotically Plausible Conditions. *Nature* **2009**, *459* (7244), 239–242. <https://doi.org/10.1038/nature08013>.
- (14) Fishkis, M. Steps Towards the Formation of A Protocell: The Possible Role of Short Peptides. *Orig. Life Evol. Biospheres* **2007**, *37* (6), 537–553. <https://doi.org/10.1007/s11084-007-9111-4>.
- (15) Gilbert, W. Origin of Life: The RNA World. *Nature* **1986**, *319* (6055), 618–618. <https://doi.org/10.1038/319618a0>.
- (16) Schwartz, A. W.; de Graaf, R. M. The Prebiotic Synthesis of Carbohydrates: A Reassessment. *J. Mol. Evol.* **1993**, *36* (2), 101–106. <https://doi.org/10.1007/BF00166245>.
- (17) Gorlero, M.; Wieczorek, R.; Adamala, K.; Giorgi, A.; Schininà, M. E.; Stano, P.; Luisi, P. L. Ser-His Catalyses the Formation of Peptides and PNAs. *FEBS Lett.* **2009**, *583* (1), 153–156. <https://doi.org/10.1016/j.febslet.2008.11.052>.
- (18) van der Gulik, P.; Speijer, D. How Amino Acids and Peptides Shaped the RNA World. *Life* **2015**, *5* (1), 230–246. <https://doi.org/10.3390/life5010230>.
- (19) Brack, A. From Interstellar Amino Acids to Prebiotic Catalytic Peptides: A Review. *Chem. Biodivers.* **2007**, *4* (4), 665–679. <https://doi.org/10.1002/cbdv.200790057>.
- (20) Bailey, J. Chirality and the Origin of Life. *Acta Astronaut.* **2000**, *46* (10), 627–631. [https://doi.org/10.1016/S0094-5765\(00\)00024-2](https://doi.org/10.1016/S0094-5765(00)00024-2).
- (21) Blackmond, D. G. The Origin of Biological Homochirality. *Cold Spring Harb. Perspect. Biol.* **2019**, *11* (3), a032540. <https://doi.org/10.1101/cshperspect.a032540>.
- (22) Meierhenrich, U. Stereochemistry for the Study of the Origin of Life. In *Amino Acids and the Asymmetry of Life: Caught in the Act of Formation*; Meierhenrich, U., Ed.; Advances in Astrobiology and Biogeophysics; Springer: Berlin, Heidelberg, 2008; pp 17–46. [https://doi.org/10.1007/978-3-540-76886-9\\_2](https://doi.org/10.1007/978-3-540-76886-9_2).
- (23) de Marcellus, P.; Meinert, C.; Nuevo, M.; Filippi, J.-J.; Danger, G.; Deboffle, D.; Nahon, L.; Le Sergeant d'Hendecourt, L.; Meierhenrich, U. J. NON-RACEMIC AMINO ACID PRODUCTION BY ULTRAVIOLET IRRADIATION OF ACHIRAL INTERSTELLAR ICE ANALOGS WITH CIRCULARLY POLARIZED LIGHT. *Astrophys. J.* **2011**, *727* (2), L27. <https://doi.org/10.1088/2041-8205/727/2/L27>.

- (24) Hazen, R. M.; Filley, T. R.; Goodfriend, G. A. Selective Adsorption of L- and D-Amino Acids on Calcite: Implications for Biochemical Homochirality. *Proc. Natl. Acad. Sci.* **2001**, *98* (10), 5487–5490. <https://doi.org/10.1073/pnas.101085998>.
- (25) Avetisov, V. A.; Kuz'min, V. V.; Goldanskii, V. I. Handedness, Origin of Life and Evolution. *Phys. Today* **2008**, *44* (7), 33. <https://doi.org/10.1063/1.881264>.
- (26) Joyce, G. F.; Visser, G. M.; van Boeckel, C. A. A.; van Boom, J. H.; Orgel, L. E.; van Westrenen, J. Chiral Selection in Poly(C)-Directed Synthesis of Oligo(G). *Nature* **1984**, *310* (5978), 602–604. <https://doi.org/10.1038/310602a0>.
- (27) Goldanskii, V. I.; Avetisov, V. A.; Kuz'min, V. V. Chiral Purity of Nucleotides as a Necessary Condition of Complementarity. *FEBS Lett.* **1986**, *207* (1), 181–183. [https://doi.org/10.1016/0014-5793\(86\)80036-9](https://doi.org/10.1016/0014-5793(86)80036-9).
- (28) Lambert, J.-F. Adsorption and Polymerization of Amino Acids on Mineral Surfaces: A Review. *Orig. Life Evol. Biospheres* **2008**, *38* (3), 211–242. <https://doi.org/10.1007/s11084-008-9128-3>.
- (29) Shock, E. L. Stability of Peptides in High-Temperature Aqueous Solutions. *Geochim. Cosmochim. Acta* **1992**, *56* (9), 3481–3491. [https://doi.org/10.1016/0016-7037\(92\)90392-V](https://doi.org/10.1016/0016-7037(92)90392-V).
- (30) Ferris, J. P. Catalysis and Prebiotic Synthesis. *Rev. Mineral. Geochem.* **2005**, *59* (1), 187–210. <https://doi.org/10.2138/rmg.2005.59.8>.
- (31) Cleaves, H. J.; Aubrey, A. D.; Bada, J. L. An Evaluation of the Critical Parameters for Abiotic Peptide Synthesis in Submarine Hydrothermal Systems. *Orig. Life Evol. Biosphere J. Int. Soc. Study Orig. Life* **2009**, *39* (2), 109–126. <https://doi.org/10.1007/s11084-008-9154-1>.
- (32) Kawamura, K.; Konagaya, N.; Maruoka, Y. Enhancement and Inhibitory Activities of Minerals for Alanine Oligopeptide Elongation Under Hydrothermal Conditions. *Astrobiology* **2018**, *18* (11), 1403–1413. <https://doi.org/10.1089/ast.2017.1732>.
- (33) Bernal, J. D. *The Physical Basis of Life*; Routledge and Kegan Paul: Londres, 1951.
- (34) Paecht-Horowitz, M.; Katchalsky, A. Synthesis of Amino Acyl-Adenylates under Prebiotic Conditions. *J. Mol. Evol.* **1973**, *2* (2–3), 91–98. <https://doi.org/10.1007/BF01653989>.
- (35) Paecht-Horowitz, M. Clays and Other Minerals in Prebiotic Processes. *Orig. Life* **1984**, *14* (1–4), 307–314. <https://doi.org/10.1007/BF00933672>.
- (36) Wächtershäuser, G. Pyrite Formation, the First Energy Source for Life: A Hypothesis. *Syst. Appl. Microbiol.* **1988**, *10* (3), 207–210. [https://doi.org/10.1016/S0723-2020\(88\)80001-8](https://doi.org/10.1016/S0723-2020(88)80001-8).
- (37) Pascal, R.; Boiteau, L.; Commeyras, A. From the Prebiotic Synthesis of  $\alpha$ -Amino Acids Towards a Primitive Translation Apparatus for the Synthesis of Peptides. In *Prebiotic Chemistry*; Walde, P., Ed.; Topics in Current Chemistry; Springer: Berlin, Heidelberg, 2005; pp 69–122. <https://doi.org/10.1007/b136707>.
- (38) Danger, G.; Boiteau, L.; Cottet, H.; Pascal, R. The Peptide Formation Mediated by Cyanate Revisited. N-Carboxyanhydrides as Accessible Intermediates in the Decomposition of N-Carbamoylamino Acids. *J. Am. Chem. Soc.* **2006**, *128* (23), 7412–7413. <https://doi.org/10.1021/ja061339+>.
- (39) Gao, X.; Liu, Y.; Xu, P. X.; Cai, Y. M.; Zhao, Y. F.  $\alpha$ -Amino Acid Behaves Differently from  $\beta$ - or  $\gamma$ -Amino Acids as Treated by Trimetaphosphate. *Amino Acids* **2008**, *34* (1), 47–53. <https://doi.org/10.1007/s00726-007-0599-8>.
- (40) Marshall-Bowman, K.; Ohara, S.; Sverjensky, D. A.; Hazen, R. M.; Cleaves, H. J. Catalytic Peptide Hydrolysis by Mineral Surface: Implications for Prebiotic Chemistry. *Geochim. Cosmochim. Acta* **2010**, *74* (20), 5852–5861. <https://doi.org/10.1016/j.gca.2010.07.009>.
- (41) Fox, S.; Dalai, P.; Lambert, J.-F.; Strasdeit, H. Hypercondensation of an Amino Acid: Synthesis and Characterization of a Black Glycine Polymer. *Chem. – Eur. J.* **2015**, *21* (24), 8897–8904. <https://doi.org/10.1002/chem.201500820>.
- (42) Lambert, J.-F. Origins of Life: From the Mineral to the Biochemical World. *BIO Web Conf.* **2015**, *4*, 00012. <https://doi.org/10.1051/bioconf/20150400012>.
- (43) Lahav, N.; White, D.; Chang, S. Peptide Formation in the Prebiotic Era: Thermal Condensation of Glycine in Fluctuating Clay Environments. *Science* **1978**, *201* (4350), 67–69. <https://doi.org/10.1126/science.663639>.
- (44) White, D. H.; Erickson, J. C. Catalysis of Peptide Bond Formation by Histidyl-Histidine in a Fluctuating Clay Environment. *J. Mol. Evol.* **1980**, *16* (3–4), 279–290. <https://doi.org/10.1007/BF01804979>.
- (45) Bujdák, J.; Faybíková, K.; Eder, A.; Yongyai, Y.; Rode, B. M. Peptide Chain Elongation: A Possible Role of Montmorillonite in Prebiotic Synthesis of Protein Precursors. *Orig. Life Evol. Biosph.* **1995**, *25* (5), 431–441. <https://doi.org/10.1007/BF01581994>.

- (46) Ervin, J. N.; Bouza, M.; Fernández, F. M.; Forsythe, J. G. Proline Behavior in Model Prebiotic Peptides Formed by Wet–Dry Cycling. *ACS Earth Space Chem.* **2020**, *4* (8), 1349–1359. <https://doi.org/10.1021/acsearthspacechem.0c00119>.
- (47) Ferris, J. P. Montmorillonite Catalysis of 30–50 Mer Oligonucleotides: Laboratory Demonstration of Potential Steps in the Origin of the RNA World. *Orig. Life Evol. Biosph.* **2002**, *32* (4), 311–332. <https://doi.org/10.1023/A:1020543312109>.
- (48) Ferris, J. P.; Hill Jr, A. R.; Liu, R.; Orgel, L. E. Synthesis of Long Prebiotic Oligomers on Mineral Surfaces. *Lett. Nat.* **1996**, *381*, 59–61.
- (49) Zagorevskii, D. V.; Aldersley, M. F.; Ferris, J. P. MALDI Analysis of Oligonucleotides Directly from Montmorillonite. *J. Am. Soc. Mass Spectrom.* **2006**, *17* (9), 1265–1270. <https://doi.org/10.1021/jasms.8b02737>.
- (50) Mungi, C. V.; Bapat, N. V.; Hongo, Y.; Rajamani, S. Formation of Abasic Oligomers in Nonenzymatic Polymerization of Canonical Nucleotides. *Life Basel Switz.* **2019**, *9* (3). <https://doi.org/10.3390/life9030057>.
- (51) Fox, S. W.; Harada, K. Thermal Copolymerization of Amino Acids in the Presence of Phosphoric Acid. *Arch. Biochem. Biophys.* **1960**, *86*, 281–285.
- (52) Fox, S. W.; Harada, K. The Thermal Copolymerization of Amino Acids Common to Protein <sup>1</sup>. *J. Am. Chem. Soc.* **1960**, *82* (14), 3745–3751. <https://doi.org/10.1021/ja01499a069>.
- (53) Leyton, P.; Zárate, R. A.; Fuentes, S.; Paipa, C.; Gómez-Jeria, J. S.; Leyton, Y. Influence of Aluminum Oxide on the Prebiotic Thermal Synthesis of Gly-Glu-(Gly-Glu)<sub>n</sub> Polymer. *Biosystems* **2011**, *104* (2–3), 118–126. <https://doi.org/10.1016/j.biosystems.2011.01.008>.
- (54) Leyton, P.; Saladino, R.; Crestini, C.; Campos-Vallette, M.; Paipa, C.; Berríos, A.; Fuentes, S.; Zárate, R. A. Influence of TiO<sub>2</sub> on Prebiotic Thermal Synthesis of the Gly-Gln Polymer. *Amino Acids* **2012**, *42* (6), 2079–2088. <https://doi.org/10.1007/s00726-011-0939-6>.
- (55) Nakashima, T.; Jungck, J. R.; Fox, S. W.; Lederer, E.; Das, B. C. A Test for Randomness in Peptides Isolated from a Thermal Polyamino Acid. *Int. J. Quantum Chem.* **1977**, *12* (S4), 65–72. <https://doi.org/10.1002/qua.560120708>.
- (56) Hartmann, J.; Christel Brand, M.; Dose, K. Formation of Specific Amino Acid Sequences during Thermal Polymerization of Amino Acids. *Biosystems* **1981**, *13* (3), 141–147. [https://doi.org/10.1016/0303-2647\(81\)90055-1](https://doi.org/10.1016/0303-2647(81)90055-1).
- (57) Mosqueira, F. G.; Ramos-Bernal, S.; Negrón-Mendoza, A. A Simple Model of the Thermal Prebiotic Oligomerization of Amino Acids. *Biosystems* **2000**, *57* (2), 67–73. [https://doi.org/10.1016/S0303-2647\(00\)00089-7](https://doi.org/10.1016/S0303-2647(00)00089-7).
- (58) Bujdák, J.; Rode, B. M. Preferential Amino Acid Sequences in Alumina-Catalyzed Peptide Bond Formation. *J. Inorg. Biochem.* **2002**, *90* (1–2), 1–7. [https://doi.org/10.1016/S0162-0134\(02\)00395-1](https://doi.org/10.1016/S0162-0134(02)00395-1).
- (59) Hazen, R. M. Paleomineralogy of the Hadean Eon: A Preliminary Species List. *Am. J. Sci.* **2013**, *313* (9), 807–843. <https://doi.org/10.2475/09.2013.01>.
- (60) Morrison, S.; Runyon, S.; Hazen, R. The Paleomineralogy of the Hadean Eon Revisited. *Life* **2018**, *8* (4), 64. <https://doi.org/10.3390/life8040064>.
- (61) Legrand, A. P. *The Surface Properties of Silicas*; John Wiley & Sons: Chichester, 1998.
- (62) Zhuravlev, L. T. The Surface Chemistry of Amorphous Silica. Zhuravlev Model. *Colloids Surf. Physicochem. Eng. Asp.* **2000**, *173* (1–3), 1–38. [https://doi.org/10.1016/S0927-7757\(00\)00556-2](https://doi.org/10.1016/S0927-7757(00)00556-2).
- (63) Paecht-Horowitz, M.; Berger, J.; Katchalsky, A. Prebiotic Synthesis of Polypeptides by Heterogeneous Polycondensation of Amino-Acid Adenylates. *Nature* **1970**, *228* (5272), 636–639. <https://doi.org/10.1038/228636a0>.
- (64) Paecht-Horowitz, M.; Lahav, N. Polymerization of Alanine in the Presence of a Non-Swelling Montmorillonite. *J. Mol. Evol.* **1977**, *10* (1), 73–76. <https://doi.org/10.1007/BF01796135>.
- (65) Jackson, T. A. Preferential Polymerization and Adsorption of L-Optical Isomers of Amino Acids Relative to D-Optical Isomers on Kaolinite Templates. *Chem. Geol.* **1971**, *7* (4), 295–306. [https://doi.org/10.1016/0009-2541\(71\)90014-3](https://doi.org/10.1016/0009-2541(71)90014-3).
- (66) Basiuk, V. A.; Gromovoy, T. Y.; Glukhoy, A. M.; Golovaty, V. G. Chemical Transformations of Proteinogenic Amino Acids during Their Sublimation in the Presence of Silica. *Orig. Life Evol. Biosph.* **1991**, *21* (3), 129–144. <https://doi.org/10.1007/BF01809442>.
- (67) Basiuk, V. A.; Gromovoy, T. Yu.; Golovaty, V. G.; Glukhoy, A. M. Mechanisms of Amino Acid Polycondensation on Silica and Alumina Surfaces. *Orig. Life Evol. Biosph.* **1990**, *20* (6), 483–498. <https://doi.org/10.1007/BF01808196>.

- (68) Gromovoy TYu, null; Basiuk, V. A.; Chuiko, A. A. Growth of Peptide Chains on Silica in Absence of Amino Acid Access from Without. *Orig. Life Evol. Biosphere J. Int. Soc. Study Orig. Life* **1991**, *21* (3), 119–128. <https://doi.org/10.1007/BF01809441>.
- (69) Basiuk, V. A.; Gromovoy, T. Y.; Chuiko, A. A.; Soloshonok, V. A.; Kukhar, V. P. A Novel Approach to the Synthesis of Symmetric Optically Active 2,5-Dioxopiperazines. *Synthesis* **1992**, *1992* (5), 449–451. <https://doi.org/10.1055/s-1992-26131>.
- (70) Basiuk, V. A.; Gromovoy, T. Yu. Reactions of Vaporous Proteinogenic  $\alpha$ -Amino Acids on Silica and Alumina Surfaces. *React. Kinet. Catal. Lett.* **1993**, *50* (1–2), 297–303. <https://doi.org/10.1007/BF02062227>.
- (71) Basiuk, V. A.; Gromovoy, T. Y. The Gas-Solid-Phase 2,5-Dioxopiperazine Synthesis. Cyclization of Vaporous Dipeptides on Silica Surface. *Collect. Czechoslov. Chem. Commun.* **1994**, *59* (2), 461–466. <https://doi.org/10.1135/cccc19940461>.
- (72) Basiuk, V. A.; Navarro-Gonzalez, R.; Basiuk, E. V. Behavior of Amino Acids When Volatilized in the Presence of Silica Gel and Pulverized Basaltic Lava. *Orig. Life Evol. Biosph.* **1998**, *28* (2), 167–193. <https://doi.org/10.1023/A:1006596111124>.
- (73) Bujdák, J.; Rode, B. M. Silica, Alumina, and Clay-Catalyzed Alanine Peptide Bond Formation. *J. Mol. Evol.* **1997**, *45* (5), 457–466. <https://doi.org/10.1007/PL00006250>.
- (74) Bujdák, J.; Rode, B. M. Glycine Oligomerization on Silica and Alumina. *React. Kinet. Catal. Lett.* **1997**, *62* (2), 281–286. <https://doi.org/10.1007/BF02475464>.
- (75) Le Son, H.; Suwannachot, Y.; Bujdak, J.; Rode, B. M. Salt-Induced Peptide Formation from Amino Acids in the Presence of Clays and Related Catalysts. *Inorganica Chim. Acta* **1998**, *272* (1), 89–94. [https://doi.org/10.1016/S0020-1693\(97\)05975-6](https://doi.org/10.1016/S0020-1693(97)05975-6).
- (76) Meng, M.; Stievano, L.; Lambert, J.-F. Adsorption and Thermal Condensation Mechanisms of Amino Acids on Oxide Supports. 1. Glycine on Silica. *Langmuir* **2004**, *20* (3), 914–923. <https://doi.org/10.1021/la035336b>.
- (77) Stievano, L.; Yu Piao, L.; Lopes, I.; Meng, M.; Costa, D.; Lambert, J.-F. Glycine and Lysine Adsorption and Reactivity on the Surface of Amorphous Silica. *Eur. J. Mineral.* **2007**, 321–331. <https://doi.org/10.1127/0935-1221/2007/0019-1731>.
- (78) Lambert, J.-F.; Stievano, L.; Lopes, I.; Gharsallah, M.; Piao, L. The Fate of Amino Acids Adsorbed on Mineral Matter. *Planet. Space Sci.* **2009**, *57* (4), 460–467. <https://doi.org/10.1016/j.pss.2009.01.012>.
- (79) Lopes, I.; Piao, L.; Stievano, L.; Lambert, J.-F. Adsorption of Amino Acids on Oxide Supports: A Solid-State NMR Study of Glycine Adsorption on Silica and Alumina. *J. Phys. Chem. C* **2009**, *113* (42), 18163–18172. <https://doi.org/10.1021/jp906891y>.
- (80) Bouchoucha, M.; Jaber, M.; Onfroy, T.; Lambert, J.-F.; Xue, B. Glutamic Acid Adsorption and Transformations on Silica. *J. Phys. Chem. C* **2011**, *115* (44), 21813–21825. <https://doi.org/10.1021/jp206967b>.
- (81) Folliet, N.; Gervais, C.; Costa, D.; Laurent, G.; Babonneau, F.; Stievano, L.; Lambert, J.-F.; Tielens, F. A Molecular Picture of the Adsorption of Glycine in Mesoporous Silica through NMR Experiments Combined with DFT-D Calculations. *J. Phys. Chem. C* **2013**, *117* (8), 4104–4114. <https://doi.org/10.1021/jp312195a>.
- (82) Guo, C.; Holland, G. P. Investigating Lysine Adsorption on Fumed Silica Nanoparticles. *J. Phys. Chem. C* **2014**, *118* (44), 25792–25801. <https://doi.org/10.1021/jp508627h>.
- (83) Guo, C.; Holland, G. P. Alanine Adsorption and Thermal Condensation at the Interface of Fumed Silica Nanoparticles: A Solid-State NMR Investigation. *J. Phys. Chem. C* **2015**, *119* (45), 25663–25672. <https://doi.org/10.1021/acs.jpcc.5b10236>.
- (84) Guo, C.; Jordan, J. S.; Yarger, J. L.; Holland, G. P. Highly Efficient Fumed Silica Nanoparticles for Peptide Bond Formation: Converting Alanine to Alanine Anhydride. *ACS Appl. Mater. Interfaces* **2017**, *9* (20), 17653–17661. <https://doi.org/10.1021/acsami.7b04887>.
- (85) Ben Shir, I.; Kababya, S.; Amitay-Rosen, T.; Balazs, Y. S.; Schmidt, A. Molecular Level Characterization of the Inorganic–Bioorganic Interface by Solid State NMR: Alanine on a Silica Surface, a Case Study. *J. Phys. Chem. B* **2010**, *114* (18), 5989–5996. <https://doi.org/10.1021/jp100114v>.
- (86) Ben Shir, I.; Kababya, S.; Schmidt, A. Binding Specificity of Amino Acids to Amorphous Silica Surfaces: Solid-State NMR of Glycine on SBA-15. *J. Phys. Chem. C* **2012**, *116* (17), 9691–9702. <https://doi.org/10.1021/jp302431t>.
- (87) Ben Shir, I.; Kababya, S.; Schmidt, A. Molecular Details of Amorphous Silica Surfaces Determine Binding Specificity to Small Amino Acids. *J. Phys. Chem. C* **2014**, *118* (15), 7901–7909. <https://doi.org/10.1021/jp409030w>.

- (88) Martra, G.; Deiana, C.; Sakhno, Y.; Barberis, I.; Fabbiani, M.; Pazzi, M.; Vincenti, M. The Formation and Self-Assembly of Long Prebiotic Oligomers Produced by the Condensation of Unactivated Amino Acids on Oxide Surfaces. *Angew. Chem. Int. Ed.* **2014**, *53* (18), 4671–4674. <https://doi.org/10.1002/anie.201311089>.
- (89) Sakhno, Y.; Battistella, A.; Mezzetti, A.; Jaber, M.; Georgelin, T.; Michot, L.; Lambert, J.-F. One Step up the Ladder of Prebiotic Complexity: Formation of Nonrandom Linear Polypeptides from Binary Systems of Amino Acids on Silica. *Chem. - Eur. J.* **2019**, *25* (5), 1275–1285. <https://doi.org/10.1002/chem.201803845>.
- (90) White, D. H.; Kennedy, R. M.; Macklin, J. Acyl Silicates and Acyl Aluminates as Activated Intermediates in Peptide Formation on Clays. *Orig. Life* **1984**, *14* (1), 273–278. <https://doi.org/10.1007/BF00933668>.
- (91) Abadian, H. Prebiotic Organization of Biomolecules on Mineral Surfaces, Sorbonne Université, Paris, 2021.
- (92) Rimola, A.; Civalleri, B.; Ugliengo, P. Neutral vs Zwitterionic Glycine Forms at the Water/Silica Interface: Structure, Energies, and Vibrational Features from B3LYP Periodic Simulations. *Langmuir* **2008**, *24* (24), 14027–14034. <https://doi.org/10.1021/la8029352>.
- (93) Jaber, M.; Georgelin, T.; Bazzi, H.; Costa-Torro, F.; Lambert, J.-F.; Bolbach, G.; Clodic, G. Selectivities in Adsorption and Peptidic Condensation in the (Arginine and Glutamic Acid)/Montmorillonite Clay System. *J. Phys. Chem. C* **2014**, *118* (44), 25447–25455. <https://doi.org/10.1021/jp507335e>.
- (94) Lambert, J.-F.; Jaber, M.; Georgelin, T.; Stievano, L. A Comparative Study of the Catalysis of Peptide Bond Formation by Oxide Surfaces. *Phys. Chem. Chem. Phys.* **2013**, *15* (32), 13371. <https://doi.org/10.1039/c3cp51282g>.
- (95) de Castro Silva, F.; Lima, L. C. B.; Silva-Filho, E. C.; Fonseca, M. G.; Lambert, J.-F.; Jaber, M. A Comparative Study of Alanine Adsorption and Condensation to Peptides in Two Clay Minerals. *Appl. Clay Sci.* **2020**, *192*, 105617. <https://doi.org/10.1016/j.clay.2020.105617>.
- (96) Jaber, M.; Spadavecchia, J.; Bazzi, H.; Georgelin, T.; Costa-Torro, F.; Lambert, J.-F. Non-Biological Selectivity in Amino Acids Polymerization on TiO<sub>2</sub> Nanoparticles. *Amino Acids* **2013**, *45* (2), 403–406. <https://doi.org/10.1007/s00726-013-1516-y>.
- (97) Georgelin, T.; Akouche, M.; Jaber, M.; Sakhno, Y.; Matheron, L.; Fournier, F.; Méthivier, C.; Martra, G.; Lambert, J.-F. Iron(III) Oxide Nanoparticles as Catalysts for the Formation of Linear Glycine Peptides. *Eur. J. Inorg. Chem.* **2017**, *2017* (1), 198–211. <https://doi.org/10.1002/ejic.201601296>.
- (98) Shanker, U.; Bhushan, B.; Bhattacharjee, G.; Kamaluddin. Oligomerization of Glycine and Alanine Catalyzed by Iron Oxides: Implications for Prebiotic Chemistry. *Orig. Life Evol. Biospheres* **2012**, *42* (1), 31–45. <https://doi.org/10.1007/s11084-012-9266-5>.
- (99) Huber, C.; Eisenreich, W.; Hecht, S.; Wächtershäuser, G. A Possible Primordial Peptide Cycle. *Science* **2003**, *301* (5635), 938–940. <https://doi.org/10.1126/science.1086501>.
- (100) Holger Försterling, F. *Spin Dynamics: Basics of Nuclear Magnetic Resonance, Second Edition. Med. Phys.* **2009**, *37* (1), 406–407. <https://doi.org/10.1118/1.3273534>.
- (101) Yu, S.-S.; Krishnamurthy, R.; Fernández, F. M.; Hud, N. V.; Schork, F. J.; Grover, M. A. Kinetics of Prebiotic Dipeptide Formation from the Ester–Amide Exchange Reaction. *Phys. Chem. Chem. Phys.* **2016**, *18* (41), 28441–28450. <https://doi.org/10.1039/C6CP0527C>.
- (102) Forsythe, J. G.; Petrov, A. S.; Millar, W. C.; Yu, S.-S.; Krishnamurthy, R.; Grover, M. A.; Hud, N. V.; Fernández, F. M. Surveying the Sequence Diversity of Model Prebiotic Peptides by Mass Spectrometry. *Proc. Natl. Acad. Sci.* **2017**, *114* (37), E7652–E7659. <https://doi.org/10.1073/pnas.1711631114>.
- (103) Alpert, A. J. Hydrophilic-Interaction Chromatography for the Separation of Peptides, Nucleic Acids and Other Polar Compounds. *J. Chromatogr. A* **1990**, *499*, 177–196. [https://doi.org/10.1016/S0021-9673\(00\)96972-3](https://doi.org/10.1016/S0021-9673(00)96972-3).
- (104) Hao, Z.; Lu, C.-Y. J.; Xiao, B.; Weng, N.; Parker, B.; Knapp, M.; Ho, C.-T. Separation of Amino Acids, Peptides and Corresponding Amadori Compounds on a Silica Column at Elevated Temperature. *J. Chromatogr. A* **2007**, *1147* (2), 165–171. <https://doi.org/10.1016/j.chroma.2007.02.057>.
- (105) Yu, S.-S.; Solano, M. D.; Blanchard, M. K.; Soper-Hopper, M. T.; Krishnamurthy, R.; Fernández, F. M.; Hud, N. V.; Schork, F. J.; Grover, M. A. Elongation of Model Prebiotic Proto-Peptides by Continuous Monomer Feeding. *Macromolecules* **2017**, *50* (23), 9286–9294. <https://doi.org/10.1021/acs.macromol.7b01569>.
- (106) Harscoat-Schiavo, C.; Raminosa, F.; Ronat-Heit, E.; Vanderesse, R.; Marc, I. Modeling the Separation of Small Peptides by Cation-Exchange Chromatography. *J. Sep. Sci.* **2010**, *33* (16), 2447–2457. <https://doi.org/10.1002/jssc.201000112>.

- (107) Xie, W.; Teraoka, I.; Gross, R. A. Reversed Phase Ion-Pairing Chromatography of an Oligolysine Mixture in Different Mobile Phases: Effort of Searching Critical Chromatography Conditions. *J. Chromatogr. A* **2013**, *1304*, 127–132. <https://doi.org/10.1016/j.chroma.2013.06.078>.
- (108) Xie, W.; Qin, X.; Teraoka, I.; Gross, R. A. Comparison of Retention Behavior of Oligolysine and Oligoarginine in Ion-Pairing Chromatography Using Heptafluorobutyric Acid. *Anal. Bioanal. Chem.* **2013**, *405* (30), 9739–9746. <https://doi.org/10.1007/s00216-013-7397-9>.
- (109) Rode, B. M.; Schwendinger, M. G. Copper-Catalyzed Amino Acid Condensation in Water — A Simple Possible Way of Prebiotic Peptide Formation. *Orig. Life Evol. Biosph.* **1990**, *20* (5), 401–410. <https://doi.org/10.1007/BF01808134>.
- (110) Ohara, S.; Kakegawa, T.; Nakazawa, H. Pressure Effects on the Abiotic Polymerization of Glycine. *Orig. Life Evol. Biosphere J. Int. Soc. Study Orig. Life* **2007**, *37* (3), 215–223. <https://doi.org/10.1007/s11084-007-9067-4>.
- (111) Dalai, P.; Pleyer, H. L.; Strasdeit, H.; Fox, S. The Influence of Mineral Matrices on the Thermal Behavior of Glycine. *Orig. Life Evol. Biospheres* **2017**, *47* (4), 427–452. <https://doi.org/10.1007/s11084-016-9523-0>.
- (112) Campbell, T. D.; Febrian, R.; Kleinschmidt, H. E.; Smith, K. A.; Bracher, P. J. Quantitative Analysis of Glycine Oligomerization by Ion-Pair Chromatography. *ACS Omega* **2019**, *4* (7), 12745–12752. <https://doi.org/10.1021/acsomega.9b01492>.
- (113) Rodriguez-Garcia, M.; Surman, A. J.; Cooper, G. J. T.; Suárez-Marina, I.; Hosni, Z.; Lee, M. P.; Cronin, L. Formation of Oligopeptides in High Yield under Simple Programmable Conditions. *Nat. Commun.* **2015**, *6* (1). <https://doi.org/10.1038/ncomms9385>.
- (114) Eckstein, J. A.; Ammerman, G. M.; Reveles, J. M.; Ackermann, B. L. Analysis of Glutamine, Glutamate, Pyroglutamate, and GABA in Cerebrospinal Fluid Using Ion Pairing HPLC with Positive Electrospray LC/MS/MS. *J. Neurosci. Methods* **2008**, *171* (2), 190–196. <https://doi.org/10.1016/j.jneumeth.2008.02.019>.
- (115) Hoaglund, C. S.; Valentine, S. J.; Sporleder, C. R.; Reilly, J. P.; Clemmer, D. E. Three-Dimensional Ion Mobility/TOFMS Analysis of Electrosprayed Biomolecules. *Anal. Chem.* **1998**, *70* (11), 2236–2242. <https://doi.org/10.1021/ac980059c>.
- (116) Silveira, J. A.; Ridgeway, M. E.; Park, M. A. High Resolution Trapped Ion Mobility Spectrometry of Peptides. *Anal. Chem.* **2014**, *86* (12), 5624–5627. <https://doi.org/10.1021/ac501261h>.
- (117) Gabelica, V. Intérêt de la spectrométrie de mobilité ionique pour l'étude de la conformation et des assemblages non-covalents de biomolécules. **2006**, *7*.
- (118) Revercomb, H. E.; Mason, E. A. Theory of Plasma Chromatography/Gaseous Electrophoresis. Review. *Anal. Chem.* **1975**, *47* (7), 970–983. <https://doi.org/10.1021/ac60357a043>.
- (119) Gabelica, V.; Marklund, E. Fundamentals of Ion Mobility Spectrometry. *Curr. Opin. Chem. Biol.* **2018**, *42*, 51–59. <https://doi.org/10.1016/j.cbpa.2017.10.022>.
- (120) McLean, J. A.; Ruotolo, B. T.; Gillig, K. J.; Russell, D. H. Ion Mobility–Mass Spectrometry: A New Paradigm for Proteomics. *Int. J. Mass Spectrom.* **2005**, *240* (3), 301–315. <https://doi.org/10.1016/j.ijms.2004.10.003>.
- (121) Domalain, V.; Tognetti, V.; Hubert-Roux, M.; Lange, C. M.; Joubert, L.; Baudoux, J.; Rouden, J.; Afonso, C. Role of Cationization and Multimers Formation for Diastereomers Differentiation by Ion Mobility–Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2013**, *24* (9), 1437–1445. <https://doi.org/10.1007/s13361-013-0690-1>.
- (122) Domalain, V.; Hubert-Roux, M.; Tognetti, V.; Joubert, L.; Lange, C. M.; Rouden, J.; Afonso, C. Enantiomeric Differentiation of Aromatic Amino Acids Using Traveling Wave Ion Mobility–Mass Spectrometry. *Chem Sci* **2014**, *5* (8), 3234–3239. <https://doi.org/10.1039/C4SC00443D>.
- (123) Dwivedi, P.; Wu, C.; Matz, L. M.; Clowers, B. H.; Siems, W. F.; Hill, H. H. Gas-Phase Chiral Separations by Ion Mobility Spectrometry. *Anal. Chem.* **2006**, *78* (24), 8200–8206. <https://doi.org/10.1021/ac0608772>.
- (124) Laphorn, C.; Pullen, F.; Chowdhry, B. Z. Ion Mobility Spectrometry–Mass Spectrometry (IMS–MS) of Small Molecules: Separating and Assigning Structures to Ions. *Mass Spectrom. Rev.* **2013**, *32* (1), 43–71. <https://doi.org/10.1002/mas.21349>.
- (125) Bruins, A. P. Mechanistic Aspects of Electrospray Ionization. *J. Chromatogr. A* **1998**, *794* (1–2), 345–357. [https://doi.org/10.1016/S0021-9673\(97\)01110-2](https://doi.org/10.1016/S0021-9673(97)01110-2).
- (126) Alves, S.; Rathahao-Paris, E.; Tabet, J. C. Potential of Fourier Transform Mass Spectrometry for High-Throughput Metabolomics Analysis. In *Advances in Botanical Research*; Elsevier, 2013; Vol. 67, pp 219–302. <https://doi.org/10.1016/B978-0-12-397922-3.00005-8>.



- (127) 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. [https://doi.org/10.1002/\(SICI\)1098-2787\(1998\)17:1<1::AID-MAS1>3.0.CO;2-K](https://doi.org/10.1002/(SICI)1098-2787(1998)17:1<1::AID-MAS1>3.0.CO;2-K).
- (128) Moco, S.; Vervoort, J.; Moco, S.; Bino, R. J.; De Vos, R. C. H.; Bino, R. Metabolomics Technologies and Metabolite Identification. *TrAC Trends Anal. Chem.* **2007**, *26* (9), 855–866. <https://doi.org/10.1016/j.trac.2007.08.003>.
- (129) Cech, N. B.; Enke, C. G. Practical Implications of Some Recent Studies in Electrospray Ionization Fundamentals. *Mass Spectrom. Rev.* **2001**, *20* (6), 362–387. <https://doi.org/10.1002/mas.10008>.
- (130) Annesley, T. M. Ion Suppression in Mass Spectrometry. *Clin. Chem.* **2003**, *49* (7), 1041–1044. <https://doi.org/10.1373/49.7.1041>.
- (131) 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. [https://doi.org/10.1002/\(SICI\)1098-2787\(1998\)17:1<1::AID-MAS1>3.0.CO;2-K](https://doi.org/10.1002/(SICI)1098-2787(1998)17:1<1::AID-MAS1>3.0.CO;2-K).
- (132) Bayat, P.; Lesage, D.; Cole, R. B. Tutorial: Ion Activation in Tandem Mass Spectrometry Using Ultra-High Resolution Instrumentation. *Mass Spectrom. Rev.* **2020**, *39* (5–6), 680–702. <https://doi.org/10.1002/mas.21623>.
- (133) Peltz, C.; Drahos, L.; Vékey, K. SORI Excitation: Collisional and Radiative Processes. *J. Am. Soc. Mass Spectrom.* **2007**, *18* (12), 2119–2126. <https://doi.org/10.1016/j.jasms.2007.09.011>.
- (134) Aebersold, R.; Goodlett, D. R. Mass Spectrometry in Proteomics. *Chem. Rev.* **2001**, *101* (2), 269–296. <https://doi.org/10.1021/cr990076h>.
- (135) Roepstorff, P.; Fohlman, J. Letter to the Editors. *Biol. Mass Spectrom.* **1984**, *11* (11), 601–601. <https://doi.org/10.1002/bms.1200111109>.
- (136) Harrison, A. G.; Csizmadia, I. G.; Tang, T.-H. Structure and Fragmentation of B2 Ions in Peptide Mass Spectra. *J. Am. Soc. Mass Spectrom.* **2000**, *11* (5), 427–436. [https://doi.org/10.1016/S1044-0305\(00\)00104-5](https://doi.org/10.1016/S1044-0305(00)00104-5).
- (137) Guigue, J.; Harir, M.; Mathieu, O.; Lucio, M.; Ranjard, L.; Lévêque, J.; Schmitt-Kopplin, P. Ultrahigh-Resolution FT-ICR Mass Spectrometry for Molecular Characterisation of Pressurised Hot Water-Extractable Organic Matter in Soils. *Biogeochemistry* **2016**, *128* (3), 307–326. <https://doi.org/10.1007/s10533-016-0209-5>.
- (138) Hertkorn, N.; Frommberger, M.; Witt, M.; Koch, B. P.; Schmitt-Kopplin, Ph.; Perdue, E. M. Natural Organic Matter and the Event Horizon of Mass Spectrometry. *Anal. Chem.* **2008**, *80* (23), 8908–8919. <https://doi.org/10.1021/ac800464g>.
- (139) Schmitt-Kopplin, P.; Gabelica, Z.; Gougeon, R. D.; Fekete, A.; Kanawati, B.; Harir, M.; Gebefuegi, I.; Eckel, G.; Hertkorn, N. High Molecular Diversity of Extraterrestrial Organic Matter in Murchison Meteorite Revealed 40 Years after Its Fall. *Proc. Natl. Acad. Sci.* **2010**, *107* (7), 2763–2768. <https://doi.org/10.1073/pnas.0912157107>.
- (140) Hertkorn, N.; Harir, M.; Schmitt-Kopplin, Ph. Nontarget Analysis of Murchison Soluble Organic Matter by High-Field NMR Spectroscopy and FTICR Mass Spectrometry: Murchison Organic Matter NMR. *Magn. Reson. Chem.* **2015**, *53* (9), 754–768. <https://doi.org/10.1002/mrc.4249>.
- (141) Danger, G.; Ruf, A.; Maillard, J.; Hertzog, J.; Vinogradoff, V.; Schmitt-Kopplin, P.; Afonso, C.; Carrasco, N.; Schmitz-Afonso, I.; d’Hendecourt, L. L. S.; Remusat, L. Unprecedented Molecular Diversity Revealed in Meteoritic Insoluble Organic Matter: The Paris Meteorite’s Case. *Planet. Sci. J.* **2020**, *1* (3), 55. <https://doi.org/10.3847/PSJ/abb60f>.
- (142) Briois, C.; Thissen, R.; Thirkell, L.; Aradj, K.; Bouabdellah, A.; Boukrara, A.; Carrasco, N.; Chalumeau, G.; Chapelon, O.; Colin, F.; Coll, P.; Cottin, H.; Engrand, C.; Grand, N.; Lebreton, J.-P.; Orthous-Daunay, F.-R.; Pennanech, C.; Szopa, C.; Vuitton, V.; Zapf, P.; Makarov, A. Orbitrap Mass Analyser for in Situ Characterisation of Planetary Environments: Performance Evaluation of a Laboratory Prototype. *Planet. Space Sci.* **2016**, *131*, 33–45. <https://doi.org/10.1016/j.pss.2016.06.012>.
- (143) Wollrab, E.; Scherer, S.; Aubriet, F.; Carré, V.; Carlomagno, T.; Codutti, L.; Ott, A. Chemical Analysis of a “Miller-Type” Complex Prebiotic Broth: Part I: Chemical Diversity, Oxygen and Nitrogen Based Polymers. *Orig. Life Evol. Biospheres* **2016**, *46* (2–3), 149–169. <https://doi.org/10.1007/s11084-015-9468-8>.
- (144) Bouza, M.; Li, A.; Forsythe, J. G.; Petrov, A.; Wang, Z. L.; Fernández, F. M. Compositional Characterization of Complex Protopeptide Libraries via Triboelectric Nanogenerator Orbitrap Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2019**, *33* (16), 1293–1300. <https://doi.org/10.1002/rcm.8469>.

- (145) Wincel, H.; Fokkens, R. H.; Nibbering, N. M. M. Peptide Bond Formation in Gas-phase Ion/Molecule Reactions of Amino Acids: A Novel Proposal for the Synthesis of Prebiotic Oligopeptides. *Rapid Commun Mass Spectrom* **2000**, *6*.
- (146) Counterman, A. E.; Clemmer, D. E. Magic Number Clusters of Serine in the Gas Phase. *J. Phys. Chem. B* **2001**, *105* (34), 8092–8096. <https://doi.org/10.1021/jp011421l>.
- (147) Cooks, R. G.; Zhang, D.; Koch, K. J.; Gozzo, F. C.; Eberlin, M. N. Chiroselective Self-Directed Octamerization of Serine: Implications for Homochirogenesis. *Anal. Chem.* **2001**, *73* (15), 3646–3655. <https://doi.org/10.1021/ac010284l>.
- (148) Hodyss, R.; Julian, R. R.; Beauchamp, J. L. Spontaneous Chiral Separation in Noncovalent Molecular Clusters. *Chirality* **2001**, *13* (10), 703–706. <https://doi.org/10.1002/chir.10021>.
- (149) Takats, Z.; Nanita, S. C.; Cooks, R. G. Serine Octamer Reactions: Indicators of Prebiotic Relevance. *Angew. Chem. Int. Ed.* **2003**, *42* (30), 3521–3523. <https://doi.org/10.1002/anie.200351210>.
- (150) Illos, R. A.; Bisogno, F. R.; Clodic, G.; Bolbach, G.; Weissbuch, I.; Lahav, M. Oligopeptides and Copeptides of Homochiral Sequence, via  $\beta$ -Sheets, from Mixtures of Racemic  $\alpha$ -Amino Acids, in a One-Pot Reaction in Water; Relevance to Biochirogenesis. *J. Am. Chem. Soc.* **2008**, *130* (27), 8651–8659. <https://doi.org/10.1021/ja709969v>.
- (151) Granold, M.; Hajieva, P.; Toşa, M. I.; Irimie, F.-D.; Moosmann, B. Modern Diversification of the Amino Acid Repertoire Driven by Oxygen. *Proc. Natl. Acad. Sci.* **2018**, *115* (1), 41–46. <https://doi.org/10.1073/pnas.1717100115>.
- (152) Shibue, R.; Sasamoto, T.; Shimada, M.; Zhang, B.; Yamagishi, A.; Akanuma, S. Comprehensive Reduction of Amino Acid Set in a Protein Suggests the Importance of Prebiotic Amino Acids for Stable Proteins. *Sci. Rep.* **2018**, *8* (1), 1227. <https://doi.org/10.1038/s41598-018-19561-1>.
- (153) Cronin, J. R. Acid-Labile Amino Acid Precursors in the Murchison Meteorite. *Orig. Life* **1976**, *7* (4), 337–342. <https://doi.org/10.1007/BF00927928>.
- (154) Bertrand, M.; Brack, A. Conformational Variety of Polyanionic Peptides at Low Salt Concentrations. *Orig. Life Evol. Biosphere J. Int. Soc. Study Orig. Life* **1997**, *27* (5–6), 585–595. <https://doi.org/10.1023/a:1006585925390>.
- (155) Kuge, K.; Brack, A.; Fujii, N. Conformation-Dependent Racemization of Aspartyl Residues in Peptides. *Chem. - Eur. J.* **2007**, *13* (19), 5617–5621. <https://doi.org/10.1002/chem.200601677>.
- (156) Miller, S. L. Which Organic Compounds Could Have Occurred on the Prebiotic Earth? *Cold Spring Harb. Symp. Quant. Biol.* **1987**, *52* (0), 17–27. <https://doi.org/10.1101/SQB.1987.052.01.005>.
- (157) van der Gulik, P.; Massar, S.; Gilis, D.; Buhrman, H.; Roodman, M. The First Peptides: The Evolutionary Transition between Prebiotic Amino Acids and Early Proteins. *J. Theor. Biol.* **2009**, *261*, 531–539. <https://doi.org/10.1016/j.jtbi.2009.09.004>.
- (158) Ikehara, K. Possible Steps to the Emergence of Life: The [GADV]-Protein World Hypothesis. *Chem. Rec.* **2005**, *5* (2), 107–118. <https://doi.org/10.1002/tcr.20037>.
- (159) Ikehara, K. [GADV]-Protein World Hypothesis on the Origin of Life. *Orig. Life Evol. Biosph.* **2014**, *44* (4), 299–302. <https://doi.org/10.1007/s11084-014-9383-4>.
- (160) Walker, M. C.; van der Donk, W. A. The Many Roles of Glutamate in Metabolism. *J. Ind. Microbiol. Biotechnol.* **2016**, *43* (0), 419–430. <https://doi.org/10.1007/s10295-015-1665-y>.
- (161) Baglioni, C. The Role of Pyrrolidone Carboxylic Acid in the Initiation of Immunoglobulin Peptide Chains. *Biochem. Biophys. Res. Commun.* **1970**, *38* (2), 212–219. [https://doi.org/10.1016/0006-291X\(70\)90698-4](https://doi.org/10.1016/0006-291X(70)90698-4).
- (162) Melville, J. Labile Glutamine Peptides, and Their Bearing on the Origin of the Ammonia Set Free during the Enzymic Digestion of Proteins. *Biochem. J.* **1935**, *29* (1), 179–186. <https://doi.org/10.1042/bj0290179>.
- (163) Messer, M.; Ottesen, M. Isolation and Properties of Glutamine Cyclotransferase of Dried Papaya Latex. *Comptes-Rendus Trav. Lab. Carlsberg* **1965**, *35* (1), 1–24.
- (164) Blombäck, B. [44] Derivatives of Glutamine in Peptides. In *Methods in Enzymology*; Enzyme Structure; Academic Press, 1967; Vol. 11, pp 398–411. [https://doi.org/10.1016/S0076-6879\(67\)11046-X](https://doi.org/10.1016/S0076-6879(67)11046-X).
- (165) Rehder, D. S.; Dillon, T. M.; Pipes, G. D.; Bondarenko, P. V. Reversed-Phase Liquid Chromatography/Mass Spectrometry Analysis of Reduced Monoclonal Antibodies in Pharmaceuticals. *J. Chromatogr. A* **2006**, *1102* (1–2), 164–175. <https://doi.org/10.1016/j.chroma.2005.10.053>.
- (166) Twardzik, D. R.; Peterkofsky, A. Glutamic Acid as a Precursor to N-Terminal Pyroglutamic Acid in Mouse Plasmacytoma Protein. *Proc. Natl. Acad. Sci.* **1972**, *69* (1), 274–277. <https://doi.org/10.1073/pnas.69.1.274>.

- (167) Jones, G. H. Cell-Free Synthesis of Amino-Terminal L-Pyroglutamic Acid. *Biochemistry* **1974**, *13* (5), 855–860. <https://doi.org/10.1021/bi00702a003>.
- (168) Chelius, D.; Jing, K.; Lueras, A.; Rehder, D. S.; Dillon, T. M.; Vizel, A.; Rajan, R. S.; Li, T.; Treuheit, M. J.; Bondarenko, P. V. Formation of Pyroglutamic Acid from N-Terminal Glutamic Acid in Immunoglobulin Gamma Antibodies. *Anal. Chem.* **2006**, *78* (7), 2370–2376. <https://doi.org/10.1021/ac051827k>.
- (169) Shih, F. F. Analysis of Glutamine, Glutamic Acid and Pyroglutamic Acid in Protein Hydrolysates by High-Performance Liquid Chromatography. *J. Chromatogr. A* **1985**, *322*, 248–256. [https://doi.org/10.1016/S0021-9673\(01\)97681-2](https://doi.org/10.1016/S0021-9673(01)97681-2).
- (170) Wilson, H.; Cannan, R. K. The Glutamic Acid-Pyrrolidonecarboxylic Acid System. *J. Biol. Chem.* **1937**, *119* (1), 309–331.
- (171) Clayton, D. W.; Kenner, G. W.; Sheppard, R. C. 76. Peptides. Part V. Condensation of the  $\gamma$ -Carboxyl Group of  $\alpha$ -Glutamyl Peptides with the Peptide Chain. *J Chem Soc* **1956**, *0* (0), 371–380. <https://doi.org/10.1039/JR9560000371>.
- (172) Howard, J. B.; Vermeulen, M.; Swenson, R. P. The Temperature-Sensitive Bond in Human Alpha 2-Macroglobulin Is the Alkylamine-Reactive Site. *J. Biol. Chem.* **1980**, *255* (9), 3820–3823.
- (173) Khan, S. A.; Erickson, B. W. Synthesis and Regioselective Hydrolysis of Peptides Containing an Internal Residue of Pyroglutamic Acid. *J. Am. Chem. Soc.* **1984**, *106* (3), 798–799. <https://doi.org/10.1021/ja00315a060>.
- (174) Khan, S. A.; Sekulski, J. M.; Erickson, B. W. Peptide Models of Protein Metastable Binding Sites: Competitive Kinetics of Isomerization and Hydrolysis. *Biochemistry* **1986**, *25* (18), 5165–5171. <https://doi.org/10.1021/bi00366a027>.
- (175) Harada, K.; Fox, S. W. The Thermal Condensation of Glutamic Acid and Glycine to Linear Peptides <sup>1</sup>. *J. Am. Chem. Soc.* **1958**, *80* (11), 2694–2697. <https://doi.org/10.1021/ja01544a027>.
- (176) Melius, P.; Hubbard, W.-L. Pyroglutamyl N-Termini of Thermal Polyamino Acids. *Biosystems* **1987**, *20* (3), 213–217. [https://doi.org/10.1016/0303-2647\(87\)90027-X](https://doi.org/10.1016/0303-2647(87)90027-X).
- (177) Meister, A.; Anderson, M. E.; Hwang, O. Intracellular Cysteine and Glutathione Delivery Systems. *J. Am. Coll. Nutr.* **1986**, *5* (2), 137–151. <https://doi.org/10.1080/07315724.1986.10720121>.
- (178) Yang, J.; Bai, W.; Zeng, X.; Cui, C. Gamma Glutamyl Peptides: The Food Source, Enzymatic Synthesis, Kokumi-Active and the Potential Functional Properties – A Review. *Trends Food Sci. Technol.* **2019**, *91*, 339–346. <https://doi.org/10.1016/j.tifs.2019.07.022>.
- (179) Waley, S. G. Naturally Occurring Peptides. In *Advances in Protein Chemistry*; Anfinsen, C. B., Anson, M. L., Edsall, J. T., Richards, F. M., Eds.; Academic Press, 1966; Vol. 21, pp 1–112. [https://doi.org/10.1016/S0065-3233\(08\)60127-8](https://doi.org/10.1016/S0065-3233(08)60127-8).
- (180) Kasai, T.; Larsen, P. O. Chemistry and Biochemistry of  $\gamma$ -Glutamyl Derivatives from Plants Including Mushrooms (Basidiomycetes). In *Fortschritte der Chemie organischer Naturstoffe / Progress in the Chemistry of Organic Natural Products*; Anderson, R. C., Fraser-Reid, B., Jones, H., Kasai, T., Larsen, P. O., Liaaen-Jensen, S., Rasmussen, G. H., Herz, W., Grisebach, H., Kirby, G. W., Eds.; Fortschritte der Chemie organischer Naturstoffe / Progress in the Chemistry of Organic Natural Products; Springer: Vienna, 1980; pp 173–285. [https://doi.org/10.1007/978-3-7091-8551-3\\_4](https://doi.org/10.1007/978-3-7091-8551-3_4).
- (181) Frerot, E.; Chen, T. Identification and Quantitation of New Glutamic Acid Derivatives in Soy Sauce by UPLC/MS/MS. *Chem. Biodivers.* **2013**, *10* (10), 1842–1850. <https://doi.org/10.1002/cbdv.201300150>.
- (182) Roudot-Algaron, F.; Kerhoas, L.; Le Bars, D.; Einhorn, J.; Gripon, J. C. Isolation of  $\gamma$ -Glutamyl Peptides from Comté Cheese. *J. Dairy Sci.* **1994**, *77* (5), 1161–1166. [https://doi.org/10.3168/jds.S0022-0302\(94\)77053-3](https://doi.org/10.3168/jds.S0022-0302(94)77053-3).
- (183) Shibata, M.; Hirotsuka, M.; Mizutani, Y.; Takahashi, H.; Kawada, T.; Matsumiya, K.; Hayashi, Y.; Matsumura, Y. Diversity of  $\gamma$ -Glutamyl Peptides and Oligosaccharides, the “Kokumi” Taste Enhancers, in Seeds from Soybean Mini Core Collections. *Biosci. Biotechnol. Biochem.* **2018**, *82* (3), 507–514. <https://doi.org/10.1080/09168451.2018.1436960>.
- (184) Ichinose, H.; Togari, A.; Suzuki, H.; Kumagai, H.; Nagatsu, T. Increase of Catecholamines in Mouse Brain by Systemic Administration of  $\gamma$ -Glutamyl L-3,4-Dihydroxyphenylalanine. *J. Neurochem.* **1987**, *49* (3), 928–932. <https://doi.org/10.1111/j.1471-4159.1987.tb00982.x>.
- (185) Main, P. A.; Angley, M. T.; O’Doherty, C. E.; Thomas, P.; Fenech, M. The Potential Role of the Antioxidant and Detoxification Properties of Glutathione in Autism Spectrum Disorders: A Systematic Review and Meta-Analysis. *Nutr. Metab.* **2012**, *9* (1), 35. <https://doi.org/10.1186/1743-7075-9-35>.
- (186) Oakley, A. J.; Yamada, T.; Liu, D.; Coggan, M.; Clark, A. G.; Board, P. G. The Identification and Structural Characterization of C7orf24 as  $\gamma$ -Glutamyl Cyclotransferase. *J. Biol. Chem.* **2008**, *283* (32), 22031–22042. <https://doi.org/10.1074/jbc.M803623200>.

- (187) Kuroda, M.; Miyamura, N. Mechanism of the Perception of “Kokumi” Substances and the Sensory Characteristics of the “Kokumi” Peptide,  $\gamma$ -Glu-Val-Gly. *Flavour* **2015**, *4* (1), 11. <https://doi.org/10.1186/2044-7248-4-11>.
- (188) Miyaki, T.; Kawasaki, H.; Kuroda, M.; Miyamura, N.; Kouda, T. Effect of a Kokumi Peptide,  $\gamma$ -Glutamyl-Valyl-Glycine, on the Sensory Characteristics of Chicken Consommé. *Flavour* **2015**, *4* (1), 17. <https://doi.org/10.1186/2044-7248-4-17>.
- (189) Zhang, H.; Hu, C.-A. A.; Kovacs-Nolan, J.; Mine, Y. Bioactive Dietary Peptides and Amino Acids in Inflammatory Bowel Disease. *Amino Acids* **2015**, *47* (10), 2127–2141. <https://doi.org/10.1007/s00726-014-1886-9>.
- (190) Zhang, H.; Kodera, T.; Eto, Y.; Mine, Y.  $\gamma$ -Glutamyl Valine Supplementation-Induced Mitigation of Gut Inflammation in a Porcine Model of Colitis. *J. Funct. Foods* **2016**, *24*, 558–567. <https://doi.org/10.1016/j.jff.2016.04.038>.
- (191) Kasai, T.; Nishitoba, T.; Sakamura, S. Transformation of Glutamyl Dipeptides by Heating in Aqueous Solution. *Agric. Biol. Chem.* **1983**, *47* (11), 2647–2649. <https://doi.org/10.1080/00021369.1983.10866007>.
- (192) Ching, C. B.; Hidajat, K.; Uddin, M. S. Evaluation of Equilibrium and Kinetic Parameters of Smaller Molecular Size Amino Acids on KX Zeolite Crystals via Liquid Chromatographic Techniques. *Scopus* **1989**.
- (193) Nuruzzaman, Md.; Rahman, M. M.; Liu, Y.; Naidu, R. Nanoencapsulation, Nano-Guard for Pesticides: A New Window for Safe Application. *J. Agric. Food Chem.* **2016**, *64* (7), 1447–1483. <https://doi.org/10.1021/acs.jafc.5b05214>.
- (194) Biondi, E.; Branciamore, S.; Maurel, M.-C.; Gallori, E. Montmorillonite Protection of an UV-Irradiated Hairpin Ribozyme: Evolution of the RNA World in a Mineral Environment. *BMC Evol. Biol.* **2007**, *7* (2), S2. <https://doi.org/10.1186/1471-2148-7-S2-S2>.
- (195) Banin, A.; Lawless, J. G.; Mazzurco, J.; Church, F. M.; Margulies, L.; Orenberg, J. B. PH Profile of the Adsorption of Nucleotides onto Montmorillonite. *Orig. Life Evol. Biosph.* **1985**, *15* (2), 89–101. <https://doi.org/10.1007/BF01809491>.
- (196) Vieira Coelho, A. C.; Rocha, G. A.; Souza Santos, P.; Souza Santos, H.; Kiyohara, P. K. Specific Surface Area and Structures of Aluminas from Fibrillar Pseudoboehmite. *Matér. Rio Jan.* **2008**, *13* (2), 329–341. <https://doi.org/10.1590/S1517-70762008000200011>.
- (197) Levin, I.; Brandon, D. Metastable Alumina Polymorphs: Crystal Structures and Transition Sequences. *J. Am. Ceram. Soc.* **1998**, *81* (8), 1995–2012. <https://doi.org/10.1111/j.1151-2916.1998.tb02581.x>.
- (198) Busca. The Surface of Transitional Aluminas: A Critical Review. *Catal. Today* **2014**, *226*, 2–13. <https://doi.org/10.1016/j.cattod.2013.08.003>.
- (199) Terterov, I.; Vyatkina, K.; Kononikhin, A. S.; Boitsov, V.; Vyazmin, S.; Popov, I. A.; Nikolaev, E. N.; Pevzner, P.; Dubina, M. Application of *de Novo* Sequencing Tools to Study Abiogenic Peptide Formations by Tandem Mass Spectrometry. The Case of Homo-Peptides from Glutamic Acid Complicated by Substitutions of Hydrogen by Sodium or Potassium Atoms: Imaging of Photocleavable Peptide Bead-Arrays. *Rapid Commun. Mass Spectrom.* **2014**, *28* (1), 33–41. <https://doi.org/10.1002/rcm.6757>.
- (200) Bergquist, J.; Baykut, G.; Bergquist, M.; Witt, M.; Mayer, F.-J.; Baykut, D. Human Myocardial Protein Pattern Reveals Cardiac Diseases. *Int. J. Proteomics* **2012**, *2012*, 1–17. <https://doi.org/10.1155/2012/342659>.
- (201) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: A Primer. 35.
- (202) Gabrielse, G.; Haarsma, L.; Rolston, S. L. Open-Endcap Penning Traps for High Precision Experiments. *Int. J. Mass Spectrom. Ion Process.* **1989**, *88* (2), 319–332. [https://doi.org/10.1016/0168-1176\(89\)85027-X](https://doi.org/10.1016/0168-1176(89)85027-X).
- (203) Tolmachev, A. V.; Robinson, E. W.; Wu, S.; Kang, H.; Lourette, N. M.; Paša-Tolić, L.; Smith, R. D. Trapped-Ion Cell with Improved DC Potential Harmonicity for FT-ICR MS. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (4), 586–597. <https://doi.org/10.1016/j.jasms.2008.01.006>.
- (204) Brustkern, A. M.; Rempel, D. L.; Gross, M. L. An Electrically Compensated Trap Designed to Eighth Order for FT-ICR Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (9), 1281–1285. <https://doi.org/10.1016/j.jasms.2008.05.016>.
- (205) Bruce, J. E.; Anderson, G. A.; Lin, C. Y.; Gorshkov, M.; Rockwood, A. L.; Smith, R. D. A Novel High-Performance Fourier Transform Ion Cyclotron Resonance Cell for Improved Biopolymer

- Characterization. *J. Mass Spectrom. JMS* **2000**, *35* (1), 85–94. [https://doi.org/10.1002/\(SICI\)1096-9888\(200001\)35:1<85::AID-JMS910>3.0.CO;2-7](https://doi.org/10.1002/(SICI)1096-9888(200001)35:1<85::AID-JMS910>3.0.CO;2-7).
- (206) Nikolaev, E. N.; Boldin, I. A.; Jertz, R.; Baykut, G. Initial Experimental Characterization of a New Ultra-High Resolution FTICR Cell with Dynamic Harmonization. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (7), 1125–1133. <https://doi.org/10.1007/s13361-011-0125-9>.
- (207) Ridgeway, M. E.; Lubeck, M.; Jordens, J.; Mann, M.; Park, M. A. Trapped Ion Mobility Spectrometry: A Short Review. *Int. J. Mass Spectrom.* **2018**, *425*, 22–35. <https://doi.org/10.1016/j.ijms.2018.01.006>.
- (208) Silveira, J. A.; Ridgeway, M. E.; Laukien, F. H.; Mann, M.; Park, M. A. Parallel Accumulation for 100% Duty Cycle Trapped Ion Mobility-Mass Spectrometry. *Int. J. Mass Spectrom.* **2017**, *C* (413), 168–175. <https://doi.org/10.1016/j.ijms.2016.03.004>.
- (209) Nunes, R. S.; Cavalheiro, É. T. G. Thermalbehavior of Glutamic Acid and Its Sodium, Lithium and Ammonium Salts. *J. Therm. Anal. Calorim.* **2007**, *87* (3), 627–630. <https://doi.org/10.1007/s10973-006-7788-7>.
- (210) Lill, J. Proteomic Tools for Quantitation by Mass Spectrometry. *Mass Spectrom. Rev.* **2003**, *22* (3), 182–194. <https://doi.org/10.1002/mas.10048>.
- (211) Ehler, K. W.; Orgel, L. E. N,N'-Carbonyldiimidazole-Induced Peptide Formation in Aqueous Solution. *Biochim. Biophys. Acta BBA - Protein Struct.* **1976**, *434* (1), 233–243. [https://doi.org/10.1016/0005-2795\(76\)90055-6](https://doi.org/10.1016/0005-2795(76)90055-6).
- (212) Brack, A.; Ehler, K. W.; Orgel, L. E. N, N'-Carbonyldiimidazole-Induced Diketopiperazine Formation in Aqueous Solution in the Presence of Adenosine-5'-Monophosphate. *J. Mol. Evol.* **1976**, *8* (4), 307–310. <https://doi.org/10.1007/BF01739255>.
- (213) Chu, B. C. F.; Orgel, L. E. Inhibition of Oligo(Glutamine) Precipitation by Glutamine-Containing Peptides. *Biochem. Biophys. Res. Commun.* **2001**, *283* (2), 351–355. <https://doi.org/10.1006/bbrc.2001.4778>.
- (214) Adami, C. What Is Complexity? *BioEssays* **2002**, *24* (12), 1085–1094. <https://doi.org/10.1002/bies.10192>.
- (215) Shannon, C. E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27* (3), 379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>.
- (216) Balendiran, G. K.; Dabur, R.; Fraser, D. The Role of Glutathione in Cancer. *Cell Biochem. Funct.* **2004**, *22* (6), 343–352. <https://doi.org/10.1002/cbf.1149>.
- (217) Frenkel-Pinter, M.; Haynes, J. W.; C, M.; Petrov, A. S.; Burcar, B. T.; Krishnamurthy, R.; Hud, N. V.; Leman, L. J.; Williams, L. D. Selective Incorporation of Proteinaceous over Nonproteinaceous Cationic Amino Acids in Model Prebiotic Oligomerization Reactions. *Proc. Natl. Acad. Sci.* **2019**, *116* (33), 16338–16346. <https://doi.org/10.1073/pnas.1904849116>.
- (218) Fuchida, S.; Masuda, H.; Shinoda, K. Peptide Formation Mechanism on Montmorillonite Under Thermal Conditions. *Orig. Life Evol. Biospheres* **2014**, *44* (1), 13–28. <https://doi.org/10.1007/s11084-014-9359-4>.
- (219) Bujdák, J.; Rode, B. M. Activated Alumina as an Energy Source for Peptide Bond Formation: Consequences for Mineral-Mediated Prebiotic Processes. *Amino Acids* **2001**, *21* (3), 281–291. <https://doi.org/10.1007/s007260170014>.
- (220) Basiuk, V. A.; Sainz-Rojas, J. Catalysis of Peptide Formation by Inorganic Oxides: High Efficiency of Alumina under Mild Conditions on the Earth-like Planets. *Adv. Space Res.* **2001**, *27* (2), 225–230. [https://doi.org/10.1016/S0273-1177\(01\)00051-5](https://doi.org/10.1016/S0273-1177(01)00051-5).
- (221) Greiner, E.; Kumar, K.; Sumit, M.; Giuffre, A.; Zhao, W.; Pedersen, J.; Sahai, N. Adsorption of L-Glutamic Acid and L-Aspartic Acid to  $\gamma$ -Al<sub>2</sub>O<sub>3</sub>. *Geochim. Cosmochim. Acta* **2014**, *133*, 142–155. <https://doi.org/10.1016/j.gca.2014.01.004>.
- (222) Kelley, D. S.; Karson, J. A.; Früh-Green, G. L.; Yoerger, D. R.; Shank, T. M.; Butterfield, D. A.; Hayes, J. M.; Schrenk, M. O.; Olson, E. J.; Proskurowski, G.; Jakuba, M.; Bradley, A.; Larson, B.; Ludwig, K.; Glickson, D.; Buckman, K.; Bradley, A. S.; Brazelton, W. J.; Roe, K.; Elend, M. J.; Delacour, A.; Bernasconi, S. M.; Lilley, M. D.; Baross, J. A.; Summons, R. E.; Sylva, S. P. A Serpentinite-Hosted Ecosystem: The Lost City Hydrothermal Field. *Science* **2005**, *307* (5714), 1428–1434. <https://doi.org/10.1126/science.1102556>.
- (223) Knauth, L. P.; Lowe, D. R. High Archean Climatic Temperature Inferred from Oxygen Isotope Geochemistry of Cherts in the 3.5 Ga Swaziland Supergroup, South Africa. *GSA Bull.* **2003**, *115* (5), 566–580. [https://doi.org/10.1130/0016-7606\(2003\)115<0566:HACTIF>2.0.CO;2](https://doi.org/10.1130/0016-7606(2003)115<0566:HACTIF>2.0.CO;2).
- (224) Zhao, X.; Wang, H.; Zhao, R. G.; Yang, W. S. Self-Assembly of Amino Acids on the Cu(001) Surface. *Mater. Sci. Eng. C* **2001**, *1–2* (16), 41–50.

- (225) Rimola, A.; Sodupe, M.; Ugliengo, P. Amide and Peptide Bond Formation: Interplay between Strained Ring Defects and Silanol Groups at Amorphous Silica Surfaces. *J. Phys. Chem. C* **2016**, *120* (43), 24817–24826. <https://doi.org/10.1021/acs.jpcc.6b07945>.
- (226) Rimola, A.; Fabbiani, M.; Sodupe, M.; Ugliengo, P.; Martra, G. How Does Silica Catalyze the Amide Bond Formation under Dry Conditions? Role of Specific Surface Silanol Pairs. *ACS Catal.* **2018**, *8* (5), 4558–4568. <https://doi.org/10.1021/acscatal.7b03961>.
- (227) Steinberg, S.; Bada, J. L. Diketopiperazine Formation During Investigations of Amino Acid Racemization in Dipeptides. *Science* **1981**, *213* (4507), 544–545. <https://doi.org/10.1126/science.213.4507.544>.
- (228) Nagayama, M.; Takaoka, O.; Inomata, K.; Yamagata, Y. Diketopiperazine-Mediated Peptide Formation in Aqueous Solution. *Orig. Life Evol. Biosph.* **1990**, *20* (3), 249–257. <https://doi.org/10.1007/BF01808107>.
- (229) Bernstein, J. Polymorphism of L -Glutamic Acid: Decoding the  $\alpha$ - $\beta$  Phase Relationship *via* Graph-Set Analysis. *Acta Crystallogr. B* **1991**, *47* (6), 1004–1010. <https://doi.org/10.1107/S0108768191009345>.
- (230) Griffith, O. W.  $\beta$ -Amino Acids: Mammalian Metabolism and Utility as  $\alpha$ -Amino Acid Analogues. *Annu. Rev. Biochem.* **1986**, *55* (1), 855–878. <https://doi.org/10.1146/annurev.bi.55.070186.004231>.
- (231) Robinson, P.; Neelon, K.; Schreier, H. J.; Roberts, M. F.  $\beta$ -Glutamate as a Substrate for Glutamine Synthetase. *Appl. Environ. Microbiol.* **2001**, *67* (10), 4458–4463. <https://doi.org/10.1128/AEM.67.10.4458-4463.2001>.
- (232) Ruzicka, F. J.; Frey, P. A. Glutamate 2,3-Aminomutase: A New Member of the Radical SAM Superfamily of Enzymes. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **2007**, *1774* (2), 286–296. <https://doi.org/10.1016/j.bbapap.2006.11.008>.
- (233) Lyons, B.; Kwan, A. H.; Truscott, R. Spontaneous Cyclization of Polypeptides with a Penultimate Asp, Asn or IsoAsp at the N-Terminus and Implications for Cleavage by Aminopeptidase. *FEBS J.* **2014**, *281* (13), 2945–2955. <https://doi.org/10.1111/febs.12833>.
- (234) Geiger, T.; Clarke, S. Deamidation, Isomerization, and Racemization at Asparaginyl and Aspartyl Residues in Peptides. Succinimide-Linked Reactions That Contribute to Protein Degradation. *J. Biol. Chem.* **1987**, *262* (2), 785–794.
- (235) Athmer, L.; Kindrachuk, J.; Georges, F.; Napper, S. The Influence of Protein Structure on the Products Emerging from Succinimide Hydrolysis \*. *J. Biol. Chem.* **2002**, *277* (34), 30502–30507. <https://doi.org/10.1074/jbc.M205314200>.
- (236) Aswad, D. W.; Paranandi, M. V.; Schurter, B. T. Isoaspartate in Peptides and Proteins: Formation, Significance, and Analysis. *J. Pharm. Biomed. Anal.* **2000**, *21* (6), 1129–1136. [https://doi.org/10.1016/S0731-7085\(99\)00230-7](https://doi.org/10.1016/S0731-7085(99)00230-7).
- (237) Fox, S. W.; Harada, K. Thermal Copolymerization of Amino Acids to a Product Resembling Protein. *Science* **1958**, *128* (3333), 1214–1214. <https://doi.org/10.1126/science.128.3333.1214>.
- (238) Harada, K.; Fox, S. W. The Thermal Copolymerization of Aspartic Acid and Glutamic Acid. *Arch. Biochem. Biophys.* **1960**, *86* (2), 274–280. [https://doi.org/10.1016/0003-9861\(60\)90418-5](https://doi.org/10.1016/0003-9861(60)90418-5).
- (239) Cope, A. C.; Trumbull, E. R. Olefins from Amines: The Hofmann Elimination Reaction and Amine Oxide Pyrolysis. In *Organic Reactions*; American Cancer Society, 2011; pp 317–493. <https://doi.org/10.1002/0471264180.or011.05>.
- (240) Laerdahl, J. K.; Bache-Andreassen, L.; Uggerud, E. Nucleophilic Identity Substitution Reactions. The Reaction between Ammonia and Protonated Amines. *Org. Biomol. Chem.* **2003**, *1* (16), 2943–2950. <https://doi.org/10.1039/B302270F>.
- (241) Bojesen, G.; Breindahl, T. On the Proton Affinity of Some  $\alpha$ -Amino Acids and the Theory of the Kinetic Method. *J. Chem. Soc. Perkin Trans. 2* **1994**, No. 5, 1029–1037. <https://doi.org/10.1039/P29940001029>.
- (242) Hunter, E. P. L.; Lias, S. G. Evaluated Gas Phase Basicities and Proton Affinities of Molecules: An Update. *J. Phys. Chem. Ref. Data* **1998**, *27* (3), 413–656. <https://doi.org/10.1063/1.556018>.
- (243) Afonso, C.; Modeste, F.; Breton, P.; Fournier, F.; Tabet, J.-C. Proton Affinities of the Commonly Occurring L-Amino Acids by Using Electrospray Ionization-Ion Trap Mass Spectrometry. *Eur. J. Mass Spectrom.* **2000**, *6* (5), 443–449. <https://doi.org/10.1255/ejms.369>.
- (244) Sagadeev, E. V.; Gimadeev, A. A.; Barabanov, V. P. The Enthalpies of Formation and Sublimation of Amino Acids and Peptides. *Russ. J. Phys. Chem. A* **2010**, *84* (2), 209–214. <https://doi.org/10.1134/S0036024410020093>.
- (245) Miller, 1987, Which Organic Compounds Could Have Occurred on the Prebiotic Earth.Pdf.

- (246) Zhang, H.; Zhang, D.; Ray, K.; Zhu, M. Mass Defect Filter Technique and Its Applications to Drug Metabolite Identification by High-Resolution Mass Spectrometry. *J. Mass Spectrom. JMS* **2009**, *44* (7), 999–1016. <https://doi.org/10.1002/jms.1610>.
- (247) Sleno, L. The Use of Mass Defect in Modern Mass Spectrometry. *J. Mass Spectrom. JMS* **2012**, *47* (2), 226–236. <https://doi.org/10.1002/jms.2953>.
- (248) Kim, S.; Kramer, R. W.; Hatcher, P. G. Graphical Method for Analysis of Ultrahigh-Resolution Broadband Mass Spectra of Natural Organic Matter, the Van Krevelen Diagram. *Anal. Chem.* **2003**, *75* (20), 5336–5344. <https://doi.org/10.1021/ac034415p>.
- (249) Hertzog, J.; Carré, V.; Le Brech, Y.; Logan Mackay, C.; Dufour, A.; Mašek, O.; Aubriet, F. Combination of Electrospray Ionization, Atmospheric Pressure Photoionization and Laser Desorption Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry for the Investigation of Complex Mixtures – Application to the Petroleomic Analysis of Bio-Oils. *Anal. Chim. Acta* **2017**, *969*, 26–34. <https://doi.org/10.1016/j.aca.2017.03.022>.
- (250) Roullier-Gall, C.; Boutegrabet, L.; Gougeon, R. D.; Schmitt-Kopplin, P. A Grape and Wine Chemodiversity Comparison of Different Appellations in Burgundy: Vintage vs Terroir Effects. *Food Chem.* **2014**, *152*, 100–107. <https://doi.org/10.1016/j.foodchem.2013.11.056>.
- (251) Doerfler, H.; Sun, X.; Wang, L.; Engelmeier, D.; Lyon, D.; Weckwerth, W. MzGroupAnalyzer-Predicting Pathways and Novel Chemical Structures from Untargeted High-Throughput Metabolomics Data. *PLoS ONE* **2014**, *9* (5), e96188. <https://doi.org/10.1371/journal.pone.0096188>.
- (252) Maillard, J.; Carrasco, N.; Schmitz-Afonso, I.; Gautier, T.; Afonso, C. Comparison of Soluble and Insoluble Organic Matter in Analogues of Titan’s Aerosols. *Earth Planet. Sci. Lett.* **2018**, *495*, 185–191. <https://doi.org/10.1016/j.epsl.2018.05.014>.
- (253) Imanaka, H.; Smith, M. A. Formation of Nitrogenated Organic Aerosols in the Titan Upper Atmosphere. *Proc. Natl. Acad. Sci.* **2010**, *107* (28), 12423–12428. <https://doi.org/10.1073/pnas.0913353107>.
- (254) Gautier, T.; Carrasco, N.; Schmitz-Afonso, I.; Touboul, D.; Szopa, C.; Buch, A.; Pernot, P. Nitrogen Incorporation in Titan’s Tholins Inferred by High Resolution Orbitrap Mass Spectrometry and Gas Chromatography–Mass Spectrometry. *Earth Planet. Sci. Lett.* **2014**, *404*, 33–42. <https://doi.org/10.1016/j.epsl.2014.07.011>.
- (255) Pernot, P.; Carrasco, N.; Thissen, R.; Schmitz-Afonso, I. Tholinomics—Chemical Analysis of Nitrogen-Rich Polymers. *Anal. Chem.* **2010**, *82* (4), 1371–1380. <https://doi.org/10.1021/ac902458q>.
- (256) Marshall, A. G.; Rodgers, R. P. Petroleomics: The Next Grand Challenge for Chemical Analysis. *Acc. Chem. Res.* **2004**, *37* (1), 53–59. <https://doi.org/10.1021/ar020177t>.
- (257) Fouquet, T.; Sato, H. Improving the Resolution of Kendrick Mass Defect Analysis for Polymer Ions with Fractional Base Units. *Mass Spectrom.* **2017**, *6* (1), A0055–A0055. <https://doi.org/10.5702/massspectrometry.A0055>.
- (258) Svec, H. J.; Clyde, D. D. Vapor Pressures of Some  $\alpha$ -Amino Acids. *J. Chem. Eng. Data* **1965**, *10* (2), 151–152. <https://doi.org/10.1021/je60025a024>.
- (259) Bartel, D. P.; Szostak, J. W. Isolation of New Ribozymes from a Large Pool of Random Sequences [See Comment]. *Science* **1993**, *261* (5127), 1411–1418. <https://doi.org/10.1126/science.7690155>.
- (260) Kirschning, A. Coenzymes and Their Role in the Evolution of Life. *Angew. Chem. Int. Ed. n/a* (n/a). <https://doi.org/10.1002/anie.201914786>.
- (261) Hunt, D. F.; Yates, J. R.; Shabanowitz, J.; Winston, S.; Hauer, C. R. Protein Sequencing by Tandem Mass Spectrometry. *Proc. Natl. Acad. Sci.* **1986**, *83* (17), 6233–6237. <https://doi.org/10.1073/pnas.83.17.6233>.
- (262) Dongré, A. R.; Jones, J. L.; Somogyi, Á.; Wysocki, V. H. Influence of Peptide Composition, Gas-Phase Basicity, and Chemical Modification on Fragmentation Efficiency: Evidence for the Mobile Proton Model. *J. Am. Chem. Soc.* **1996**, *118* (35), 8365–8374. <https://doi.org/10.1021/ja9542193>.
- (263) Harrison, A. G.; Csizmadia, I. G.; Tang, T.-H. Structure and Fragmentation of B2 Ions in Peptide Mass Spectra. *J. Am. Soc. Mass Spectrom.* **2000**, *11* (5), 427–436. [https://doi.org/10.1016/S1044-0305\(00\)00104-5](https://doi.org/10.1016/S1044-0305(00)00104-5).
- (264) Schwartz, B. L.; Bursey, M. M. Some Proline Substituent Effects in the Tandem Mass Spectrum of Protonated Pentaalanine. *Biol. Mass Spectrom.* **1992**, *21* (2), 92–96. <https://doi.org/10.1002/bms.1200210206>.
- (265) Loo, J. A.; Edmonds, C. G.; Smith, R. D. Tandem Mass Spectrometry of Very Large Molecules. 2. Dissociation of Multiply Charged Proline-Containing Proteins from Electrospray Ionization. *Anal. Chem.* **1993**, *65* (4), 425–438. <https://doi.org/10.1021/ac00052a020>.

- (266) Goodlett, D. R.; Gale, D. C.; Guiles, S.; Crowther, J. B. Mass Spectrometry in Pharmaceutical Analysis. In *Encyclopedia of Analytical Chemistry*; American Cancer Society, 2006. <https://doi.org/10.1002/9780470027318.a1911>.
- (267) Mák, M.; Mezö, G.; Skribanek, Z.; Hudecz, F. Stability of Asp-Pro Bond under High and Low Energy Collision Induced Dissociation Conditions in the Immunodominant Epitope Region of Herpes Simplex Virion Glycoprotein D. *Rapid Commun. Mass Spectrom.* **1998**, *12* (13), 837–842. [https://doi.org/10.1002/\(SICI\)1097-0231\(19980715\)12:13<837::AID-RCM248>3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1097-0231(19980715)12:13<837::AID-RCM248>3.0.CO;2-Z).
- (268) Liu, C.; Topchiy, E.; Lehmann, T.; Basile, F. Characterization of the Dehydration Products Due to Thermal Decomposition of Peptides by Liquid Chromatography-Tandem Mass Spectrometry. *J. Mass Spectrom.* **2015**, *50* (3), 625–632. <https://doi.org/10.1002/jms.3570>.
- (269) Dookeran, N. N.; Yalcin, T.; Harrison, A. G. Fragmentation Reactions of Protonated  $\alpha$ -Amino Acids. *J. Mass Spectrom.* **1996**, *31* (5), 500–508. [https://doi.org/10.1002/\(SICI\)1096-9888\(199605\)31:5<500::AID-JMS327>3.0.CO;2-Q](https://doi.org/10.1002/(SICI)1096-9888(199605)31:5<500::AID-JMS327>3.0.CO;2-Q).
- (270) Pivcová, H.; Saudek, V.; Drobnik, H.  $^{13}\text{C}$  n.m.r. Study of the Structure of Poly(Aspartic Acid). *Polymer* **1982**, *23* (8), 1237–1241. [https://doi.org/10.1016/0032-3861\(82\)90386-X](https://doi.org/10.1016/0032-3861(82)90386-X).
- (271) Hurtado, P. P.; O'Connor, P. B. Differentiation of Isomeric Amino Acid Residues in Proteins and Peptides Using Mass Spectrometry. *Mass Spectrom. Rev.* **2012**, *31* (6), 609–625. <https://doi.org/10.1002/mas.20357>.
- (272) Sargaeva, N. P.; Lin, C.; O'Connor, P. B. Identification of Aspartic and Isoaspartic Acid Residues in Amyloid  $\beta$  Peptides, Including A $\beta$ 1–42, Using Electron-Ion Reactions. *Anal. Chem.* **2009**, *81* (23), 9778–9786. <https://doi.org/10.1021/ac901677t>.
- (273) Sargaeva, N. P.; Lin, C.; O'Connor, P. B. Unusual Fragmentation of  $\beta$ -Linked Peptides by ExD Tandem Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (3), 480–491. <https://doi.org/10.1021/jasms.8b03982>.
- (274) Li, X.; Lin, C.; O'Connor, P. B. Glutamine Deamidation: Differentiation of Glutamic Acid and  $\gamma$ -Glutamic Acid in Peptides by Electron Capture Dissociation. *Anal. Chem.* **2010**, *82* (9), 3606–3615. <https://doi.org/10.1021/ac9028467>.
- (275) Harrison, A. G.; Tu, Y.-P. Ion Chemistry of Protonated Aspartic Acid Derivatives. *J. Mass Spectrom.* **1998**, *33* (6), 532–542. [https://doi.org/10.1002/\(SICI\)1096-9888\(199806\)33:6<532::AID-JMS657>3.0.CO;2-U](https://doi.org/10.1002/(SICI)1096-9888(199806)33:6<532::AID-JMS657>3.0.CO;2-U).
- (276) González, L. J.; Shimizu, T.; Satomi, Y.; Betancourt, L.; Besada, V.; Padrón, G.; Orlando, R.; Shirasawa, T.; Shimonishi, Y.; Takao, T. Differentiating  $\alpha$ - and  $\beta$ -Aspartic Acids by Electrospray Ionization and Low-Energy Tandem Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2000**, *14* (22), 2092–2102. [https://doi.org/10.1002/1097-0231\(20001130\)14:22<2092::AID-RCM137>3.0.CO;2-V](https://doi.org/10.1002/1097-0231(20001130)14:22<2092::AID-RCM137>3.0.CO;2-V).
- (277) Harrison, Alex. G. Fragmentation Reactions of Protonated Peptides Containing Glutamine or Glutamic Acid. *J. Mass Spectrom.* **2003**, *38* (2), 174–187. <https://doi.org/10.1002/jms.427>.
- (278) Guo, Y.-C.; Cao, S.-X.; Zong, X.-K.; Liao, X.-C.; Zhao, Y.-F. ESI-MS<sup>n</sup> Study on the Fragmentation of Protonated Cyclic-Dipeptides. *Spectroscopy* **2009**, *23* (3–4), 131–139. <https://doi.org/10.1155/2009/580182>.
- (279) Andreotti, A. H. Native State Proline Isomerization: An Intrinsic Molecular Switch. *Biochemistry* **2003**, *42* (32), 9515–9524. <https://doi.org/10.1021/bi0350710>.
- (280) Nishikawa, T.; Hayashi, Y.; Suzuki, S.; Kubo, H.; Ohtani, H. Cis-Trans Isomerization of Proline Dipeptides during Liquid Chromatography: Kinetic Analysis of the Elution Profile. *Anal. Sci.* **1996**, *12* (4), 561–564. <https://doi.org/10.2116/analsci.12.561>.
- (281) Bouabdallah, S.; Ben Dhia, M. T.; Driss, M. R. Study of a Conformational Equilibrium of Lisinopril by HPLC, NMR, and DFT. *Int. J. Anal. Chem.* **2014**, *2014*, 1–8. <https://doi.org/10.1155/2014/494719>.





# Annexes

**Annexe 1. *ACS Earth Space Chem.* 2020, 4, 1802-1812**

# Origins of Life and Molecular Information: Selectivity in Mineral Surface-Induced Prebiotic Amino Acid Polymerization

Lise Bedoin, Sandra Alves,\* and Jean-François Lambert\*

Cite This: *ACS Earth Space Chem.* 2020, 4, 1802–1812

Read Online

ACCESS |



Metrics &amp; More



Article Recommendations

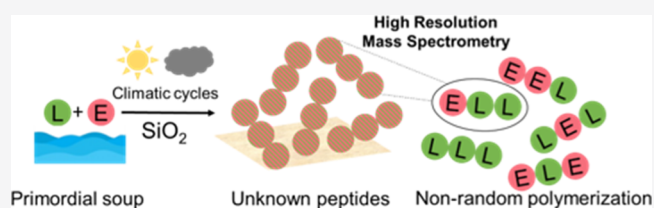


Supporting Information

**ABSTRACT:** In current living matter, biopolymers follow specific sequences that give them special properties, such as the sequence of amino acids (AAs) in proteins and peptides. A major challenge for the elucidation of the origins of life lies in understanding how, for example, nonrandom polypeptides have been selected among all the possible ones. While many investigations established plausible prebiotic polymerization pathways, surprisingly, only a few attempted to study the selectivity of these processes. We

studied a mineral surface polymerization scenario based on moderate thermal activation of leucine + glutamic acid mixtures on silica. Oligopeptides up to octamers were quantitatively formed in a “clean” prebiotic reaction and analyzed by high-resolution mass spectrometry and Fourier transform ion cyclotron resonance spectrometry for unambiguous molecular assignments. Nontrivial oligomerization selectivities are evidenced in both stoichiometric compositions and AA sequences, while comparable selectivities are not observed in other polymerization scenarios. This must therefore be due to specific catalytic reaction pathways occurring on the SiO<sub>2</sub> surface. A statistical measure of information contained in oligopeptide distributions is proposed. It could be used to follow the evolution of potentially meaningful complexity in biopolymers from the mineral to the biochemical world.

**KEYWORDS:** amino acids, peptides, origins of life, information, mass spectrometry, FT-ICR



## 1. INTRODUCTION

The study of the origins of life on the primitive Earth is a complex puzzle, involving several successive steps of matter organization. The emergence of simple biomolecules has been demonstrated over 60 years ago, when the famous Miller experiments,<sup>1</sup> recently reanalyzed,<sup>2</sup> showed that amino acids (AAs) could emerge through nonbiological chemical reactions. Alternatively, amino-acid blocks could also have been synthesized in interplanetary space and brought to Earth by meteorites.<sup>3</sup> The next step in the emergence of biochemistry-like complexity would have been their polymerization into oligopeptides, small-scale versions of proteins that can exhibit chemical functions, such as enzyme-like catalysis or self-replication,<sup>4–6</sup> even for relatively small-sized oligomers.<sup>7</sup> AA polymerization involves a condensation reaction between the amino and carboxylic groups of two AAs, forming a peptide bond. This reaction is endergonic and thus disfavored in water solution. Several alternative scenarios have been proposed for peptide bond formation in primitive Earth conditions, including the use of plausible sacrificial activators, allowing polymerization through *N*-carboxyanhydride chemistry.<sup>8</sup> The simplest scenarios, however, involve drying steps: removal of water, a reaction product, makes AA polymerization more favorable according to basic thermodynamic principles. Indeed, simple drying of aqueous solutions with moderate temperature activation has been shown to induce the formation of potentially useful peptides when applied to pure AAs or

mixtures of them.<sup>9,10</sup> However, in realistic geochemical settings, it is likely that drying would have occurred in the presence of minerals rather than in a bulk phase. Around 1950, Bernal<sup>11</sup> had already proposed the implication of mineral surfaces, presumably present on the primitive Earth, in AA polymerization. In addition to theoretical discussions,<sup>12,13</sup> experimental evidence of AA oligomerization on different surfaces was presented by many investigators,<sup>14–18</sup> including our previous works on high-surface fumed silica.<sup>19–22</sup> In the latter case, in addition to its thermodynamic effect on the polymerization reaction by providing a drying platform, the silica surface also has a kinetic effect as it helps AA condensation by mild acid catalysis, allowing the reaction to proceed at lower temperatures than in the bulk.<sup>19</sup>

Thus, surface AA polymerization is well-established, but strangely enough, some key questions relevant to the prebiotic potential of this scenario have hardly been investigated. Among them is the question of selectivity. Most of surface polymerization studies concern single AAs. Not only is this situation unlikely to have happened in realistic settings where AA

Received: July 9, 2020

Revised: August 15, 2020

Accepted: September 14, 2020

Published: September 15, 2020



synthesis pathways generally produce a mixture of several species, but even more importantly, if the next step is to be the development of function in oligopeptides, then their AA sequences are crucial in conferring on them whatever special properties they may have. In extant organisms, protein sequences are precisely coded in the genetic blueprint, with little tolerance for substitution. Even though there may have been a somewhat higher tolerance for variation in primitive functional oligopeptides, the probability of forming the “right” sequence in a random polymerization falls to astronomically low values very quickly when the oligomer length increases. As a consequence, the expected concentration of functional peptides would be very low, compromising the possibility of starting any “interesting” chemistry.

In the literature, the issues of polymerization selectivity in prebiotic oligopeptide synthesis and of the existence of nonrandom peptide sequences were only addressed in a few cases (somewhat more developed reflections are available for nucleotides polymerization<sup>23</sup>). In early studies, a nonrandom peptide oligomerization has been postulated under thermal activation of a mixed Glu + Asp solid phase.<sup>24,25</sup> Nakashima *et al.*<sup>26</sup> copolymerized a mixture of glycine, glutamic acid, and tyrosine and reported that among the eight possible tripeptide sequences, two (PyroGlu-Tyr-Gly and PyroGlu-Gly-Tyr) were largely predominant. In order to explain these experimental observations, Hartmann *et al.*<sup>27</sup> proposed a polymerization mechanism attempting to rationalize di- and tripeptides abundances in the system. These authors assumed that pyroglutamic acid reacts with cyclic dimers of the other two AAs (substituted diketopiperazines or DKPs for short) to specifically form two tripeptides. Mosqueira *et al.*<sup>28</sup> further elaborated on these results with a probabilistic model but did not provide new experimental data. Later on, Bujdák and Rode<sup>29</sup> studied surface condensation products from neutral hydrophobic AA couples. They observed a degree of polymerization selectivity, as Gly-X sequences were preferentially formed over X-Gly sequences when X = Leu or Val, whereas the reverse was observed when X = Pro or Ala. The authors explained such a nonrandom peptide synthesis by inductive effects of the AA side chains. Thus, while not absent from the literature and sometimes tantalizing, studies of AA polymerization selectivity remain few and often anecdotal. This overview would not be complete without mentioning the recent works by Krishnamurthy *et al.*<sup>30–32</sup> on depsipeptides, the products of co-condensation of  $\alpha$ -hydroxy acids with AAs on SiO<sub>2</sub> nanoparticle surfaces. In these more complicated systems, the authors established interesting selectivity patterns, including a preference for proteinaceous over nonproteinaceous AAs.

We undertook the present work in order to obtain hard data on the occurrence or not of selectivity in surface mineral-mediated AA polymerization. Nonporous fumed silica (SiO<sub>2</sub>) nanoparticles were chosen as the mineral phase because silica is known to constitute an efficient platform for simple heating-induced AA condensation. The simplest possible systems to study selectivity consist of a pair of different AAs: here, a mixture of leucine (Leu or L) and glutamic acid (Glu or E) was chosen. The choice of these AAs was dictated by several considerations. First, it is quite logical to start with two AAs with different intrinsic properties: Glu is hydrophilic and acidic, whereas Leu is hydrophobic. This has practical consequences on peptide structuring as alternating Glu–Leu peptides are known to produce well-defined peptide secondary

structures.<sup>33,34</sup> Second, upon single-step thermal activation, many AAs<sup>18,19,22,35</sup> or couples of AAs<sup>27</sup> will overwhelmingly produce cyclic dimers (DKPs), which are not at first sight a significant step forward toward complex biomolecules. It has been shown by several authors that cyclic dimers may be “recycled” to longer polymers by using wetting-and-drying cycles,<sup>36,37</sup> and these scenarios are highly interesting as pathways to complexity. However, they involve successive treatments that are rather difficult to analyze, and as a first step toward a mechanistic understanding of peptide formation, we focused the present study on a simple thermal activation procedure.

Obtaining long peptides instead of cyclic dipeptides involves a first level of reaction selectivity. In the literature, the formation of polypeptides up to 16-mers was observed on silica, using a vapor phase deposition method,<sup>38</sup> but such a scenario is less relevant for prebiotic chemistry than adsorption from an aqueous solution. Using the latter method, Sakhno *et al.* reported the production of linear oligopeptides up to hexamers in a single thermal activation step from two AA couples deposited on silica,<sup>39</sup> that is Glu + Leu and Asp + Val pairs, each system involving one trifunctional AA. However, the products of (Asp + Val) gave rise to practical difficulties in the analysis procedure (see **Materials and Methods**). Finally, in an origin-of-life perspective, both glutamic acid and leucine are reasonable choices. Even if they do not belong to the “short list” singled out in the “GAVD hypothesis”,<sup>40,41</sup> they are a part of a longer list of 10 or so AAs that are synthesized by prebiotic reactions and/or found in meteorites.<sup>42</sup>

In the present work, analysis of AA condensation products was performed on desorption solutions using mass spectrometry (MS). Commonly used in proteomics, MS, in particular Fourier transform ion cyclotron resonance (FT-ICR) spectrometry, appears as the most suitable analytical technique; thanks to its selectivity, sensitivity, and specificity to characterize mixtures as complex as our condensation products. Especially, FT-ICR is an extremely powerful MS device providing high analytical performances in terms of mass resolution, mass measurement accuracy, and wide dynamic range.<sup>43</sup> Thus, the present study also aims to demonstrate the potential of an FT-ICR instrument equipped with a dynamically harmonized cell for a prebiotic chemistry application. The use of accurate mass measurements demonstrates the formation of nonrandom oligopeptide sequences.

## 2. MATERIALS AND METHODS

The mineral support used for the adsorption of AAs was the fumed silica (SiO<sub>2</sub>) Aerosil 380 provided by Evonik Industries (Hanau, Germany), with a Brunauer–Emmett–Teller surface area of 380 m<sup>2</sup>/g. L-Leucine (L) and L-Glutamic acid (E) were purchased from Sigma Aldrich (Steinheim, Germany). The tripeptides LEL, ELL, ELE, LEE, and LLE were synthesized by the Institut de Biologie Paris Seine (Sorbonne Université, Paris), and LLL, EEE, and EEL were purchased from Bachem (Bubendorf, Switzerland).

Single AAs and/or mixtures of both were deposited on silica from aqueous solutions using the “incipient wetness impregnation” method, derived from the field of supported catalyst synthesis: the required amounts of the AAs were dissolved in water, and the resulting solution was then added to silica in a ratio of 10 mL for 1 g of silica. This resulted in a slurry without a separate liquid phase, which was then dried under air flow at

room temperature overnight. A reference sample was prepared with 1.25% Leu and 1.4% Glu (all loadings are in weight percent), corresponding to a Leu/Glu molar ratio of 1:1. In this communication, “Leu + Glu/SiO<sub>2</sub>”, without further specification, refers to the reference sample. The reference loading corresponds to 0.15 Leu and 0.15 Glu molecules per nm<sup>2</sup>. This would be about one tenth of a physical monolayer based on the sizes of the two AAs: the AAs are then relatively “scarce” on the surface.

Additionally, 2.5% Leu + 1.4% Glu/SiO<sub>2</sub> and 6.25% Leu + 1.4% Glu/SiO<sub>2</sub>, corresponding to 2:1 and 5:1 Leu/Glu ratios, and single AAs 1.25% Leu/SiO<sub>2</sub> and 1.4% Glu/SiO<sub>2</sub> samples were also prepared. They will be denoted as “Leu + Glu/SiO<sub>2</sub> 2:1” and “Leu + Glu/SiO<sub>2</sub> 5:1”.

Thermal activation of the AA(s)/silica systems was carried out in a U-shaped cell under argon flow, placed in a tubular oven controlled by a temperature programmer. A 1 °C/min ramp was used to reach a final value of 160 °C, where a plateau was maintained for 30 min. In order to obtain reproducible assays, the linearity of the temperature ramp must be carefully controlled. The resulting powders were stored in a desiccator.

The Leu + Glu/SiO<sub>2</sub> samples were analyzed by different methods:

Thermogravimetric analyses (TGA) of freshly dried samples (not thermally activated) were recorded using a TA instrument with a STD Q600 analyzer. Analyses were performed with a heating rate of 1 °C/min up to 200 °C, under a 100 mL/min dry air flow. The use of air rather than Ar (as in the thermal activation procedure) is a common practice in the TG study of AA/SiO<sub>2</sub> systems; it has generally been observed that the reactivity of O<sub>2</sub> does not influence the AA polymerization reactions that are the focus of the present study (although O<sub>2</sub> definitely influences the decomposition reactions at higher temperatures).

Infrared (IR) spectroscopic characterization was carried out on a Nicolet Magna-IR 550 spectrometer equipped with an MCT detector, with nominal 4 cm<sup>-1</sup> resolution. Self-supporting pellets were prepared by pressing the sample in a powder form under a pressure of 8 tons/cm<sup>2</sup>. The pellets were then placed in the FTIR sample compartment under N<sub>2</sub> flow. 128 scans were accumulated in the 4000–400 cm<sup>-1</sup> spectral domain. IR spectra were normalized to the 1870 cm<sup>-1</sup> silica band. This method was applied both before and after thermal treatment.

X-ray diffraction (XRD) analysis was used to confirm that AAs are well-dispersed on the silica surface: we mostly checked that the diffractogram only showed a broad diffusion signal due to amorphous silica, without sharp peaks that would indicate the precipitation of bulk AA crystals (see Figure S1a). Diffractograms were recorded on a Bruker D8 diffractometer using Cu K $\alpha$  radiation. An amorphous plastic sample holder was used. The 2 $\theta$  range was set from 10 to 90°, with a number of steps of 3810 and an accumulation time of 1 s per step.

In order to carry out MS analyses, thermally activated samples were submitted to a desorption procedure. 10 mg of solid samples was dispersed in 1 mL of Milli-Q water, followed by centrifugation during 5 min at 6000 rpm. The study of three successive desorption steps showed that there are no significant differences between the compositions of the first and successive desorption solutions, only a global dilution. The analysis of complex polypeptide mixtures can hardly be accomplished *in situ* in the adsorbed phase, so efficient desorption is a necessary requirement for quantitative analysis

of the reaction products. In the related Asp + Val/SiO<sub>2</sub> system, also studied by Sakhno *et al.*<sup>39</sup> as already mentioned, some of the reaction products had a very high affinity for the surface and hardly desorbed at all, so the composition of the desorption solution did not faithfully reflect that of the condensation products. The importance of peptide desorption was underlined, from another point of view, by de Duve and Miller.<sup>12</sup>

In order to better prove the specificities of the silica surface scenario with respect to other AA polymerization methods, peptide condensation was also performed in solution using a sacrificial activator. More precisely, amide bond formation was carried out using carbonyldiimidazole (CDI). 1.5 mM L-Leu, 1.5 mM L-Glu, and 10 mM KCl were dissolved in 10 mL of water. 6 mM CDI was then added, and the solution was incubated at room temperature for 24 h. The mechanism of CDI-induced polymerization of AAs was described by Ehler.<sup>44</sup>

The products of both the silica surface and CDI condensation scenarios were analyzed by MS. Direct infusion MS experiments were performed in the positive electrospray ionization (ESI) mode on a QhFT-ICR 7 T instrument (solariX XR, Bruker Daltonics, Germany) equipped with a dynamically harmonized cell. All mass spectra were acquired over a mass range of  $m/z$  100–2000. Each sample was diluted 50 times in 50:50 (v/v) water/acetonitrile (containing 0.1% v/v of CH<sub>3</sub>COOH) to reach a total concentration of approximately 5  $\mu\text{g}\cdot\text{mL}^{-1}$  (supposing that the desorption step is quantitative). A preliminary study was carried out to determine the optimal sample dilution. A dilution factor of 50 allows sufficient peak intensities while avoiding space charge phenomena inside the trapping device, which may cause mass measurement shifts.<sup>45</sup> The resulting desorption solution was directly infused at a flow rate of 180  $\mu\text{L}\cdot\text{min}^{-1}$  into the ESI source. The following conditions were used: a time of flight of 1.2 ms, ion accumulation time of 0.2 s, drying gas flow of 2 L/min at 200 °C, and nebulizer gas pressure of 1.5 bar. A capillary voltage of 4.5 kV was applied. The method was calibrated using a solution of acetate and sodium formate clusters. MS/MS experiments were performed using collision-induced dissociation (CID) activation in the hexapole collision cell in order to study the fragmentation patterns of condensation products, that is, synthesized oligopeptides. The isolation window of precursor ions was about 1  $m/z$  unit, and the collisional energy was optimized depending on precursor ion relative intensity (6–12 V).

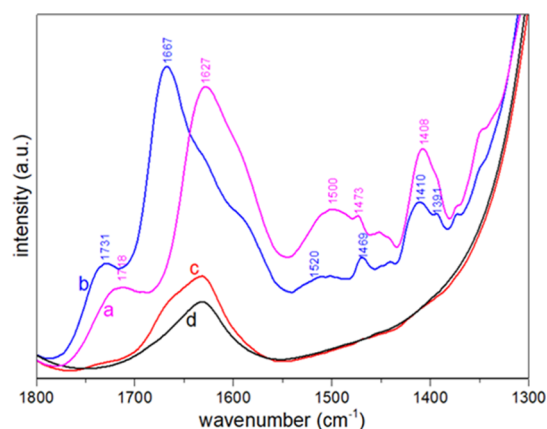
### 3. RESULTS

**3.1. Evidence of Peptide Bond Formation under Thermal Activation.** First, direct analysis of the solid Leu + Glu/SiO<sub>2</sub> system was performed using TGA in order to observe water emission events that may be assigned to peptide bond formation under thermal activation conditions (Figure S1b in Supporting Information). The TGA traces of Leu + Glu/SiO<sub>2</sub> systems show two endothermic events at 115 and 148 °C. As previously demonstrated in the studies of AA/SiO<sub>2</sub> systems,<sup>21,22,46</sup> peaks in this range can be assigned to condensation reactions between amine and carboxylic acid functions, forming amide bonds and accompanied by H<sub>2</sub>O loss. For most of the AAs, a single amide condensation event is observed in this region. The existence of two separate events is typical of systems containing some trifunctional AAs such as glutamic acid since in addition to peptide condensation, they can also give rise to intramolecular condensation with



formation of a lactam group;<sup>21,39</sup> in the case of glutamic acid, this leads to the formation of pyroglutamic residues. The TG results are the basis for our selection of an activation temperature of 160 °C since the water evolution events are completed at this temperature.

IR spectroscopy was used to follow peptide bond formation. Figure 1 compares the IR spectra of the Leu + Glu/SiO<sub>2</sub>



**Figure 1.** Normalized IR spectra in the 1800–1300 cm<sup>-1</sup> range of the Leu + Glu/SiO<sub>2</sub> system (a) before and (b) after thermal activation, (c) of the remaining solid material after water desorption, and (d) of the raw SiO<sub>2</sub> support.

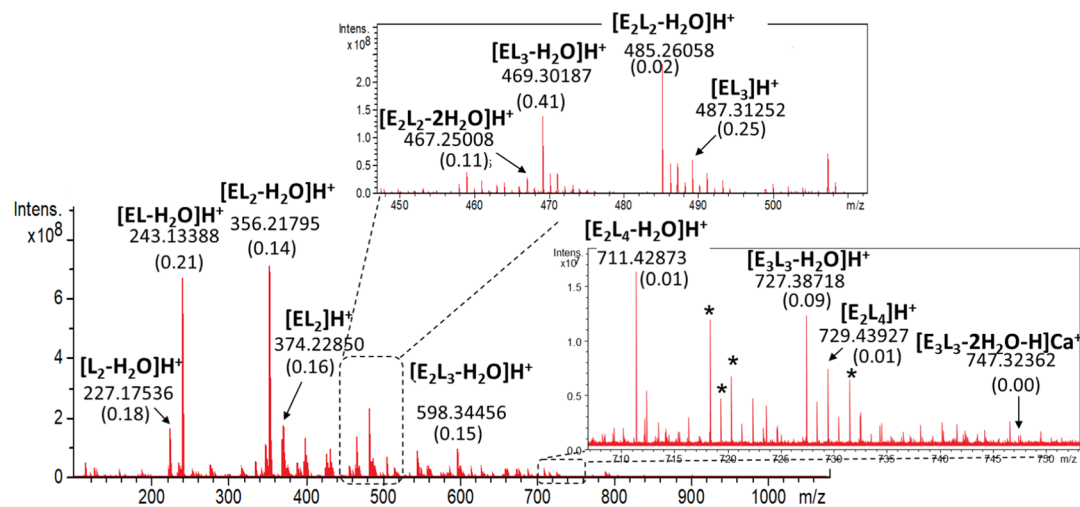
system before and after thermal activation and finally after applying the desorption step for AA and oligomer extraction prior to MS analysis. The observed bands for the unactivated samples are in keeping with previously published data, assuming the existence on the surface of AAs in the form of monomeric zwitterions (with the side chain COOH of Glu unionized). Band assignments are given in Table S1. Changes occur between Leu + Glu/SiO<sub>2</sub> IR spectra before and after thermal activation with the appearance of additional bands. The two bands at 1667 cm<sup>-1</sup>, assigned to amide I (mostly due to C=O stretching vibration), and 1520 cm<sup>-1</sup> (corresponding to amide II vibration) reveal peptide bond formation. Moreover, the presence of the amide II band suggests the formation of linear peptides<sup>39</sup> since this band is not apparent in

cyclic dipeptides (DKPs) that are the chief polymerization products in many AA/SiO<sub>2</sub> systems.<sup>22</sup>

In addition, IR results give an estimation of the desorption step efficiency.<sup>35</sup> The integrated intensity of the bands due to the organic molecules on the surface can indeed be estimated by subtracting the spectrum of raw silica support from that of the sample, after proper normalization (see Materials and Methods section). Applying this procedure to the IR spectra in Figure 1b,c reveals that the desorbed sample (spectrum c) retains less than 5% of the intensity of “organic bands” as compared to the same sample before desorption (spectrum b). Since thermal reaction products were almost quantitatively desorbed, their characterization could be undertaken by high-resolution mass spectrometry (HRMS) analysis of desorption solutions. Based on ppm mass accuracy measurements, oligopeptide-derived ions were identified, providing unambiguous molecular formulas of singly charged oligopeptides.

We first studied the condensation products of single AAs on silica, as references for comparison with the binary system Leu + Glu/SiO<sub>2</sub>. The ESI mass spectra from desorption solution of thermally activated Leu/SiO<sub>2</sub> and Glu/SiO<sub>2</sub> are reported in Figure S2 (Supporting Information). There is little or no degradation to nonpeptidic molecules and only little formation of long polymers. From Leu/SiO<sub>2</sub> analysis, the predominant peak, corresponding to [L<sub>2</sub>-H<sub>2</sub>O]<sup>+</sup>H<sup>+</sup> ions, can be assigned to a cyclic dimer cyclo(L-L), in other words, a substituted DKP, in accordance with the previous work.<sup>37</sup> At first sight, smaller amounts of trimers and tetramers could be present, but relevant peaks could as well be assigned to noncovalent complexes.

Regarding Glu/SiO<sub>2</sub>, no untransformed monomer is left after thermal activation but the product of internal condensation, PyroGlu, is detected. In addition, several Glu-containing oligopeptides seem to be present in accordance with the previous work.<sup>21</sup> However, the same ambiguity as for Leu/SiO<sub>2</sub> is encountered since they could also be attributed to noncovalent complexes. In this case, we performed specific experiments to discriminate between these two possibilities (see Figures S3 and S4), finding that noncovalent complex formation in the conditions of analysis accounts well for the observations. In summary, neither Glu/SiO<sub>2</sub> nor Leu/SiO<sub>2</sub> gives rise to the formation of long peptides.



**Figure 2.** Positive ESI mass spectrum of the desorption solution obtained from a thermally activated Glu + Leu/SiO<sub>2</sub> system using HRMS detection. The ppm accuracy of *m/z* value measurements are shown in brackets. Peaks marked with a star correspond to unidentified compounds.

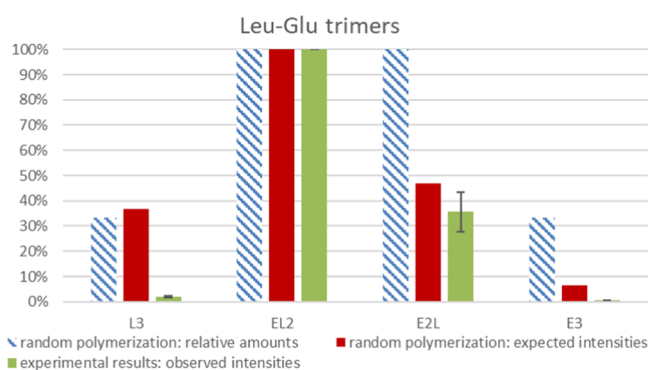
Previous spectroscopic and TGA experiments<sup>19,21,22,47</sup> suggested that thermal activation of a single AA on a silica surface does not usually produce linear oligopeptides when deposition is carried out from an aqueous solution.

The HRMS analysis of the desorption solution from activated Leu + Glu/SiO<sub>2</sub> shows a starkly different picture. Figure 2 presents the whole ESI mass spectrum of the desorption solution from activated Leu + Glu/SiO<sub>2</sub>, with close-up views of the tetrapeptide and hexapeptide mass regions shown in the insets. A vast majority of detected species may indeed be attributed to peptides. Abundant protonated and cationized ions from homo- and hetero-oligopeptides are detected, which proves the occurrence of oligopeptide formation (further confirmed by fragmentation patterns using MS/MS experiments, see below). Oligomer ions up to octamers have been detected with a mass accuracy better than 1 ppm. Thus, as previously reported,<sup>39</sup> the copolymerization of the AA pair (E and L) on silica leads to extensive formation of long oligopeptides. On the contrary, this was not significantly achieved for either AA when deposited alone. The four predominant peaks were assigned to stoichiometries [EL–H<sub>2</sub>O]H<sup>+</sup>, [EL<sub>2</sub>–H<sub>2</sub>O]H<sup>+</sup>, [EL<sub>2</sub>]H<sup>+</sup>, and [E<sub>2</sub>L<sub>2</sub>–H<sub>2</sub>O]H<sup>+</sup>, respectively. Other oligopeptide species including [L<sub>2</sub>–H<sub>2</sub>O]H<sup>+</sup>, [E<sub>2</sub>L–H<sub>2</sub>O]H<sup>+</sup>, and [E<sub>2</sub>L<sub>3</sub>–H<sub>2</sub>O]H<sup>+</sup> are also detected with quite high abundances. Interestingly, the most abundant oligopeptide species correspond to dehydrated forms (denoted as “–H<sub>2</sub>O”, *i.e.*, –18.01056 *uma* with respect to the calculated formula for a linear peptide), with the exception of the EL<sub>2</sub> tripeptide.

In order to assess the reproducibility and check the possible effect of sample aging, several HRMS analyses were performed on solutions desorbed from the activated Leu + Glu/SiO<sub>2</sub> samples after variable solid sample storage durations, either immediately after desorption or after different freezing–thawing cycles. Similar MS profiles, with comparable relative intensities, were obtained (data not shown), indicating that the conditions of sample storage do not modify the composition of the oligopeptide mixture.

**3.2. Polymerization Selectivity: Peptide Stoichiometry.** The main goal of our study was to investigate the possible existence of polymerization selectivity during peptide formation in the scenario of thermal activation on a silica surface. Caution is needed however since discrimination effects exist in MS, which strongly impact the quantification of analytes from MS. These effects stem from the phenomena occurring in different parts of MS instrumentation. In addition to ionization discrimination effects (see following paragraph), mass discrimination phenomena exist in the FTICR mass analyzer device characterized by a pseudopotential well.<sup>43</sup> So, relative abundances of detected oligomer ions using a broad 100–1000 *m/z* detection range cannot be used for a quantitative evaluation of all oligomers: only comparisons within a limited *m/z* range can be meaningful.

The first aspect of the selectivity question is whether some stoichiometries are preferred among polypeptides of a given length. Taking tripeptides as an example, the possible stoichiometries are L<sub>3</sub>, EL<sub>2</sub>, E<sub>2</sub>L, and E<sub>3</sub>. Assuming a random polymerization, with two AAs present in equimolar amounts, there would be a probability of 0.5 for each position in the chain to be occupied by Leu or Glu, respectively, so that the respective molar ratios of these different peptides would be 1:3:3:1 (the binomial coefficients) or 1/3:1:1:1/3, normalizing the highest amounts to 1. Figure 3 compares the theoretically



**Figure 3.** Comparison of the theoretically expected peak intensities for tripeptides in random polymerization to the experimentally observed peak intensities. Relative amounts in the random scenario are corrected for ionization efficiencies to give “expected intensities”. All the distributions are normalized with respect to EL<sub>2</sub>, arbitrarily set at 1.

expected tripeptide abundances to the experimentally observed peak intensities for our reference sample (a L/E molar ratio of 1:1). It must be underlined that to obtain the experimental values, the peak intensities of all the oligopeptide ions observed for a given stoichiometry (protonated or cationized species, from dehydrated or fully hydrated forms) have been summed. Average values from seven repeated analyses are plotted together with error bars corresponding to standard deviations.

At the first sight, the experimental data would seem to be strongly different from the expected amounts for random polymerization, suggesting a definite selectivity. However, discrimination effects arise at the ionization step, particularly when using atmospheric pressure ionization methods.<sup>48</sup> Variable ionization efficiencies are expected for different oligopeptide sequences. It is not straightforward to evaluate these MS discrimination effects quantitatively; typically, quantitative studies involve the use of reference labeled compounds, and developed methods are generally compound- and compound family-dependent.<sup>49</sup> To estimate such phenomena in the present work, reference tripeptides were used. ESI mass spectra of equimolar mixtures containing three reference tripeptides (LLL and EEE together with one of the six other heterotripeptides) were analyzed using the same procedure; the results are reported in Figure S5. In general, leucine-rich peptides display a higher ionization efficiency; in the most extreme case, the peak intensity of [L<sub>3</sub>]H<sup>+</sup> is higher than that of [E<sub>3</sub>]H<sup>+</sup> by almost an order of magnitude. This was expected, as analytes with hydrophobic moieties should have a higher ESI response than more polar ones due to their high affinity for electrospray droplet surfaces.<sup>48</sup> They tend to have a greater charge excess which may explain their greater ionization yields. Thus, the ionization efficiencies follow a relative order that depends on the global stoichiometry: L<sub>3</sub> > EL<sub>2</sub> > E<sub>2</sub>L > E<sub>3</sub>. Tripeptides of the same stoichiometry but with different sequences exhibit close MS response. Based on these results, we calculated an average MS response factor for the two stoichiometries having several possible sequences (*i.e.* EL<sub>2</sub> and E<sub>2</sub>L) and used them to evaluate the expected peak intensities that would correspond to random polymerization (Figure 3). These calculations remain approximate because they do not take into account possible differences in respective ionization efficiency between protonated and cationized species, nor between fully hydrated and dehydrated peptides



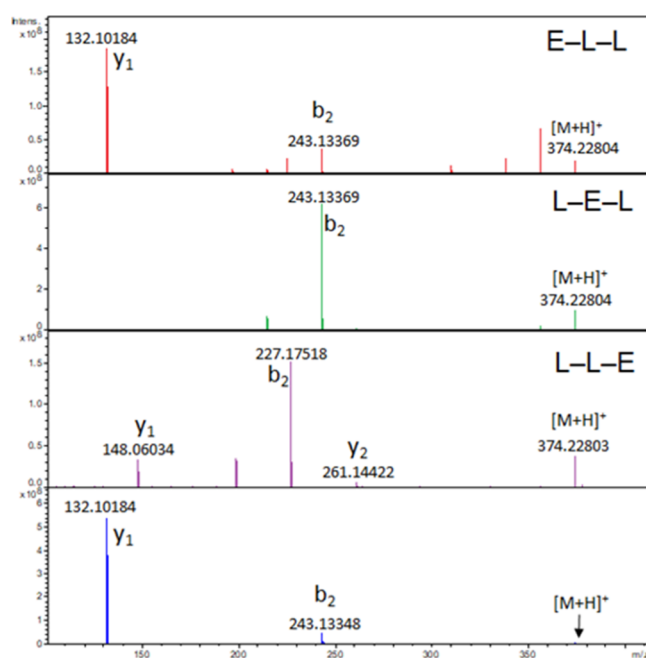
from the same peptide sequence, but they should represent an improvement as compared to raw peak intensities.

The discrepancy between (corrected) random and experimental distributions is reduced but not eliminated; with regard to the relative amounts of peptides with  $EL_2$  and  $E_2L$  stoichiometries, there is about 90% chance that the predominance of  $EL_2$  over  $E_2L$  is significant. In contrast, the quasi absence of species with stoichiometry  $L_3$  is made more conspicuous because they have a high response factor and their signal should be exalted if they were present.  $E_3$  species also appear definitely under-represented. Comparisons between the random model and experimental data are reported for other peptide lengths and also for other E/L molar ratios in Figures S6–S8. The lack of suitable reference peptides for chains longer than tripeptides did not permit even the rough quantifications that we attempted for tripeptides—not to mention the exponentially more complicated difficulty of the task when peptide length increases. Nevertheless, experimental distributions do seem to be significantly different from random ones. Especially, homopolymers seem to be very scarce. The verification of this observation was one of the reasons to investigate AA mixtures of different stoichiometries (see below). Note that the latter results are somewhat different from previous ESI results on the same system, where  $EL_4$  and  $EL_5$  species were the only detected stoichiometries for pentamers and hexamers.<sup>39</sup> In the latter study, however, less care was taken to control the residual pressure of water during thermal treatment: the effect of water activity on the polymerization process is a direction for further research.

**3.3. Polymerization Selectivity: Peptide Sequence Determination.** The next question that must be raised relates to the existence of sequence selectivity for a given peptide stoichiometry. Peptides with the same AA composition but different sequences give peaks at the same  $m/z$  values in MS, but tandem mass spectrometry (MS/MS) should provide access to a unique peptide sequence through the detection of sequence-specific series of fragment ions and more specifically the so-called b/y fragments.<sup>50,51</sup> The protonated molecular ion for a given peptide stoichiometry is selected as a precursor ion and undergoes CID fragmentations. Figure 4 presents the CID spectra of the protonated ELL, LEL, and LLE reference tripeptides and one of the peaks detected experimentally at  $m/z$  374.228 from the desorbed solution from thermally activated Leu + Glu/SiO<sub>2</sub>. The MS/MS spectrum of the precursor ion from the desorption solution displays only two major product ions at  $m/z$  243.13348 (assigned to a b<sub>2</sub> fragment) and at 132.10175 (assigned to a y<sub>1</sub> fragment). This fragmentation pattern suggests the predominant presence of the ELL sequence among tripeptides with the  $EL_2$  stoichiometry, thus demonstrating a selectivity in the sequence of the oligopeptides formed.

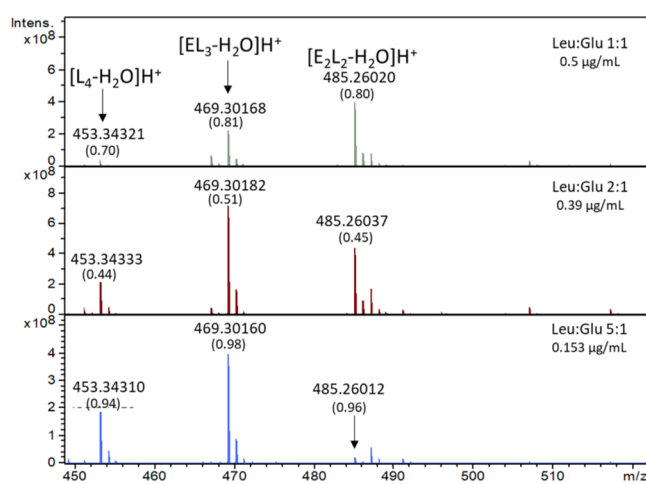
MS/MS experiments were also performed for other detected linear and dehydrated oligopeptides ions (see Figure S9 in Supporting Information). Conclusive interpretations are hard to make because of the lack of reference peptides to obtain characteristic fragmentation patterns of genuine compounds and because of the very similar fragmentation pathways for alternating Leu/Glu heteropeptides, but especially in the case of tetrapeptides, there are definite hints of sequence selectivity. To go further, an MS<sup>n</sup> study is under investigation.

**3.4. Effect of the AA Molar Ratio on Polymerization.** The effect of the molar ratio between two AAs on oligomer production was studied using the same HRMS conditions.



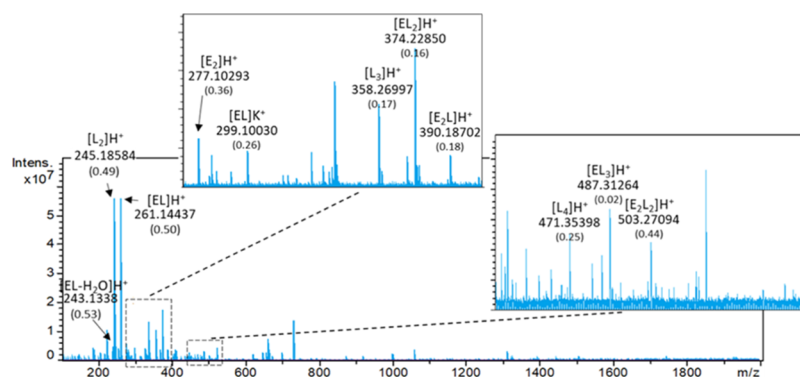
**Figure 4.** MS/MS spectra of the protonated molecule  $[M + H]^+$  at  $m/z$  374.2280 for reference tripeptides ELL, LEL, and LLE sequences and for the peak at  $m/z$  374.2280 detected from the thermally activated Leu + Glu/SiO<sub>2</sub> system (bottom spectrum).

Figure 5 shows a zoomed  $m/z$  range corresponding to the tetrapeptide ions for the three different Leu/Glu ratios



**Figure 5.** Positive ESI mass spectra in the tetrapeptide ( $m/z$  450–520) mass range of thermally activated Leu + Glu/SiO<sub>2</sub> systems prepared with different Leu:Glu ratios. Note that the sample concentrations were optimized to obtain similar MS response and avoid saturation.

investigated. Comparisons of theoretically expected tripeptide abundances to the experimentally observed peak intensities for the different oligomer lengths and three different L/E molar ratios are reported in Figures S6–S8. As expected, leucine-rich peptides are formed in higher amounts when starting from mixtures containing higher proportion of leucine. For example, the  $[L_4-H_2O]^+$  ion relative intensity ( $m/z$  453.34353) increases for 2:1 and 5:1 ratios. Similarly, the  $[E_2L_2-H_2O]^+$  ion relative intensity, ( $m/z$  485.26059) as compared to the  $[EL_3-H_2O]^+$  ion, decreases, whereas the  $[E_4-H_2O]^+$



**Figure 6.** Positive ESI mass spectrum of the condensation products obtained in solution using CDI as a sacrificial activator.

peptide is never detected. However, some polymerization selectivity is still observed. In particular, heteropeptide ions are always predominant, even when an excess of Leu should favor the  $(\text{Leu})_n$  homopolymers if the polymerization was random. Regarding the tripeptides, for which a semiquantitative comparison is possible, the trends discussed above for the 1:1 ratio are confirmed: in addition to very small amounts of  $L_3$  even for a L/E ratio of 5:1, the predominance of  $EL_2$  over  $E_2L$  is confirmed (Figure S6 in Supporting Information) and indeed more conspicuous than for the 1:1 ratio. If we look at the data for the three ratios, we see that the following oligomers seem over-represented, with respect to the random scenario:  $EL_3$  and  $E_2L_2$  for the tetramers,  $E_2L_3$  for the pentamers,  $E_2L_4$  for the hexamers,  $E_2L_5$  and  $E_3L_4$  for the heptamers, and  $E_3L_5$  for the octamers (Figures S6–S8).

**3.5. Origin of Polymerization Selectivity: Comparison with Another Condensation Method.** The observed selectivity could have a thermodynamic origin, meaning that some oligomers are formed in higher amounts because they are thermodynamically more stable, or of a kinetic origin, meaning that the particular reaction mechanism on the silica surface specifically opens certain transformation pathways by lowering their activation energy (heterogeneous catalysis). In order to shed some light on this question, a homogeneous condensation procedure using a sacrificial activator, that is, CDI, was applied to an equimolar Leu + Glu mixture. The ESI mass spectrum of the condensation products is reported in Figure 6. Oligopeptides up to pentamers are detected, but the results diverge from those obtained from surface polymerization, in two important respects.

First, little to no stoichiometry selectivity is obvious here. Indeed, the experimentally observed distributions are rather close to the random oligomerization once the relative ionization yields between Leu- and Glu-rich peptides are taken into account (see the di-, tri-, and tetra-peptide mass ranges in Figure 6), strongly suggesting that the CDI method is less selective and closer to random polymer formation. These observations underline the key role of the silica surface in polymerization selectivity. As a matter of fact, they do not definitely prove that the selectivity observed in the silica surface scenario has a kinetic origin since it may be argued that various oligomers have different stabilities on the silica surface due to different amino-acid/surface interaction energies. We will examine this question in more detail in a forthcoming publication dealing with reaction mechanisms and kinetics. Suffice it to say here that the peptide distributions discussed above may not correspond to an equilibrium situation and that

they vary significantly with the duration of the thermal treatment.

Second, no dehydrated peptide ions are detected when using the CDI procedure, even for Glu-containing oligopeptides. The absence of dehydrated peptidic forms when using a sacrificial activator in contrast to surface activation is of major interest.

#### 4. DISCUSSION

Let us recapitulate the different aspects of condensation reaction selectivity that are relevant for the origin-of-life significance of AA polymerization on mineral surfaces.

First, there is the overall yield of peptide formation. From a thermodynamic point of view, the AA condensation equilibrium can be displaced as much to the right as desired, for all practical purposes close to 100%, simply by decreasing the water activity (harsh drying conditions). From a kinetic point of view, many early studies<sup>52</sup> were carried out in temperature conditions where the reaction is rather slow and takes weeks to be completed. Although the yields reported in those low  $T$  studies may be only a few percent, TG results show that quantitative condensation can be achieved in a matter of minutes at 160 °C—although the peptide bonds will keep rearranging with time. Then, HRMS analysis shows that most observable peaks are attributable to singly charged oligopeptides. Actually, from previous studies, we know that various peptide degradation reactions may be induced by heating at higher temperatures, but on the silica surface, they do not become significant before at least 250 °C.<sup>19</sup> Therefore, AA condensation on mineral surfaces is a robust scenario, where close to 100% selectivity to the “prebiotically interesting” compounds is achieved and maintained over a large range of experimental conditions. This deserves to be underlined since many reported prebiotic pathways do form molecules of biological significance, but only in a few % yields and/or in conditions where they are quickly degraded.

Second, there is the question of selectivity for rather long oligopeptides over cyclic dipeptides (DKPs). This point was mentioned in the introduction, as one of the reasons for choosing the (Leu + Glu) AA pair, in view of the previous work<sup>39</sup> where this system did display this type of selectivity. We have confirmed that peptides containing three or more AA residues are indeed predominant, while neither single AA systems (Leu/SiO<sub>2</sub> or Glu/SiO<sub>2</sub>) form large amounts of them. Actually, it had been shown by Fox and Harada<sup>9</sup> that mixed systems containing Glu could form long polypeptide chains, although this was for a bulk polymerization scenario (no

**Table 1. Calculation of the Information Content of the Experimental Tripeptide Distribution in the Silica Surface and the CDI Activator Scenarios, as Compared to a Random Distribution<sup>a</sup>**

stoichiometry $i$	$L_3$	$EL_2$	$E_2L$	$E_3$	
$p_i$ (random)	0.125	0.375	0.375	0.125	
$p_i \log(p_i)$ (random)	-0.260	-0.368	-0.368	-0.260	$H_{\max}(X) = -\sum_{i=1} p_i \log p_i$ (rnd) = 1.256
$p_i$ (exp, SiO <sub>2</sub> )	0.010	0.570	0.407	0.013	
$p_i \log(p_i)$ (exp, SiO <sub>2</sub> )	-0.045	-0.321	-0.366	-0.057	$H(X SiO_2) = -\sum_{i=1} p_i \log p_i$ (obs) = 0.788
$p_i$ (exp, CDI)	0.300	0.495	0.205	0.000	
$p_i \log(p_i)$ (exp, CDI)	-0.361	-0.348	-0.325	0.000	$H(X CDI) = -\sum_{i=1} p_i \log p_i$ (obs) = 1.034

<sup>a</sup>Natural logarithms were used; conversion to base 2 log may be obtained multiplying by 1.4427.

surface involved) that is less likely prebiotic. Later on, the same authors established that the necessary condition for making long polypeptides in their scenario was the presence in sufficient amounts of trifunctional AAs having either a -COOH (Glu, Asp) or a -NH<sub>2</sub> group (Lys) in their side chain.<sup>24,25</sup> In the present study, the most abundant oligopeptide ions were tri- and tetramers, with a significant presence of higher polymers up to the octamers. This may seem small in comparison with the bulk polymerization scenario where peptides containing up to 100 residues were present. However, the length of chains formed on surfaces may increase with reaction time, which will be studied in a forthcoming publication. Here, we may already underline the fact that the observed trends imply a significant mobility of the AAs and peptides on the surface since the 2-D AA density was only 1/10 of a physical monolayer: in other words, if the AAs are homogeneously distributed, they are not in close contact with their neighbors.

Third and most central to the usefulness for origin of life scenarios comes the question of stoichiometry and sequence selectivity. The interpretation of experimental data is not straightforward due to the well-known difficulty of quantifying MS results. We have attempted to evaluate the ionization yields through the analysis of reference tripeptides. This process is quite cumbersome and semiquantitative at best. However, the example of tripeptides, which was worked out in detail, as well as the less detailed analysis of other formed polypeptides, definitely indicates that the mineral surface scenario is selective, in the sense that some reaction products are favored at the expense of others, most probably due to the selective opening of some reaction pathways by catalytic effects. The preferential formation of heteropeptides rather than homopeptides is a general trend; the preference of EL<sub>2</sub> over the E<sub>2</sub>L stoichiometry is less marked but still seems significant. Going from stoichiometry to sequences is even harder as it requires MS/MS experiments but can be quite rewarding, showing, for instance, that the EL<sub>2</sub> stoichiometry predominantly corresponds to the ELL sequence rather than the other two theoretically possible ones. Partial data on other oligopeptides also point at significant sequence selectivities in general.

The definition of selectivity generally used in heterogeneous/industrial catalysis, namely the amount of starting reagent transformed to the product of interest, expressed as a percentage of total reagent transformation, is insufficient to evaluate the prebiotic interest of this or any other polymerization scenario. Instead, as a first step toward evaluating to what degree a particular scenario can preferentially lead to a particular peptide or a particular category of peptides, one would like to have a quantitative measure of how much the observed distribution deviates from a “nondiscriminating one”.

The problem is made more complex because for a complete treatment, we would need to address both the distribution among various polypeptide lengths and, for a given length, the distribution among possible sequences/stoichiometries. Let us consider the second question, and more specifically, let us take a specific example, the distribution among possible tripeptides, for which we have the most experimental data.

It turns out that this problem is akin to one that was often addressed in theoretical biochemistry, namely, the assessment of genetic heterogeneity among a population of organisms. Among other existing practical measures, Adami<sup>53,54</sup> proposed to evaluate the average physical complexity of an ensemble of genetic sequences  $X$ , according to a formula based on Shannon's entropy in information theory.<sup>55</sup> In their case,  $X$  was an ensemble of known sequences, that is, of DNA chains with potentially very high lengths ( $L$ , the length of the sequence, can reach up to two million bases even for a single gene). In our case,  $X$  would be an ensemble of oligopeptides, for example of tripeptides, with  $L = 3$ . The entropy  $H^a$  is calculated along

$$H(X) = -\sum_{i=1} p_i \log p_i$$

where the summation is over all possible sequences  $i$  in ensemble  $X$ , and  $p_i$  is their respective probabilities. The number of possible sequences depends on the sequence length and on the number of monomer types,  $D$  (the “size of the alphabet”). In the case of DNA sequences  $D = 4$ , there are four possible nucleobases, while in our case  $D = 2$ , the monomers can be either L or E. The number of possible sequences is then  $N = D^L$ ; in our “world” of E + L polymerization,  $N = 2^3 = 8$  (LLL, LLE ... EEE).

Now if all sequences were equally likely, the probability of each one would be  $1/N$ , so that

$$H_{\max}(X) = -\sum_{i=1}^N \frac{1}{N} \log\left(\frac{1}{N}\right) = \log N,$$

providing a maximum entropy situation. Genomes in a real environment,  $E$ , are submitted to (natural) selection pressures so that the observed sequence probabilities  $p_i$  are greatly different from this equal probability situation. The conditional entropy of the ensemble of sequences  $X$  in environment  $E$  will then be different from  $H_{\max}$  and it is denoted  $H(X|E)$ .  $H(X|E)$  is indeed lower than  $H_{\max}$  because the sequence distribution in the real environment is “more ordered”; it is easy to realize that in the limiting case where only one sequence would be expressed in the population,  $H(X|E)$  would be zero. The difference  $H_{\max} - H(X|E)$  is denoted as  $I(X:E)$  and may be considered as an information content. Adami indeed equates it



to “the information that  $X$  has about  $E$ ”, the information that the population of genomes has about its environment.

It is easy to apply these notions to the world of oligopeptide synthesis. In our case,  $p_i$  may be equated to the relative abundances of the trimers in the analytical mixture (estimated by correcting the peak intensities for ionization efficiencies as explained above), and  $E$  will be the “silica surface environment” or in more common parlance, the silica surface AA polymerization scenario, so we will write  $\text{SiO}_2$  instead of  $E$ .  $I(X:\text{SiO}_2)$  will then constitute a quantitative evaluation of the “nonrandomness” of the observed oligopeptide distribution due to the fact that it is established in the  $\text{SiO}_2$  scenario. This extension would be rather straightforward if we had precise data on the experimental abundances of each trimeric sequence. Establishing oligopeptide sequences requires a lot of experimental work (in the results discussed above, the predominant sequence was unequivocally established only for the  $\text{EL}_2$  stoichiometry), so that we will couch the calculation in terms of experimentally observed stoichiometries instead. Then, as already stated, the maximum entropy (=random scenario) will correspond to relative intensities 1:3:3:1, which may be converted to probabilities by normalizing to 1. The steps of the calculation of  $I(X:\text{SiO}_2)$  are shown in Table 1, and this quantity is found equal to  $1.256 - 0.788 = 0.468$  entropy units. This approach may be used to compare the two polymerization scenarios, on silica and in solution with the CDI activator. It is found that  $I(X:\text{CDI})$ , is only 0.222 entropy units.

This example shows that it is possible to quantify the information content of an experimentally observed distribution of polymers obtained in a particular polymerization scenario. A generalization accounting also for the distribution between different polymer lengths could be devised, although practical measurements involve some experimental difficulties. The quantity  $I(X:\text{SiO}_2)$  calculated in this way has tantalizing implications in research on the origins of life. It was originally defined for populations of genomes, and in this context, it may be intuitively understood as the “amount of information that a population stores about the environment in which it lives”.<sup>54</sup> Nobody would of course contend that the mixed oligopeptide systems formed by surface AA polymerization are alive; to the best of our knowledge, they do not do anything useful from a systems chemistry perspective, while genomes are functional macromolecules. Still, it may be interesting to have a common measure of complexity that can be applied at such widely separated steps on the road from chemistry to life.

The course we have followed in the preceding developments is to evaluate the nonrandom character of a polypeptide distribution in order to determine whether it is promising in an OoL scenario. Thus, we took the polypeptide distribution as a piece of data to be interpreted, without trying to establish how exactly it was produced. The latter question is of course a legitimate one. Specialists of surface science and catalysis may think it is the central one, and indeed, unraveling the surface reaction mechanisms would be very helpful to determine how polymerization selectivity can be improved, other than by blind variation of the reaction parameters. This endeavor is not without precedent in the literature, in fact essentially for one system similar to ours.

Dose *et al.* have studied the copolymerization of Glu with several AAs with hydrophobic residues (thus similar to Leu), denoted  $X$ , both in the bulk polymerization<sup>27</sup> and carbodiimide<sup>56</sup> (a solution scenario with a sacrificial activator, akin to

our CDI) scenarios. They found that the trimers formed mostly consisted of PyroGlu- $X$ - $X$  ( $\text{pE-X-X}$ ), similar to the predominance of  $\text{E-L-L}$  that we observed. They explained this by hypothesizing that Glu first undergoes internal cyclization to pyroglutamic acid (“self-protecting” the  $\text{NH}_2$  terminal of Glu) while the AAs  $X$  condense to cyclic dimers, followed by the reaction  $\text{pE} + \text{cyclo}(X-X) = \text{pE-L-L}$ . It is tempting to transpose their mechanism to our systems; it would be compatible with the observed reactivity of the single AA systems and with the existence of large amounts of dehydrated peptides, in our experiments. However, our initial results (data not shown) indicate that this mechanism probably does not take place in our case and we will leave a full mechanistic discussion for a further publication.

## 5. CONCLUSIONS

In this study, close to quantitative formation of oligopeptides up to octamers has been demonstrated for Leu + Glu/ $\text{SiO}_2$  systems after a single thermal activation step, a realistic scenario in origins of life studies. Unambiguous peptide assignments with ppm mass accuracy measurements were obtained by using HRMS detection. These, and MS/MS experiments, demonstrate nontrivial distributions of oligomers between different stoichiometries and possible AA sequences, which probably have a kinetic origin—a memory of heterogeneous catalytic pathways opened up on the inorganic surface. Polymerization in aqueous solution through the use of sacrificial activators does not result in the same oligopeptide distributions, being closer to random. In order to quantify the information contained in experimentally observed distributions, we propose to use a metric derived from genomic analysis. However, it should not be overinterpreted. It provides an interesting measure to evaluate the otherwise ill-defined notion of selectivity in the progressive emergence of biopolymers. Further investigations will be performed into the polymerization process on surfaces, especially to follow polymerization kinetics and thus establish a likely mechanism, as well as investigations will be performed on the polymerization selectivity with other characterization methods (hyphenated methods as liquid chromatography/mass spectrometry or ion mobility-mass spectrometry to identify peptide sequences). More complex AA mixtures will also be investigated in order to get closer to the likely complexity of prebiotic environments.

## ■ ASSOCIATED CONTENT

### SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acsearthspacechem.0c00183>.

XRD and differential thermal analysis characterizations; additional ESI and MS/MS spectra; peptide distribution histograms; and IR bands assignments (PDF)


## ■ AUTHOR INFORMATION

### Corresponding Authors

Sandra Alves – Institut Parisien de Chimie Moléculaire (UMR 8232 CNRS), Sorbonne Université, Paris 75005, France;

Email: [sandra.alves@sorbonne-universite.fr](mailto:sandra.alves@sorbonne-universite.fr)

Jean-François Lambert – Laboratoire de Réactivité de Surface (UMR 7197 CNRS), Sorbonne Université, Paris 75005,

France;  [orcid.org/0000-0002-8124-5709](https://orcid.org/0000-0002-8124-5709); Email: [jean-francois.lambert@sorbonne-universite.fr](mailto:jean-francois.lambert@sorbonne-universite.fr)

## Author

Lise Bedoin – Institut Parisien de Chimie Moléculaire (UMR 8232 CNRS) and Laboratoire de Réactivité de Surface (UMR 7197 CNRS), Sorbonne Université, Paris 75005, France

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acsearthspacechem.0c00183>

## Author Contributions

The manuscript was written through contributions of all authors.

## Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This program was supported by a Ph.D. grant from the Paris Ile-de-France region awarded in the frame of the DIM program ACAV+ under the project name EmBioPol. A financial contribution of the CNRS MIT programme Origines 2019 is acknowledged. The National FT-ICR network (FR 3624 CNRS) providing FT-ICR facilities for conducting the research is also gratefully acknowledged.

## ADDITIONAL NOTE

“The choice of this notation for the entropy is unfortunate since it may bring about a confusion with the enthalpy in general thermodynamics, but it is already well-established.

## REFERENCES

- (1) Miller, S. L. A Production of Amino Acids under Possible Primitive Earth Conditions. *Science* **1953**, *117*, 528–529.
- (2) Parker, E. T.; Cleaves, H. J.; Dworkin, J. P.; Glavin, D. P.; Callahan, M.; Aubrey, A.; Lazcano, A.; Bada, J. L. Primordial Synthesis of Amines and Amino Acids in a 1958 Miller H<sub>2</sub>S-Rich Spark Discharge Experiment. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 5526–5531.
- (3) Meierhenrich, U. J.; Munoz Caro, G. M.; Bredehöft, J. H.; Jessberger, E. K.; Thiemann, W. H.-P. Identification of Diamino Acids in the Murchison Meteorite. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9182–9186.
- (4) Brack, A. From Interstellar Amino Acids to Prebiotic Catalytic Peptides: A Review. *Chem. Biodiversity* **2007**, *4*, 665–679.
- (5) Maury, C. P. J. Origin of Life. Primordial Genetics: Information Transfer in a Pre-RNA World Based on Self-Replicating Beta-Sheet Amyloid Conformers. *J. Theor. Biol.* **2015**, *382*, 292–297.
- (6) Issac, R.; Chmielewski, J. Approaching Exponential Growth with a Self-Replicating Peptide. *J. Am. Chem. Soc.* **2002**, *124*, 6808–6809.
- (7) Gorlero, M.; Wiczorek, R.; Adamala, K.; Giorgi, A.; Schinà, M. E.; Stano, P.; Luisi, P. L. Ser-His Catalyses the Formation of Peptides and PNAs. *FEBS Lett.* **2009**, *583*, 153–156.
- (8) Danger, G.; Boiteau, L.; Cottet, H.; Pascal, R. The Peptide Formation Mediated by Cyanate Revisited. N-Carboxyanhydrides as Accessible Intermediates in the Decomposition of N-Carbamoyl amino Acids. *J. Am. Chem. Soc.* **2006**, *128*, 7412–7413.
- (9) Harada, K.; Fox, S. W. The Thermal Condensation of Glutamic Acid and Glycine to Linear Peptides I. *J. Am. Chem. Soc.* **1958**, *80*, 2694–2697.
- (10) Rodríguez-García, M.; Surman, A. J.; Cooper, G. J. T.; Suárez-Marina, I.; Hosni, Z.; Lee, M. P.; Cronin, L. Formation of Oligopeptides in High Yield under Simple Programmable Conditions. *Nat. Commun.* **2015**, *6*, 8385.
- (11) Bernal, J. D. *The Physical Basis of Life*; Routledge and Kegan Paul: Londres, 1951.
- (12) de Duve, C.; Miller, S. L. Two-Dimensional Life? *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 10014–10017.
- (13) Smith, J. V. Biochemical Evolution. I. Polymerization on Internal, Organophilic Silica Surfaces of Dealuminated Zeolites and Feldspars. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 3370–3375.
- (14) Bujdák, J.; Rode, B. M. Silica, Alumina, and Clay-Catalyzed Alanine Peptide Bond Formation. *J. Mol. Evol.* **1997**, *45*, 457–466.
- (15) Lahav, N.; Chang, S. The Possible Role of Solid Surface Area in Condensation Reactions during Chemical Evolution: Reevaluation. *J. Mol. Evol.* **1976**, *8*, 357–380.
- (16) Bujdák, J.; Slosiarikova, H.; Texler, N.; Schwendinger, M.; Rode, B. M. On the Possible Role of Montmorillonites in Prebiotic Peptide Formation. *Monatsh. Chem.* **1994**, *125*, 1033–1039.
- (17) Basiuk, V. A.; Gromovoy, T. Y.; Glukhoy, A. M.; Golovaty, V. G. Chemical Transformations of Proteinogenic Amino Acids during Their Sublimation in the Presence of Silica. *Origins Life Evol. Biospheres* **1991**, *21*, 129–144.
- (18) Guo, C.; Holland, G. P. Alanine Adsorption and Thermal Condensation at the Interface of Fumed Silica Nanoparticles: A Solid-State NMR Investigation. *J. Phys. Chem. C* **2015**, *119*, 25663–25672.
- (19) Lambert, J.-F.; Jaber, M.; Georgelin, T.; Stievano, L. A Comparative Study of the Catalysis of Peptide Bond Formation by Oxide Surfaces. *Phys. Chem. Chem. Phys.* **2013**, *15*, 13371–13380.
- (20) Georgelin, T.; Jaber, M.; Bazzi, H.; Lambert, J.-F. Formation of Activated Biomolecules by Condensation on Mineral Surfaces – A Comparison of Peptide Bond Formation and Phosphate Condensation. *Origins Life Evol. Biospheres* **2013**, *43*, 429–443.
- (21) Bouchoucha, M.; Jaber, M.; Onfroy, T.; Lambert, J.-F.; Xue, B. Glutamic Acid Adsorption and Transformations on Silica. *J. Phys. Chem. C* **2011**, *115*, 21813–21825.
- (22) Meng, M.; Stievano, L.; Lambert, J.-F. Adsorption and Thermal Condensation Mechanisms of Amino Acids on Oxide Supports. I. Glycine on Silica. *Langmuir* **2004**, *20*, 914–923.
- (23) Miyakawa, S.; Ferris, J. P. Sequence- and Regioselectivity in the Montmorillonite-Catalyzed Synthesis of RNA. *J. Am. Chem. Soc.* **2003**, *125*, 8202–8208.
- (24) Fox, S. W.; Harada, K. The Thermal Copolymerization of Amino Acids Common to Protein. *J. Am. Chem. Soc.* **1960**, *82*, 3745–3751.
- (25) Fox, S. W.; Harada, K. Thermal Copolymerization of Amino Acids in the Presence of Phosphoric Acid. *Arch. Biochem. Biophys.* **1960**, *86*, 281–285.
- (26) Nakashima, T.; Jungck, J. R.; Fox, S. W.; Lederer, E.; Das, B. C. A Test for Randomness in Peptides Isolated from a Thermal Polyamino Acid. *Int. J. Quantum Chem.* **1977**, *12*, 65–72.
- (27) Hartmann, J.; Christel Brand, M.; Dose, K. Formation of Specific Amino Acid Sequences during Thermal Polymerization of Amino Acids. *Biosystems* **1981**, *13*, 141–147.
- (28) Mosqueira, F. G.; Ramos-Bernal, S.; Negrón-Mendoza, A. A Simple Model of the Thermal Prebiotic Oligomerization of Amino Acids. *Biosystems* **2000**, *57*, 67–73.
- (29) Bujdák, J.; Rode, B. M. Preferential Amino Acid Sequences in Alumina-Catalyzed Peptide Bond Formation. *J. Inorg. Biochem.* **2002**, *90*, 1–7.
- (30) McKee, A. D.; Solano, M.; Saydjari, A.; Bennett, C. J.; Hud, N. V.; Orlando, T. M. A Possible Path to Prebiotic Peptides Involving Silica and Hydroxy Acid-Mediated Amide Bond Formation. *ChemBioChem* **2018**, *19*, 1913–1917.
- (31) Frenkel-Pinter, M.; Haynes, J. W.; C, M.; Petrov, A. S.; Burcar, B. T.; Krishnamurthy, R.; Hud, N. V.; Leman, L. J.; Williams, L. D. Selective Incorporation of Proteinaceous over Nonproteinaceous Cationic Amino Acids in Model Prebiotic Oligomerization Reactions. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116*, 16338–16346.
- (32) Forsythe, J. G.; Yu, S.-S.; Mamajanov, I.; Grover, M. A.; Krishnamurthy, R.; Fernández, F. M.; Hud, N. V. Ester-Mediated Amide Bond Formation Driven by Wet-Dry Cycles: A Possible Path to Polypeptides on the Prebiotic Earth. *Angew. Chem., Int. Ed.* **2015**, *54*, 9871–9875.

- (33) Xiong, H.; Buckwalter, B. L.; Shieh, H. M.; Hecht, M. H. Periodicity of Polar and Nonpolar Amino Acids Is the Major Determinant of Secondary Structure in Self-Assembling Oligomeric Peptides. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 6349–6353.
- (34) Bertrand, M.; Brack, A. Conformational Variety of Polyanionic Peptides At Low Salt Concentrations. *Origins Life Evol. Biospheres* **1997**, *27*, 585–595.
- (35) Lambert, J.-F. Adsorption and Polymerization of Amino Acids on Mineral Surfaces: A Review. *Origins Life Evol. Biospheres* **2008**, *38*, 211–242.
- (36) Zamaraev, K. I.; Romannikov, V. N.; Salganik, R. I.; Wlassoff, W. A.; Khramtsov, V. V. Modelling of the Prebiotic Synthesis of Oligopeptides: Silicate Catalysts Help to Overcome the Critical Stage. *Origins Life Evol. Biospheres* **1997**, *27*, 325–337.
- (37) Sun, Y.; Frenkel-Pinter, M.; Liotta, C. L.; Grover, M. A. The pH Dependent Mechanisms of Non-Enzymatic Peptide Bond Cleavage Reactions. *Phys. Chem. Chem. Phys.* **2020**, *22*, 107–113.
- (38) Martra, G.; Deiana, C.; Sakhno, Y.; Barberis, I.; Fabbiani, M.; Pazzi, M.; Vincenti, M. The Formation and Self-Assembly of Long Prebiotic Oligomers Produced by the Condensation of Unactivated Amino Acids on Oxide Surfaces. *Angew. Chem., Int. Ed.* **2014**, *53*, 4671–4674.
- (39) Sakhno, Y.; Battistella, A.; Mezzetti, A.; Jaber, M.; Georgelin, T.; Michot, L.; Lambert, J.-F. One Step up the Ladder of Prebiotic Complexity: Formation of Nonrandom Linear Polypeptides from Binary Systems of Amino Acids on Silica. *Chem.—Eur. J.* **2019**, *25*, 1275–1285.
- (40) Ikehara, K. Possible Steps to the Emergence of Life: The [GADV]-Protein World Hypothesis. *Chem. Rec.* **2005**, *5*, 107–118.
- (41) van der Gulik, P.; Massar, S.; Gilis, D.; Buhrman, H.; Roonan, M. The First Peptides: The Evolutionary Transition between Prebiotic Amino Acids and Early Proteins. *J. Theor. Biol.* **2009**, *261*, 531–539.
- (42) Longo, L. M.; Lee, J.; Blaber, M. Simplified Protein Design Biased for Prebiotic Amino Acids Yields a Foldable, Halophilic Protein. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 2135–2139.
- (43) Gordon, E. F.; Muddiman, D. C. Impact of Ion Cloud Densities on the Measurement of Relative Ion Abundances in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: Experimental Observations of Coulombically Induced Cyclotron Radius Perturbations and Ion Cloud Dephasing Rates. *J. Mass Spectrom.* **2001**, *36*, 195–203.
- (44) Ehler, K. W.; Orgel, L. E. N,N'-carbonyldiimidazole-induced peptide formation in aqueous solution. *Biochim. Biophys. Acta, Protein Struct.* **1976**, *434*, 233–243.
- (45) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: A Primer. *Mass Spectrom. Rev.* **1998**, *17*, 1–35.
- (46) Nunes, R. S.; Cavalheiro, É. T. G. Thermal behavior of glutamic acid and its sodium, lithium and ammonium salts. *J. Therm. Anal. Calorim.* **2007**, *87*, 627–630.
- (47) Jaber, M.; Georgelin, T.; Bazzi, H.; Costa-Torro, F.; Lambert, J.-F.; Bolbach, G.; Clodic, G. Selectivities in Adsorption and Peptidic Condensation in the (Arginine and Glutamic Acid)/Montmorillonite Clay System. *J. Phys. Chem. C* **2014**, *118*, 25447–25455.
- (48) Cech, N. B.; Enke, C. G. Practical Implications of Some Recent Studies in Electrospray Ionization Fundamentals. *Mass Spectrom. Rev.* **2001**, *20*, 362–387.
- (49) Lill, J. Proteomic Tools for Quantitation by Mass Spectrometry. *Mass Spectrom. Rev.* **2003**, *22*, 182–194.
- (50) Roepstorff, P.; Fohlman, J. Proposal for a Common Nomenclature for Sequence Ions in Mass Spectra of Peptides. *Biol. Mass Spectrom.* **1984**, *11*, 601.
- (51) Downard, K. M.; Biemann, K. The Effect of Charge State and the Localization of Charge on the Collision-Induced Dissociation of Peptide Ions. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 966–975.
- (52) Lahav, N.; White, D.; Chang, S. Peptide Formation in the Prebiotic Era: Thermal Condensation of Glycine in Fluctuating Clay Environments. *Science* **1978**, *201*, 67–69.
- (53) Adami, C.; Cerf, N. J. Physical Complexity of Symbolic Sequences. *Phys. D* **2000**, *137*, 62–69.
- (54) Adami, C. What Is Complexity? *BioEssays* **2002**, *24*, 1085–1094.
- (55) Shannon, C. E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423.
- (56) Dose, K.; Hartmann, J.; Christel Brand, M. Formation of Specific Amino Acid Sequences during Carbodiimide-Mediated Condensation of Amino Acids in Aqueous Solution. *Biosystems* **1982**, *15*, 195–200.

## Annexe 2. Sélectivités stœchiométriques pour les oligopeptides de longueur variable dans l'échantillon de référence

La comparaison des intensités expérimentales obtenues avec les intensités théoriques *non corrigées* pour une polymérisation aléatoire est reportée en Figure A1. Pour toutes les tailles d'oligopeptides, les hétéro-oligopeptides montrent une abondance supérieure aux homo-oligopeptides.

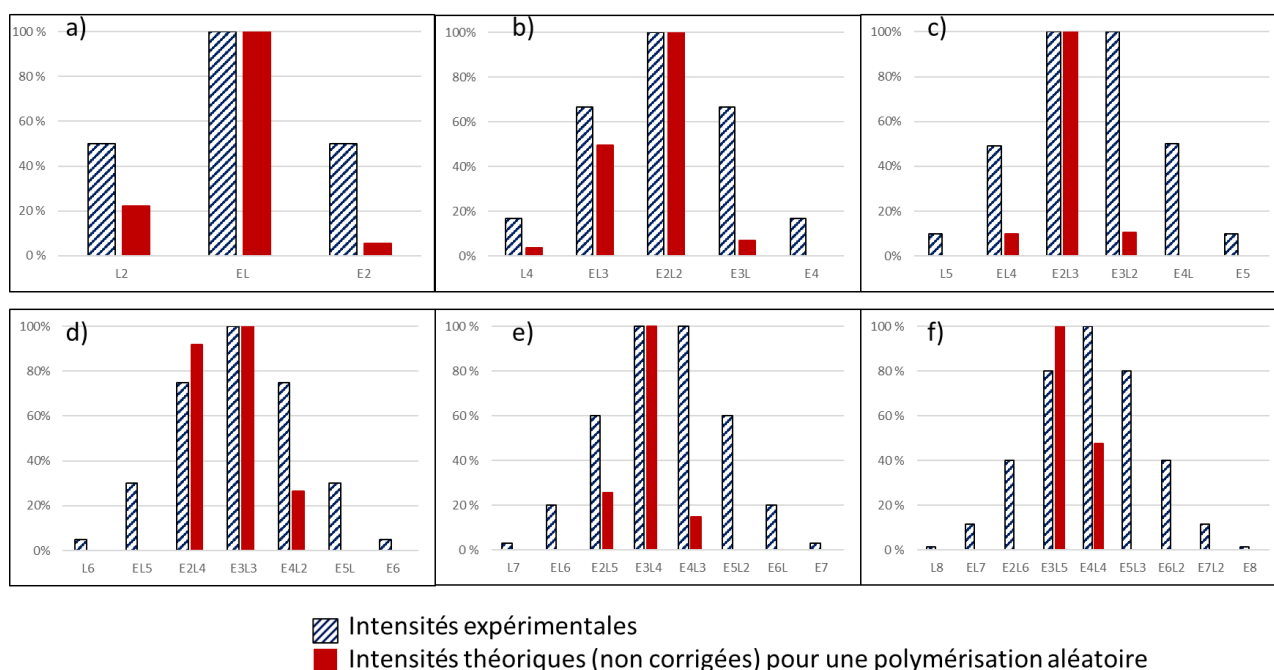


Figure A1. Comparaison des intensités expérimentales avec les intensités théoriques non corrigées attendues dans le cas d'une polymérisation aléatoire du système Leu + Glu pour toutes les longueurs a) dipeptides, b) tétrapeptides, c) pentapeptides, d) hexapeptides, e) heptapeptides, f) octapeptides. Le cas des tripeptides est montré dans le Chapitre III.

### Annexe 3. Intensités expérimentales et conservation du nombre d'acides aminés

Le Tableau A1 donne les résultats des calculs de la relation (3) du Chapitre II en utilisant les intensités des pics des monomères, dipeptides et tripeptides à la place des fractions molaires qui sont inconnues. Les valeurs obtenues sont, à l'exception de la température la plus basse, comprises entre 0,48 et 0,52, ce qui est proche des fractions molaires initialement introduites dans le mélange.

Tableau A1. Résultats des calculs de la relation (3) du Chapitre II pour les différentes températures d'activation testées.

T(°C)	$w_E^o$	$w_L^o$
100	0.60	0.40
115	0.52	0.48
130	0.51	0.49
145	0.49	0.51
160	0.48	0.52



## Annexe 4. Stœchiométries des oligopeptides formés dans l'échantillon (EE+L)/SiO<sub>2</sub> activé à 130 et 160 °C

La Figure A2 montre la répartition des différentes stœchiométries des oligopeptides formés pour l'échantillon (EE+L)/SiO<sub>2</sub> activé à 130 et 160 °C. Les pourcentages montrés correspondent à l'abondance d'une stœchiométrie donnée : les intensités des pics d'oligopeptides présentant cette stœchiométrie sont sommées et la somme obtenue est ensuite normalisée par rapport à la somme de tous les oligopeptides détectés dans l'échantillon. Les oligopeptides qui s'expliquent par l'addition de dimères et de monomères représentent 84,2 % des intensités des oligopeptides totaux pour une activation à 160 °C contre 98 % pour une activation à 130 °C.

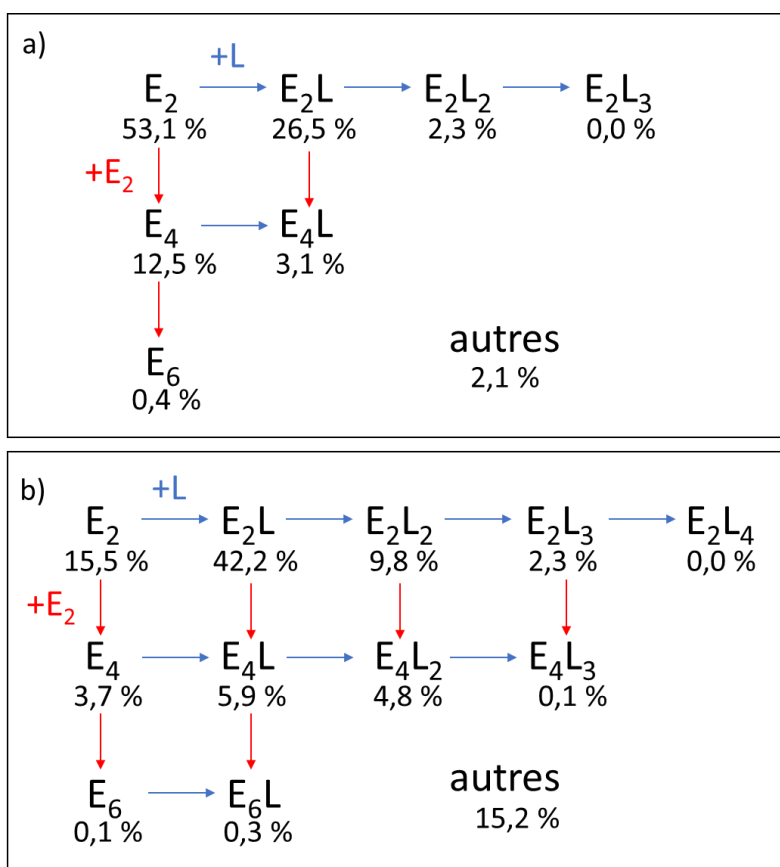


Figure A2. Répartition des différentes stœchiométries des oligopeptides formés pour l'échantillon les échantillons a) (EE+L)/SiO<sub>2</sub> activé à 130 °C et b) (EE+L)/SiO<sub>2</sub> activé à 160 °C. Les flèches bleues correspondent à une addition de résidu Leu, les rouges à une addition de E<sub>2</sub>.

## Annexe 5. Analyse des échantillons (cycloAA+E)/SiO<sub>2</sub> et (AA+E)/SiO<sub>2</sub>

Les spectres de masse des échantillons (cycloAA+E)/SiO<sub>2</sub> et (AA+E)/SiO<sub>2</sub> sont montrés sur la Figure A3. Les oligopeptides détectés sont identiques, que le dimère précurseur soit cyclique ou linéaire et quelle que soit la température d'activation. La plus longue espèce détectée est un tétramère pour chacune des activations. Les espèces détectées sont en fait des dimères non covalents. On peut donc en déduire que pour ces systèmes, aucune polymérisation (ou du moins efficace) n'a eu lieu.

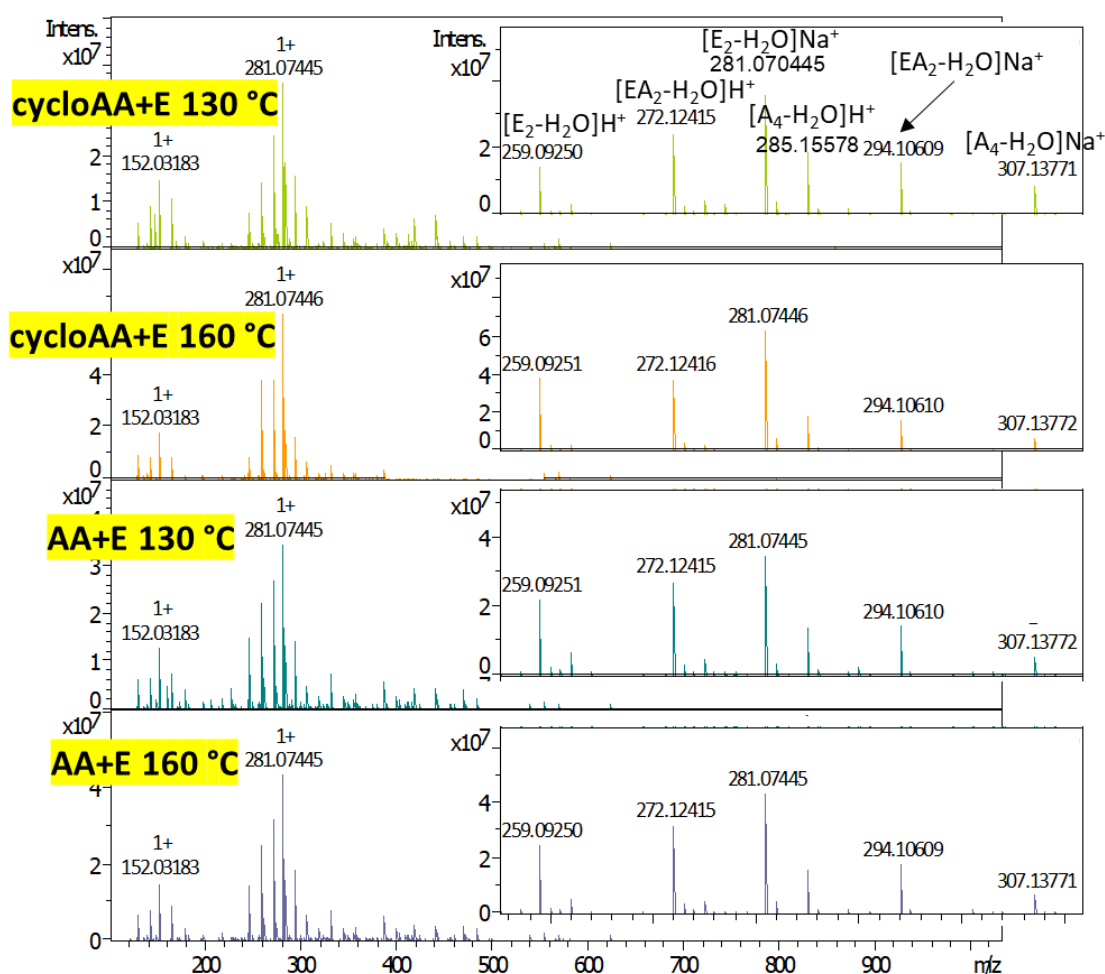


Figure A3. Spectres de masse obtenus pour les échantillons (cycloAA+E)/SiO<sub>2</sub> et (AA+E)/SiO<sub>2</sub> activés thermiquement à 130 °C et 160 °C (ESI positif, instrument FTICR).

## Annexe 6. Etude de la polymérisation de (pE+L)/SiO<sub>2</sub>

Cette étude s'intéresse à la possibilité d'effectuer une polymérisation à partir d'un monomère d'acide glutamique déjà cyclisé, l'acide pyroglutamique (pyroGlu). En théorie, si le cycle lactame ne peut pas s'ouvrir, alors la polymérisation ne pourrait se faire que par le groupement carboxyle en C-terminal et ainsi donner des oligopeptides ne contenant qu'un seul résidu pyroGlu en N-terminal (c'est-à-dire des stœchiométries (ELy-H<sub>2</sub>O)).

Le spectre ESI obtenu pour cet échantillon est montré en Figure A4. Les mêmes espèces que celles obtenues pour l'échantillon (L+E)/SiO<sub>2</sub>, ayant des longueurs allant jusqu'aux octapeptides, sont détectées y compris des espèces intactes ainsi que des espèces comportant plus de résidus Glu que de pertes d'eau (par exemple [E<sub>2</sub>L<sub>2</sub>-H<sub>2</sub>O]H<sup>+</sup>). Ces observations suggèrent que le cycle lactame peut s'ouvrir et être impliqué dans une condensation.

On a montré qu'une polymérisation était possible et que les résultats obtenus sont identiques à ceux de (E+L)/SiO<sub>2</sub>.

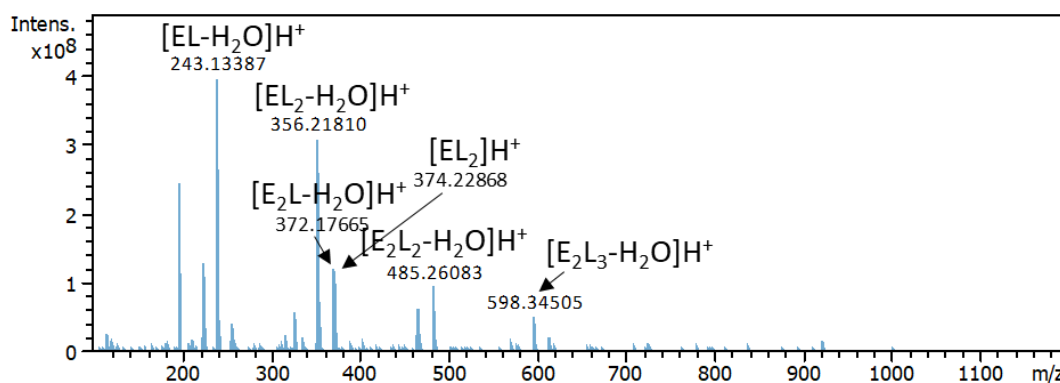


Figure A4. Spectre de masse de l'échantillon (pE+L)/SiO<sub>2</sub> activé à 160 °C (ESI positif, instrument FTICR). Il a été vérifié que les espèces présentes sont bien des espèces covalentes.

## Annexe 7. Spectres IR de l'échantillon (L+E)/Mt

Les spectres infrarouge (opérées en mode ATR) avant activation thermique, après activation thermique et après désorption sont reportés sur la Figure A5.

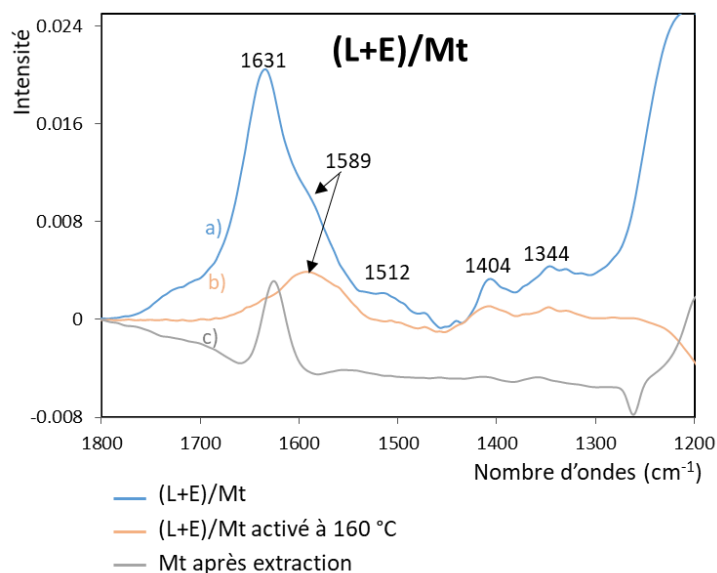


Figure A5. Spectres IR de l'échantillon (L+E)/Mt avant activation thermique (a), après activation thermique (b) et spectre de la montmorillonite après désorption (c). Le spectre de la montmorillonite a été soustrait aux spectres a et b ; le spectre c a été soustrait au spectre de l'eau. Chaque spectre a été décalé pour avoir une intensité de 0 à 1800  $\text{cm}^{-1}$ .

En se basant sur les travaux de Fuchida<sup>1</sup>, toutes les bandes observées peuvent être attribuées. Le spectre avant activation correspond à celui obtenu pour l'adsorption de glycine sur montmorillonite<sup>1</sup>. Les spectres avant et après activation présentent les mêmes bandes caractéristiques des acides aminés : la bande à 1404  $\text{cm}^{-1}$  correspond à l'élongation symétrique  $\nu_s$  ( $-\text{COO}^-$ ) de l'acide carboxylique, la bande à 1344  $\text{cm}^{-1}$  à la rotation des ( $-\text{CH}_2$ ), celles à 1589  $\text{cm}^{-1}$  et 1512  $\text{cm}^{-1}$  respectivement aux déformations asymétrique et symétrique des ( $-\text{NH}_3^+$ ). La différence majeure réside dans la présence avant activation de la bande à 1631  $\text{cm}^{-1}$  (spectre a), qui correspond à la vibration de déformation de l'eau<sup>1</sup>. Après activation thermique cette bande a disparu. Aucune bande amide, caractéristique de la formation d'une liaison peptidique, n'est observée après activation thermique ; notamment, la bande amide I serait attendue à 1671  $\text{cm}^{-1}$ .

<sup>1</sup> Fuchida, S.; Masuda, H.; Shinoda, K. Peptide Formation Mechanism on Montmorillonite Under Thermal Conditions. *Orig. Life Evol. Biospheres* **2014**, 44 (1), 13–28. <https://doi.org/10.1007/s11084-014-9359-4>.

## Annexe 8. Etude du spectre MS de l'échantillon (L+E)/Mt

Le spectre de masse de l'échantillon (L+E)/Mt est reporté en Figure A6. La plupart des pics présents correspondent aux monomères, avec des intensités faibles comparables à celles des pics du solvant.

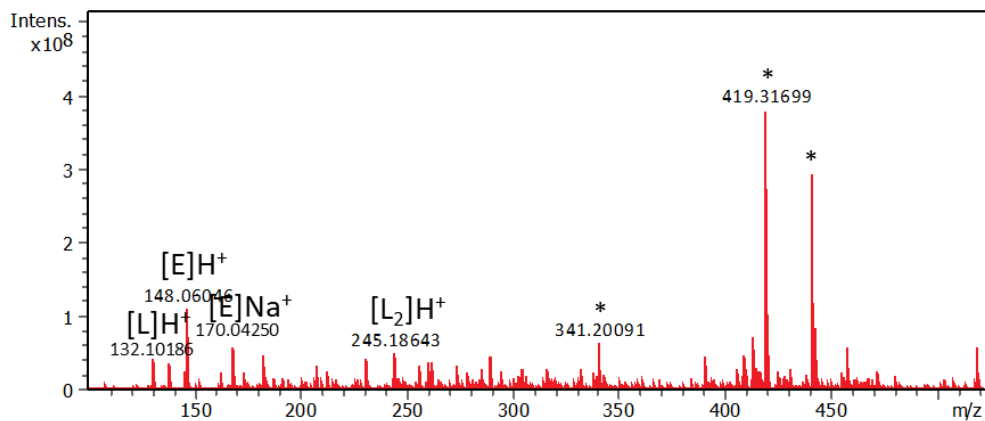


Figure A6. Spectre de masse de l'échantillon (L+E)/Mt activé à 160 °C sur la gamme m/z 100-500 (ESI positif, instrument FTICR). Les pics notés avec un astérisque correspondent à des pics de solvant.

## Annexe 9. Spectre IR de l'alumine après étape d'extraction

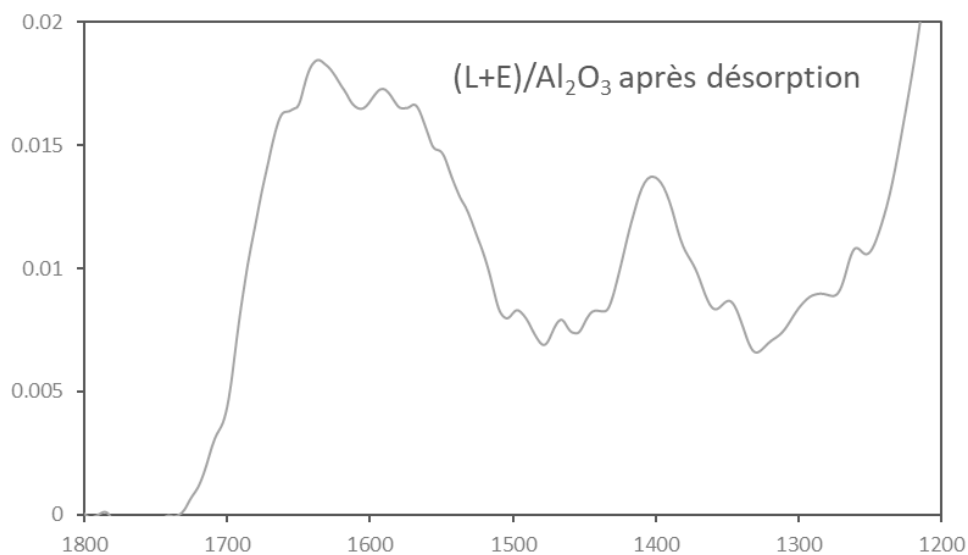
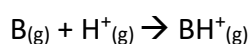
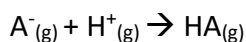


Figure A7. Spectre infrarouge de l'échantillon (L+E)/Al<sub>2</sub>O<sub>3</sub> après désorption

Malgré la mauvaise qualité de ce spectre, une absorbance non négligeable, non attribuable au support, est observée vers 1590 cm<sup>-1</sup>. Cela traduit la présence de matière organique.

## Annexe 10. Méthode cinétique de Cooks

L'affinité protonique, notée  $E_{ap}$  d'une espèce chimique est une estimation de sa basicité en phase gazeuse, correspondant à l'énergie produite par l'une des réactions :



Plus l'affinité protonique est élevée, plus la base est forte, et par conséquent plus l'acide conjugué est faible en phase gazeuse.

La méthode cinétique de Cooks permet de déterminer l'affinité protonique intrinsèque relative de deux espèces<sup>2</sup>. Elle consiste à étudier les fragmentations compétitives d'un hétérodimère non covalent protoné ( $B_1+B_2+H^+$ ). L'intensité relative des abondances des ions fragments (monomères protonés  $B_1H^+$  et  $B_2H^+$ ) renseigne, en théorie, sur l'espèce,  $B_1$  ou  $B_2$ , qui a la plus grande affinité protonique.

### I. Méthode cinétique de Cooks pour Gln, $\beta$ Glu, Glu et Asp

Afin de d'évaluer l'ordre relatif des affinités protoniques des acides aminés trifonctionnels étudiés, un mélange des quatre monomères  $\beta$ Glu, Glu\*, Gln et Asp a été préparé. « Glu\* » désigne un Glu marqué au  $^{13}C$ , choisi pour permettre la différenciation avec sa forme isomère  $\beta$ Glu. La solution a été faite à concentration élevée de manière à obtenir des dimères non covalents. Les hétérodimères protonés ont ensuite été fragmentés pour évaluer les affinités protoniques relatives. Les spectres MS/MS sont montrés à la Figure A8.

---

<sup>2</sup> Cooks, R. G.; Kruger, T. L. Intrinsic Basicity Determination Using Metastable Ions. *J. Am. Chem. Soc.* **1977**, *99* (4), 1279–1281. <https://doi.org/10.1021/ja00446a059>.

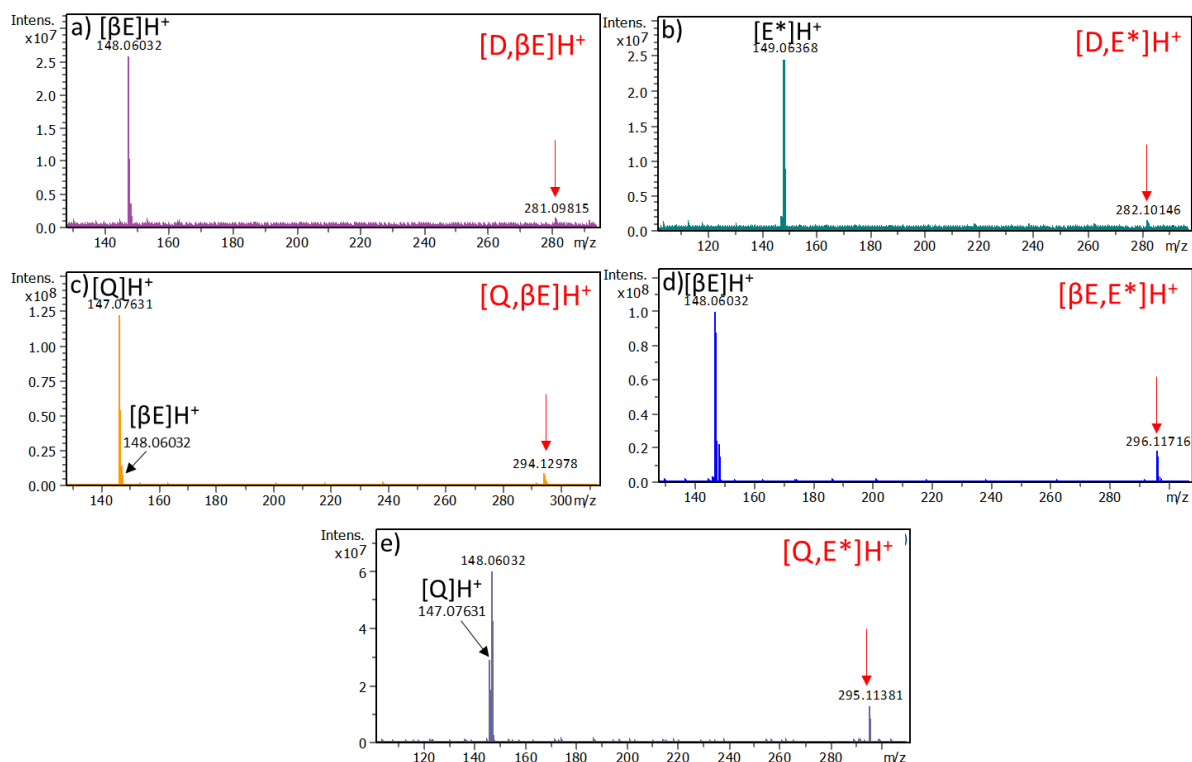


Figure A8. Spectres MS/MS des dimères non covalents.

Dans le cas de  $[Q,E^*]H^+$  (spectre e), le rapport  $m/z$  295,11381 correspond également au dimère non covalent  $[\beta E,\beta E]H^+$ , ce qui explique la présence du fragment à  $m/z$  148,06032. La comparaison de l'intensité des fragments doit se faire entre  $m/z$  147,07631 et  $m/z$  149,06368 (qui est absent).

Les MSMS permettent d'établir l'ordre relatif des affinités protoniques, en l'occurrence  $Gln > \beta Glu > Glu > Asp$ .

## II. Méthode cinétique de Cooks pour Gly, Ala, Val, Asp, Glu et Leu

Afin d'évaluer l'ordre relatif des affinités protoniques des acides aminés étudiés dans les systèmes GAVE, GAVD et GAVL, un mélange des monomères Gly, Ala, Val, Asp, Glu et Leu a été préparé. Sur tous les dimères non covalents possibles, seuls les quatre montrés à la Figure A9 ont pu être étudiés.



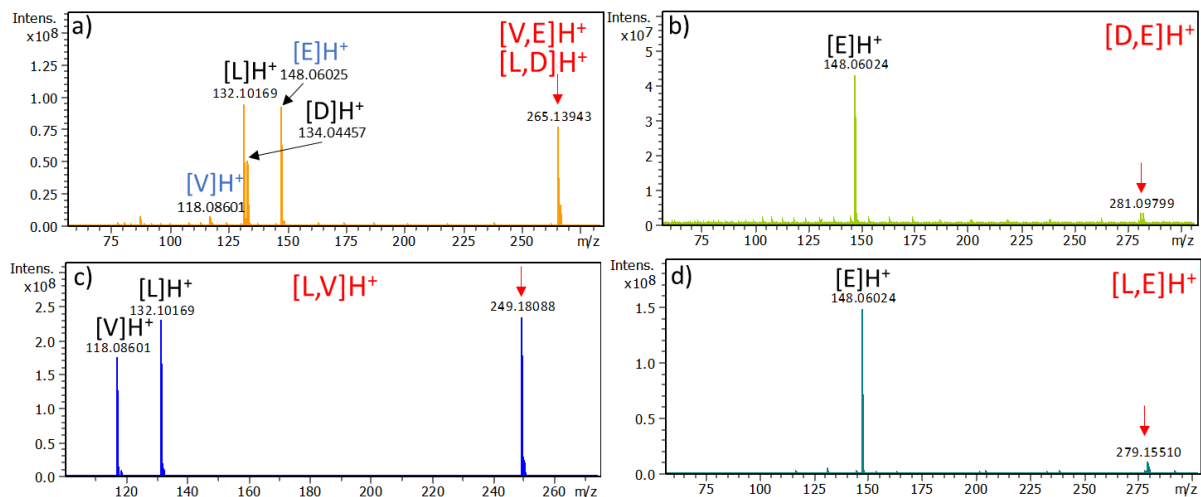


Figure A9. Spectres MS/MS des dimères non covalents.

Les dimères non covalents  $[V,E]H^+$  et  $[L,D]H^+$  ont le même  $m/z$ . Pour une meilleure compréhension, les fragments à comparer sont de la même couleur dans le spectre a. Le spectre b confirme que le marquage isotopique n'influence pas l'efficacité d'ionisation, le spectre MS/MS obtenu pour  $[D,E]H^+$  étant identique à celui obtenu à la Figure A8 pour  $[D,E^*]H^+$ .

Une telle utilisation de la méthode cinétique de Cooks permet d'établir l'ordre des affinités protoniques  $\text{Glu} > \text{Leu} > \text{Val}$ . Il n'est en revanche pas possible de conclure entre les ordres relatifs  $\text{Glu} > \text{Leu} > \text{Val} > \text{Asp}$  ou  $\text{Glu} > \text{Leu} > \text{Asp} > \text{Val}$ .

Quoi qu'il en soit, des réponses analytiques comparables sont attendues entre les peptides contenant ces résidus. En conséquence, exception faite des effets de matrice inhérents aux analyses de mélanges complexes, les analyses des produits de polymérisation doivent refléter *a priori* la composition des systèmes activés.

## Annexe 11. Analyse de l'échantillon (V+E)/SiO<sub>2</sub>

Nous avons testé une polymérisation dans laquelle c'est l'acide aminé aliphatique qui est modifié. La leucine a ainsi été remplacée par la valine, qui comporte un CH<sub>2</sub> de moins dans sa chaîne latérale. Le spectre de masse obtenu pour l'échantillon (V+E)/SiO<sub>2</sub> activé à 160 °C est montré à la Figure A10. Des oligopeptides allant jusqu'aux octapeptides sont détectés, comme dans l'échantillon (L+E)/SiO<sub>2</sub> analysé dans les mêmes conditions. Les deux systèmes montrent donc une réactivité similaire.

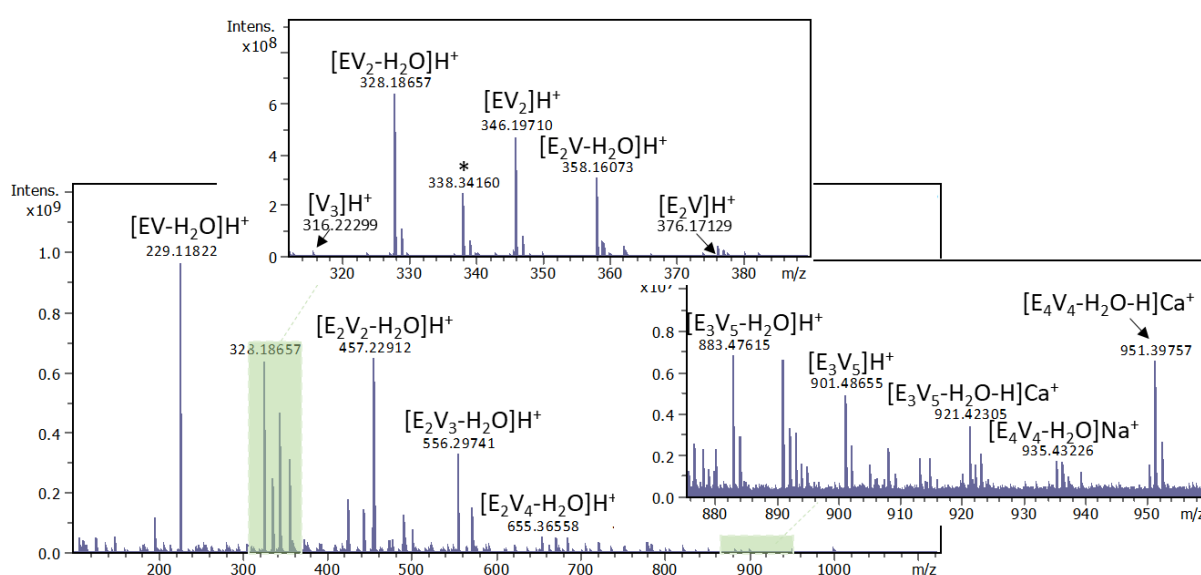


Figure A10. Spectre de masse de l'échantillon (V+E)/SiO<sub>2</sub> activé à 160 °C (ESI positif, instrument FTICR).

## Annexe 12. Etude de l'échantillon (GAVL)/SiO<sub>2</sub>

Le spectre de masse ESI obtenu pour l'échantillon (GAVL)/SiO<sub>2</sub> activé à 160 °C est montré en Figure A11. Seuls des dipeptides et tripeptides sont formés, les autres pics semblent correspondre à des espèces non covalentes. Parmi les dipeptides, les espèces déshydratées sont majoritaires, il s'agit sans doute de DKP.

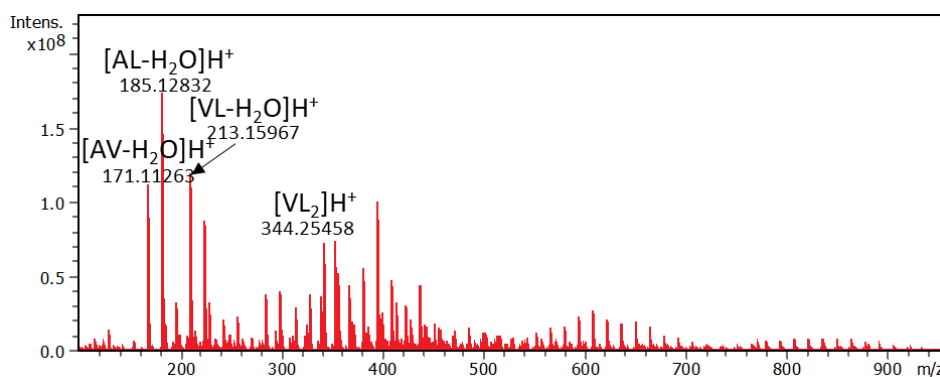


Figure A11. Spectre de masse de l'échantillon (GAVL)/SiO<sub>2</sub> activé à 160 °C (ESI positif, instrument FTICR).

## Annexe 13. Analyse de l'énantiosélectivité dans le système GAVE

Il a été montré au Chapitre III que le système Leu + Glu n'était pas énanti sélectif. L'objectif de cette étude est de s'intéresser à la possibilité d'apparition d'une énanti sélectivité dans des systèmes plus complexes, par exemple dans le système GAVE. Pour s'affranchir d'éventuels effets isotopiques, les deux échantillons  $(G_L+A_L+V_L+E_L+E_D^*)/SiO_2$  et  $(G_L+A_L+V_L+E_L^*+E_D)/SiO_2$  sont comparés.

La Figure A12 et la Figure A13 reportent les massifs isotopiques de différents oligopeptides de plusieurs longueurs observés dans le système GAVE. Les intensités observées ne correspondent pas exactement aux intensités attendues, le pic isotopique de plus haut m/z étant dans tous les cas légèrement plus intense. Cela est en réalité attendu car la quantité naturellement présente de l'isotope  $^{13}C$  dans l'acide aminé non marqué s'ajoute à la contribution de l'espèce due à l'acide aminé marqué. Au vu de ces observations, aucune énanti sélectivité significative n'est observable pour les oligopeptides allant jusqu'aux hexapeptides, ce qui signifie que les formes L et D ont la même efficacité de polymérisation.

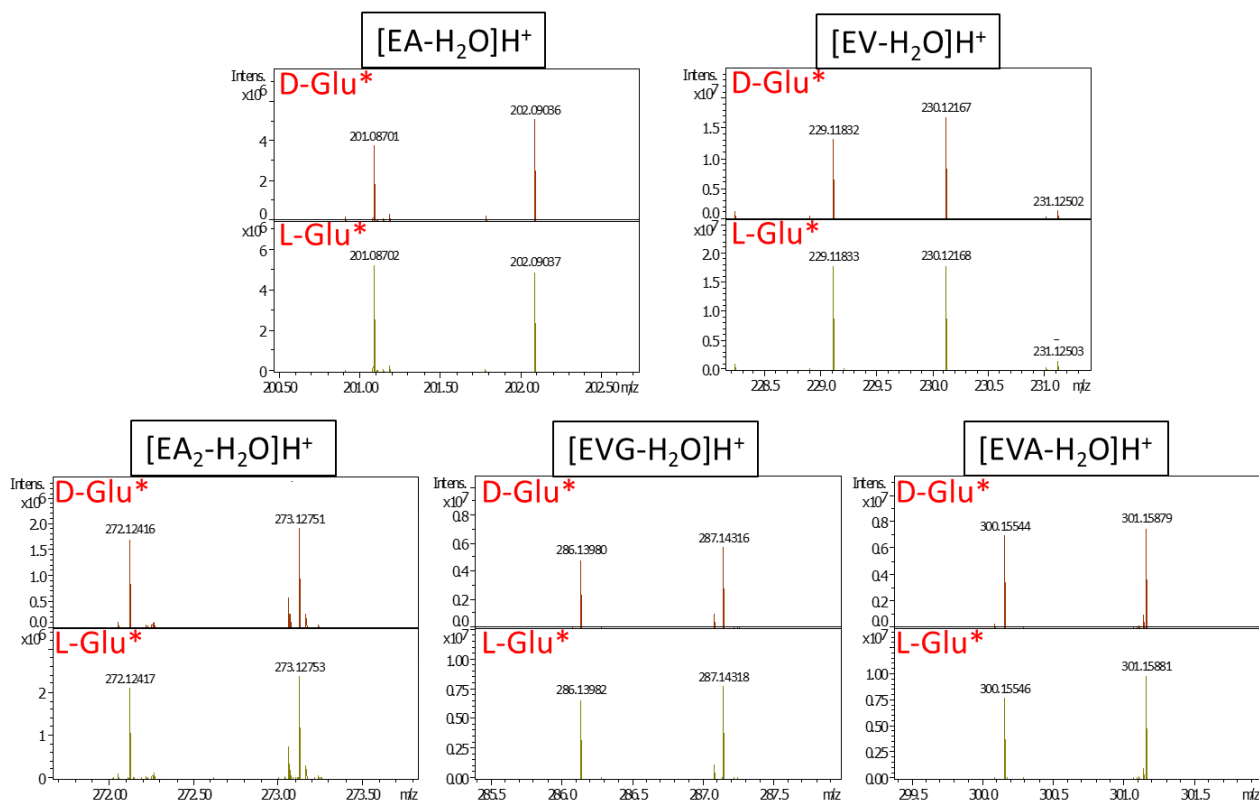


Figure A12. Massifs isotopiques obtenus pour différents dipeptides et tripeptides. Théoriquement les intensités de ces massifs doivent être de 1 : 1.

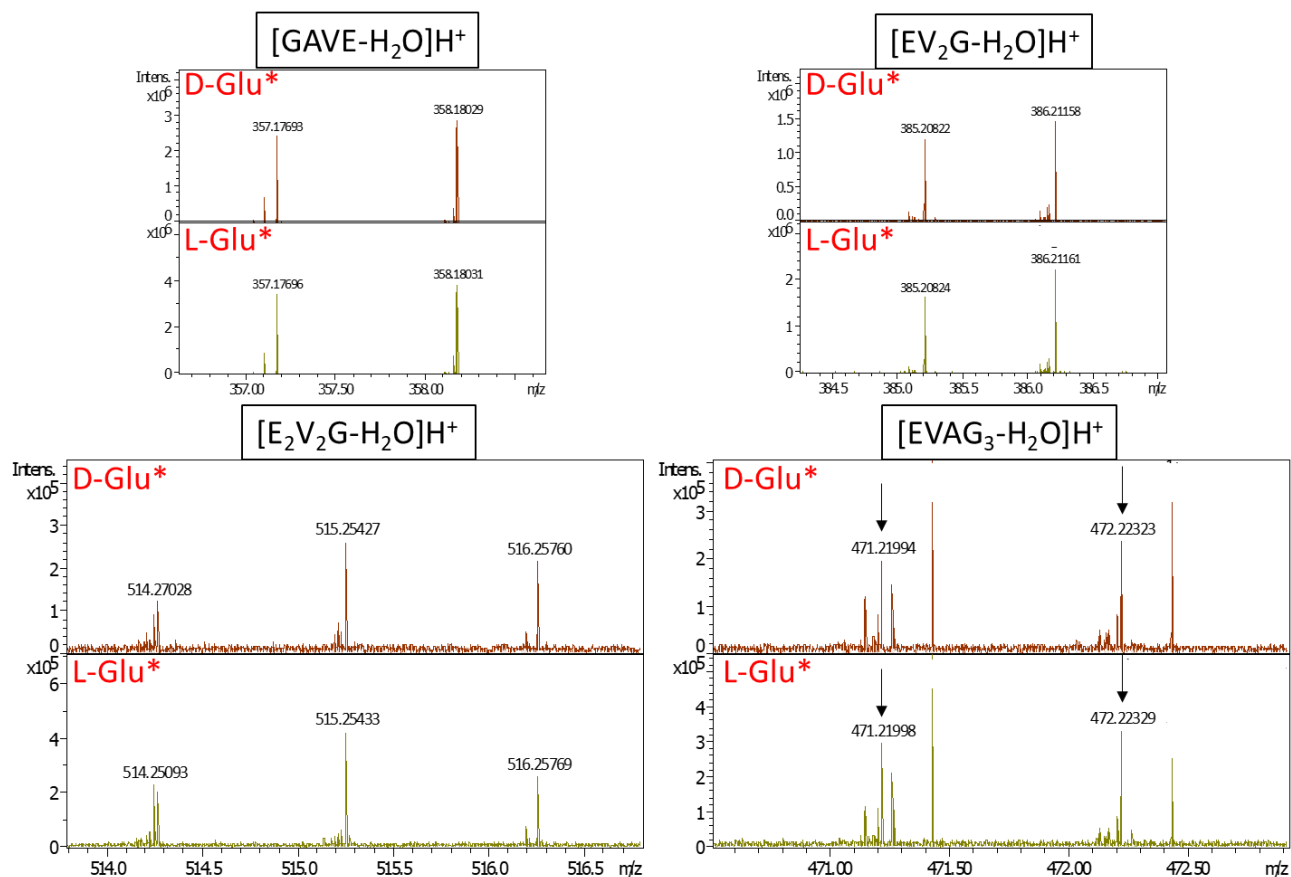


Figure A13. Massifs isotopiques obtenus pour des tétrapeptides, pentapeptides et hexapeptides. Les intensités théoriques attendues pour les massifs de [GAVE-H<sub>2</sub>O]H<sup>+</sup>, [EV<sub>2</sub>G-H<sub>2</sub>O]H<sup>+</sup> et [EVAG<sub>3</sub>-H<sub>2</sub>O]H<sup>+</sup> doivent être de 1 : 1 ; celles de [E<sub>2</sub>V<sub>2</sub>G-H<sub>2</sub>O]H<sup>+</sup> doivent être de 1 : 2 : 1.

## Annexe 14. Synthèse de peptides de référence

Afin d'identifier les composés obtenus par condensation, nous avons synthétisé des oligomères de référence par synthèse peptidique. Parmi toutes les possibilités de synthèse peptidique, nous avons choisi la synthèse peptidique sur support solide (SPPS : Solid Phase Peptide Synthesis). Des concentrations élevées et de larges excès molaires d'acides aminés sont utilisés dans ces conditions.

Cette méthode a été introduite par Merrifield dans les années 1960<sup>3</sup>. Elle consiste en l'allongement de la chaîne peptidique par son côté N-terminal (voir Figure A14), en utilisant des acides aminés dont la fonction  $\alpha$ -NH<sub>2</sub> est protégée par un groupement spécifiquement labile. Le groupement protecteur choisi pour le groupement amine est le fluorenylmethyloxycarbonyl (Fmoc). Sa déprotection se fait avec un traitement basique (ici par la pyridine). Les acides aminés protégés utilisés sont montrés Figure A15

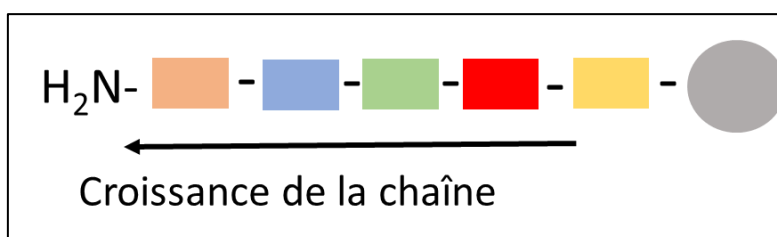


Figure A14. Schéma de l'allongement d'une chaîne peptidique par SPPS

Le support utilisé est une résine 2-chlorotriptyl chloride (2ClTrt). Pour simplification, l'acide aminé en C-terminal du peptide souhaité y est déjà greffé par son côté C-terminal. Les peptides synthétisés étant composés de Leu et de Glu, les deux résines utilisées sont donc H-Leu-2ClTrt et H-Glu-2ClTrt (voir Figure A15).

<sup>3</sup> Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* **1963**, 85 (14), 2149–2154. <https://doi.org/10.1021/ja00897a025>.

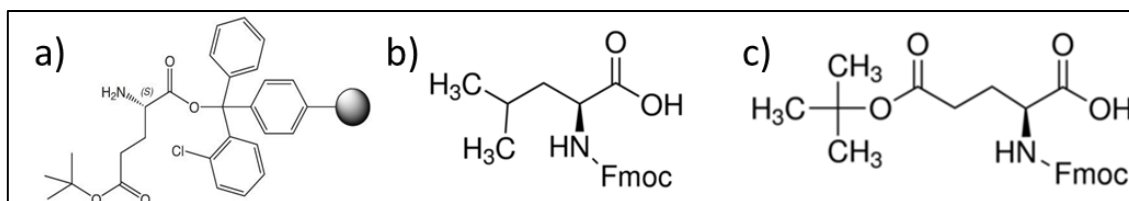


Figure A15. Réactifs utilisés a) résine H-Glu-2ClTrt, b) acide aminé Fmoc-Leu-OH, c) acide aminé Fmoc-Glu(OtBu)-OH

La résine utilisée étant déjà greffée par l'acide aminé voulu en N-terminal, la première étape repose sur le couplage avec l'acide aminé N-protégé suivant via son groupement carboxyl. Après le couplage, l'excès de réactif est éliminé par lavage et les groupements protecteurs sont éliminés de l'extrémité N-terminale du dipeptide, afin de permettre l'addition du troisième acide aminé. Ce processus est répété jusqu'à obtention de la séquence désirée. Une fois la longueur souhaitée, une étape de clivage permet de retirer le peptide du support. Cette technique permet de construire sélectivement et de façon séquentielle le peptide voulu. Le réactif de couplage utilisé est le HBTU. Le DMF est utilisé comme solvant. La solution de déprotection est une solution 80/20 DMF/Pipéridine. Le clivage se fait avec du TFA/eau/TIS à 95/2.5/2.5.

#### Protocole : exemple de la synthèse du térapeptide ELEL.

500 mg de résine H-Leu-2ClTrt sont pesés et placés dans une seringue de 20 mL avec fritté. La résine est ensuite lavée avec 60 mL de DMF. La solution de couplage est préparée dans un pilulier : 0,453 g de Fmoc-Glu-OH et 0,390 g de HBTU sont dissous dans 7 mL de DMF et 184  $\mu$ L de DIPEA sont ajoutés. La solution de couplage est placée sur la résine et l'ensemble est mis à agiter 30 min sur plaque agitatrice.

Après avoir retiré la solution de couplage de la seringue, celle-ci est rincée trois fois au DMF. Afin de vérifier que le couplage a bien fonctionné, un test Kaiser est réalisé. Le test Kaiser est un test colorimétrique permettant de mettre en évidence d'éventuelles amines primaires libres présentes sur les billes de résine<sup>4</sup>. Pour effectuer ce test, quelques billes de résine sont prélevées dans un tube à essai auxquelles sont ajoutées deux gouttes des solutions suivantes sont ajoutées : 2,5 g de ninhydrine dans 50 mL d'éthanol, 80 g de phénol liquéfié dans 20 mL d'éthanol, 2 mL d'une solution à 0,001 mol/L de cyanure de potassium dans 98 mL de pyridine. L'ensemble est agité et chauffé à 100°C pendant trois minutes. Si les billes sont toujours

<sup>4</sup> Kaiser, E.; Colecott, R. L.; Bossinger, C. D.; Cook, P. I. Color Test for Detection of Free Terminal Amino Groups in the Solid-Phase Synthesis of Peptides. *Short Commun.* **1970**, 595–598.

transparentes, le couplage a fonctionné et il est possible de passer à l'étape de déprotection. En revanche si le couplage n'a pas fonctionné, les amines libres réagissent avec la ninhydrine, entraînant une coloration bleue des billes et de la solution. Dans ce cas, la phase solide doit être de nouveau soumise à une réaction de couplage.

La déprotection se fait en remplissant la seringue d'une solution 80/20 DMF/pipéridine. L'ensemble est agité 5 minutes. La solution est jetée puis cette étape est répétée avec une agitation de 10 minutes. Une fois l'étape de déprotection terminée, la résine est rincée trois fois avec du DMF. Le second couplage est ensuite effectué (0,376 g de Leucine et 0,390 g de HBTU dissous dans 7 mL de DMF et 184  $\mu$ L de DIEA), suivi du test Kaiser et de la déprotection. A la fin de la dernière déprotection, la seringue est rincée au DMF et au méthanol puis est séchée toute la nuit dans un dessiccateur sous vide. Le retrait du le peptide du support et la déprotection des groupements tertibutyle se font simultanément en traitant la résine séchée avec 10 mL d'un mélange de TFA/eau/TIS à 95/2.5/2.5. L'ensemble est mis à agiter pendant 3 heures puis transvasé dans un ballon pour pouvoir enlever les composés volatils sous pression réduite avec un évaporateur rotatif. 30 mL d'éther froid sont ajoutés au solide obtenu et le tout est transvasé dans un tube Falcon puis centrifugé 5 min à 2500 rpm. Le surnageant est retiré et de l'éther est réintroduit afin de centrifuger à nouveau. Cette étape est répétée deux fois. Après avoir enlevé le dernier surnageant et fait évaporer le reste d'éther, le peptide sous forme de poudre ou gel aqueux concentré est lyophilisé.



## Annexe 15. Spectres MS/MS du tripeptide $[E_2L]H^+$

La Figure A16 reporte les spectres MS/MS des trois tripeptides de référence de séquence LEE, ELE et EEL et celui de l'espèce de stœchiométrie  $E_2L$  dans le mélange de désorption. Les voies de fragmentation observées pour les espèces obtenues par polymérisation ne correspondent exactement à aucune des références.

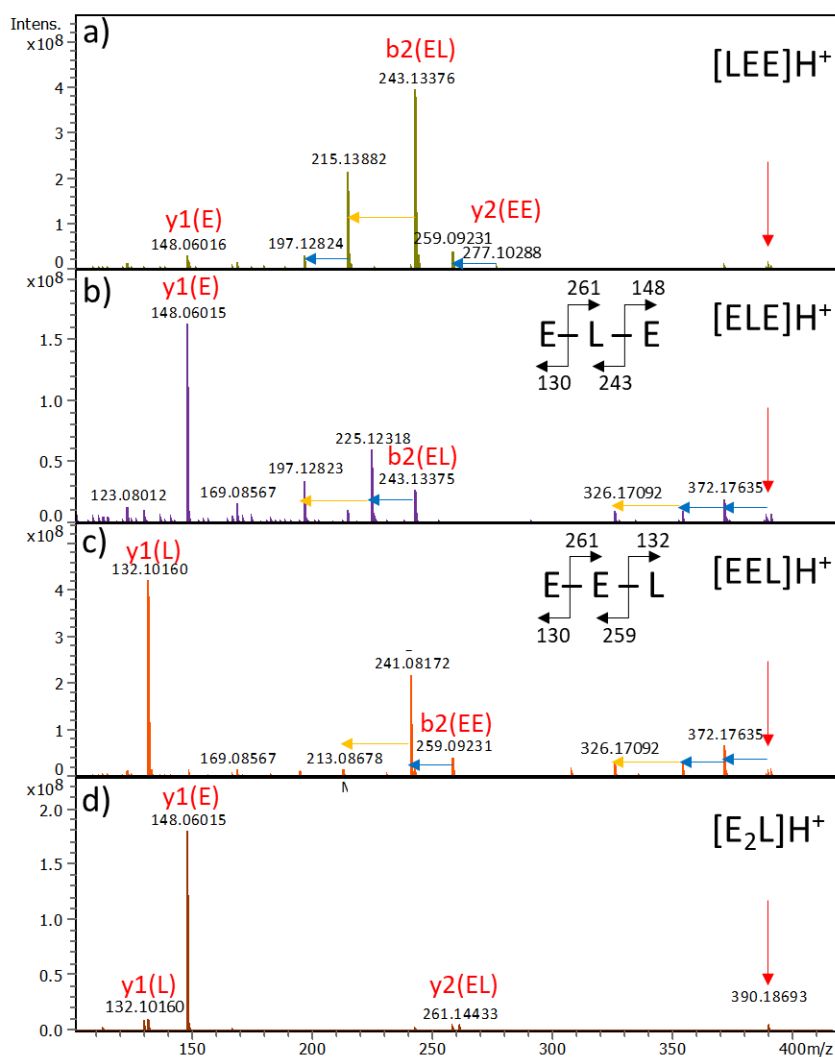


Figure A16. Spectres MS/MS des tripeptides de référence de séquences a) LEE (Elab=12 V), b) ELE (Elab=10 V), c) EEL (Elab=10 V) et d) de l'espèce de stœchiométrie  $[E_2L]H^+$  dans le produit de condensation de l'échantillon  $(L+E)/SiO_2$  (6 V) (ESI positif, instrument FTICR).

## Annexe 16. Analyse IM-MS/MS du térapeptide [EL<sub>3</sub>]H<sup>+</sup>

La Figure A17 reporte les EIMs et les spectres MS/MS de l'espèce de stœchiométrie EL<sub>3</sub> dans le mélange de désorption de l'échantillon (L+E)/SiO<sub>2</sub> activé à 160 °C et des térapeptides de référence de séquence LELL et ELLL.

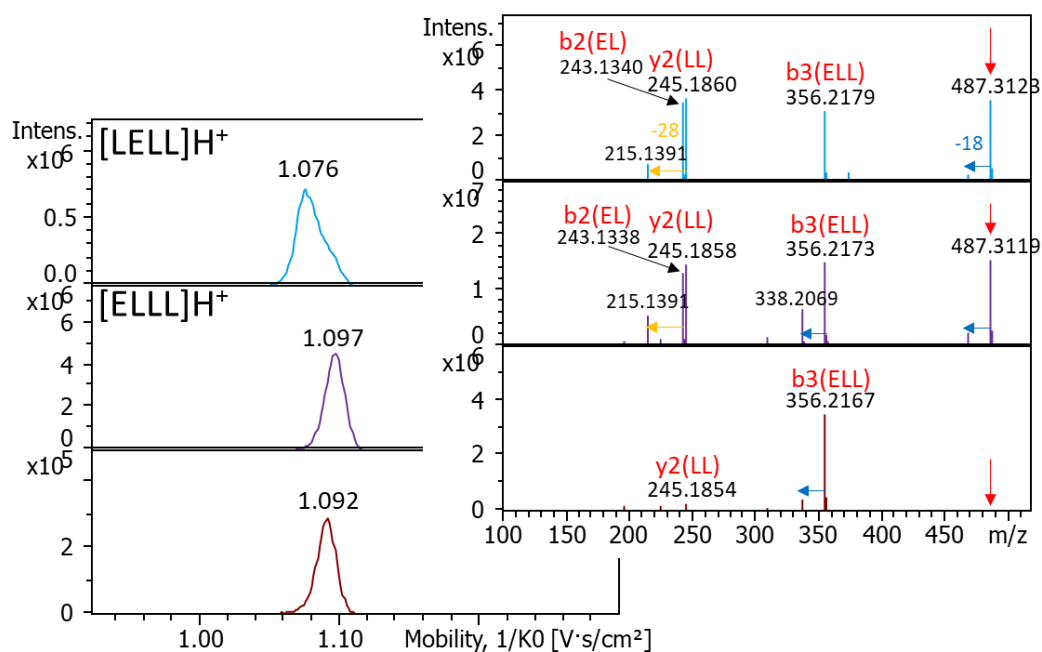


Figure A17. Comparaison des mobilogrammes extraits et des spectres MS/MS de l'espèce de stœchiométrie EL<sub>3</sub> dans l'échantillon (L+E)/SiO<sub>2</sub> activé à 160 °C avec ceux des térapeptides de référence de séquence LELL et ELLL (ESI positif, Elab = 25 eV, instrument TIMS-TOF). Les écarts calculés entre mobilités sont respectivement de 1,48 % et 0,46 % pour LELL et ELLL vis-à-vis de l'ion m/z 487 détecté dans les produits de condensation.

## Annexe 17. Comparaison de l'efficacité d'ionisation d'un peptide intact et du peptide déshydraté correspondant : exemple de ELEL

Afin d'exclure tout effet lié à une différence d'ionisation des deux formes, un mélange équimolaire des formes intactes et déshydratées du peptide ELEL a été préparé. La forme déshydratée a été obtenue par activation thermique de la forme intacte (voir Chapitre VI, § II.2.1.1). Le spectre de masse résultant suggère que les efficacités d'ionisation des deux formes sont similaires. Nous ferons la même hypothèse pour l'ensemble des systèmes testés. Un zoom sur le spectre obtenu est montré à la Figure A18.

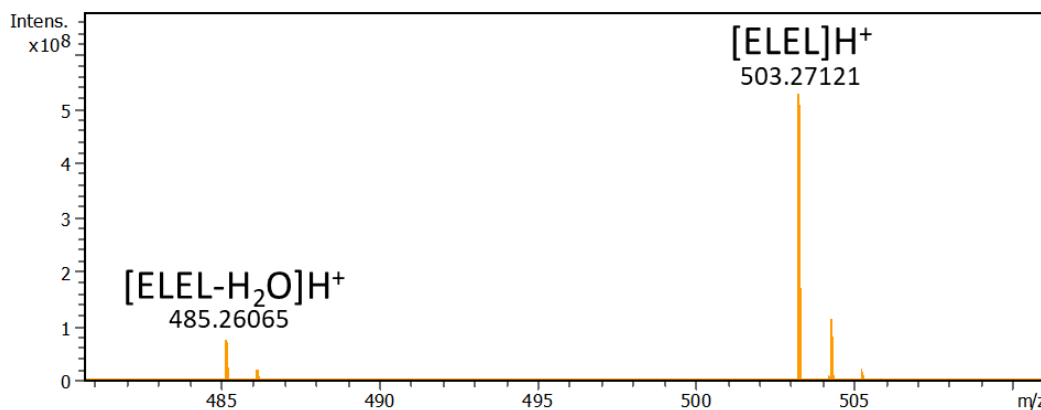


Figure A18. Spectre de masse d'un mélange du tétrapeptide  $[E_2L_2]H^+$  et de son produit d'activation (ESI positif, instrument FTICR).

## Annexe 18. Etude des spectres MS/MS de $[E_2L-H_2O]H^+$

Concernant le tripeptide déshydraté  $[E_2L-H_2O]H^+$ , les fragments observés ne correspondent à aucun des produits d'activation des peptides de références EEL, ELE et LEE, comme le montrent les spectres MS/MS ci-dessous.

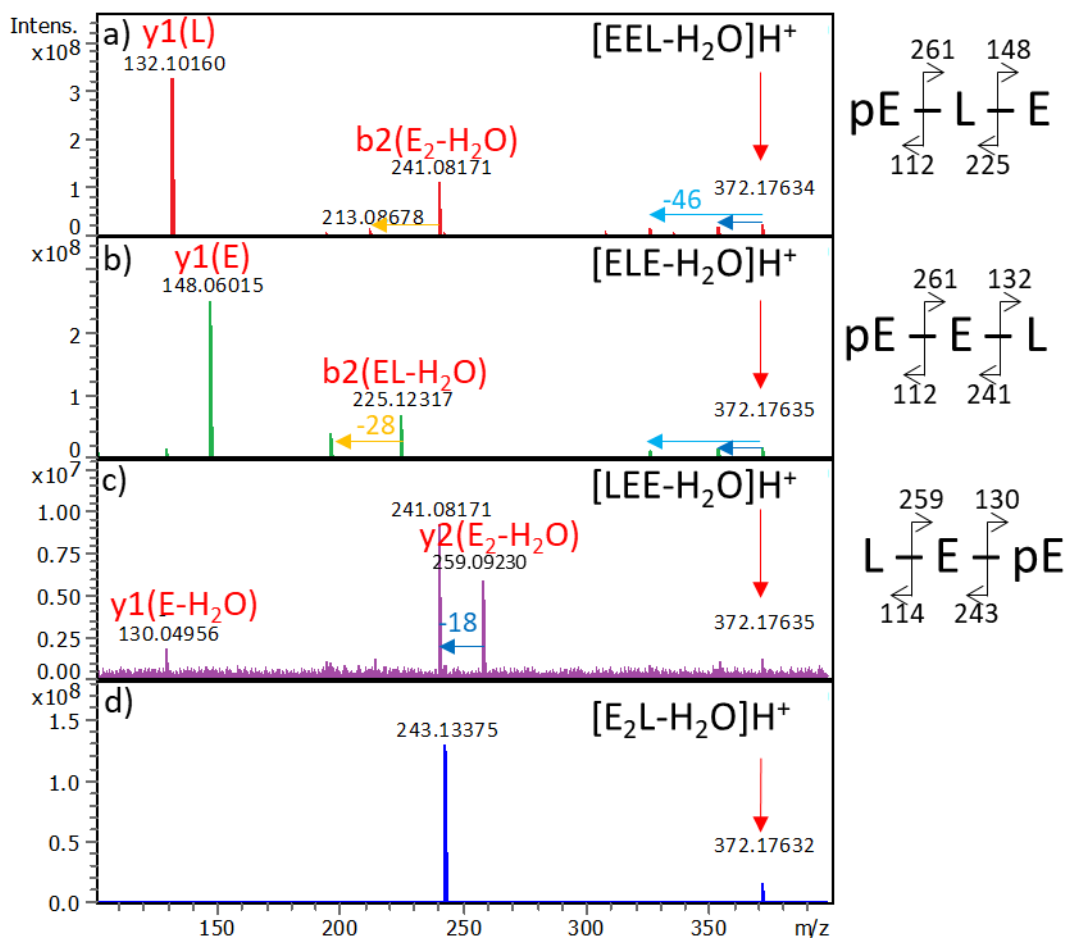


Figure A19. Comparaison des spectres MS/MS des peptides de référence activés avec celui de l'ion  $m/z$  372 dans l'échantillon  $(L+E)/SiO_2$  activé à  $160\text{ }^\circ\text{C}$  : a)  $[EEL-H_2O]H^+$  (8 eV), b)  $[ELE-H_2O]H^+$  8 eV, c)  $[LEE-H_2O]H^+$  16 eV et d)  $[E_2L-H_2O]H^+$  dans l'échantillon étudié (4 eV).

## Annexe 19. Etude des spectres MS/MS des tétrapeptides [EL<sub>3</sub>-H<sub>2</sub>O]H<sup>+</sup> et [E<sub>2</sub>L<sub>2</sub>-H<sub>2</sub>O]H<sup>+</sup>

Deux tétrapeptides obtenus par polymérisation, les ions [EL<sub>3</sub>-H<sub>2</sub>O]H<sup>+</sup> et [E<sub>2</sub>L<sub>2</sub>-H<sub>2</sub>O]H<sup>+</sup>, ont été comparés chacun à un peptide de référence activé, respectivement ELLL-H<sub>2</sub>O et ELEL-H<sub>2</sub>O. Les Figure A20 et Figure A21 reportent les spectres MS/MS correspondant. Dans les deux cas, l'ion présent dans l'échantillon (L+E)/SiO<sub>2</sub> activé à 160 °C montre des voies de fragmentation différentes de celles du peptide de référence activé. Il faut souligner que les peptides de référence activés présentent *a priori* des structures avec pyroGlu à l'extrémité N-terminale.

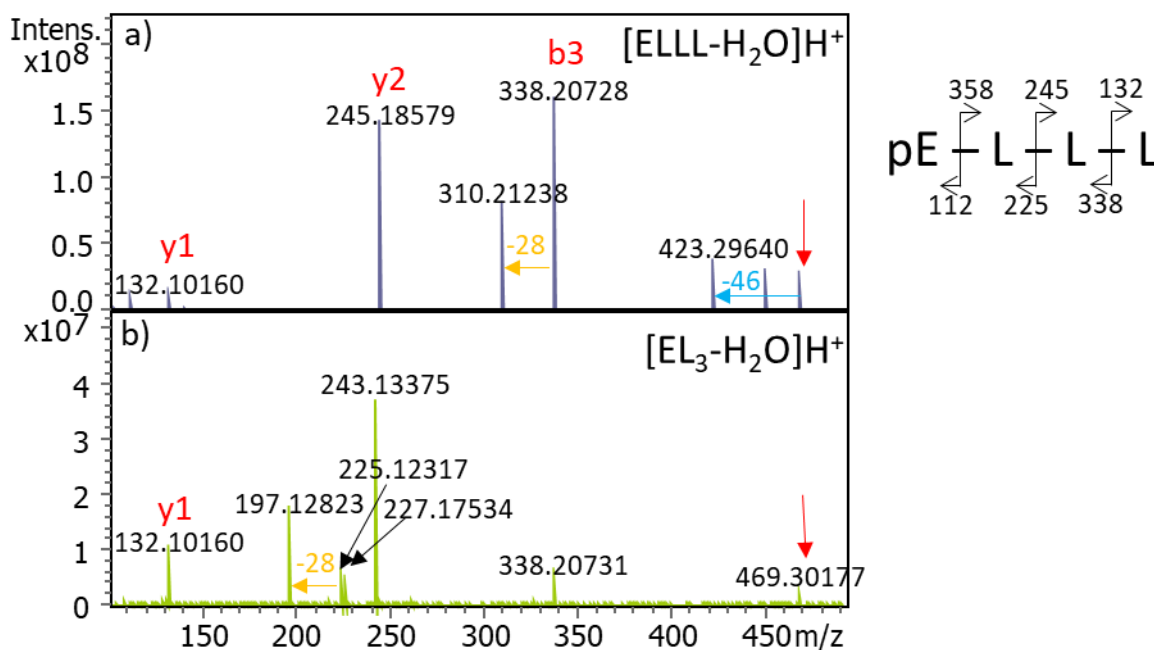


Figure A20. a) Spectre MS/MS du tétrapeptide de référence activé (ELLL-H<sub>2</sub>O). b) Spectre MS/MS de l'ion [EL<sub>3</sub>-H<sub>2</sub>O]H<sup>+</sup> dans l'échantillon (L+E)/SiO<sub>2</sub> activé à 160 °C (ESI positif, Elab = 10 eV, instrument FTICR)

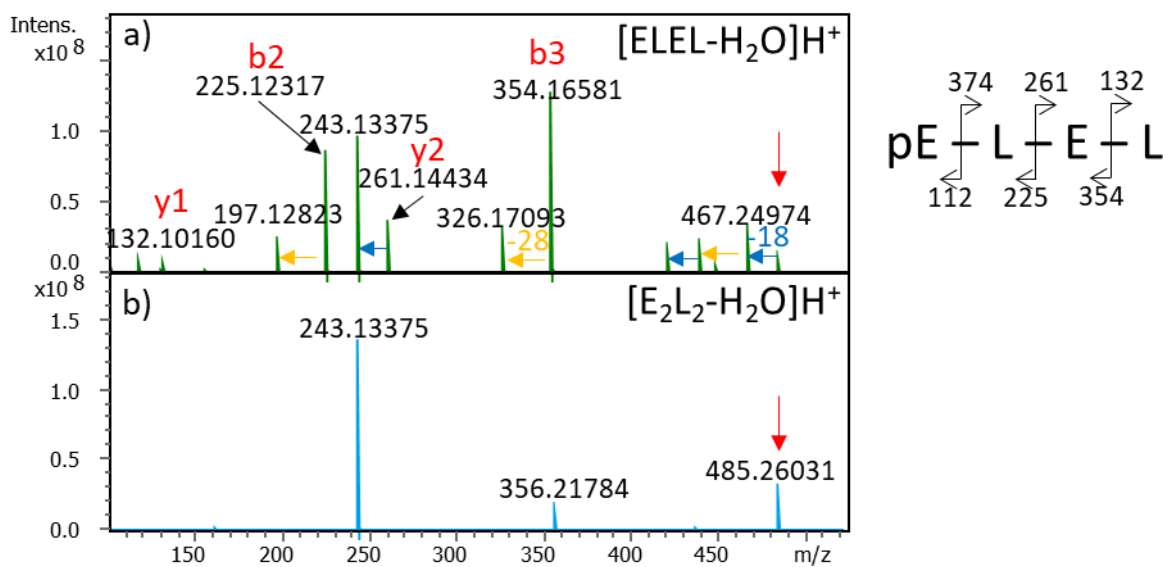


Figure A21. a) Spectre MS/MS du tétrapeptide de référence activé (ELEL- $H_2O$ ) (Elab = 10 eV).  
 b) Spectre MS/MS de l'ion  $[E_2L_2-H_2O]H^+$  dans l'échantillon (L+E)/ $SiO_2$  activé à 160 °C (6 eV)  
 (ESI positif, instrument FTICR).

## Annexe 20. Spectres MS/MS des espèces $[EG]H^+$ , $[EVG]H^+$ et $[EGA]H^+$ du système GAVE

La Figure A22 reporte les spectres MS/MS des espèces  $[EG]H^+$ ,  $[EVG]H^+$  et  $[EGA]H^+$ . Dans les trois cas, les séquences possibles possèdent un résidu Glu localisé en N-terminal.

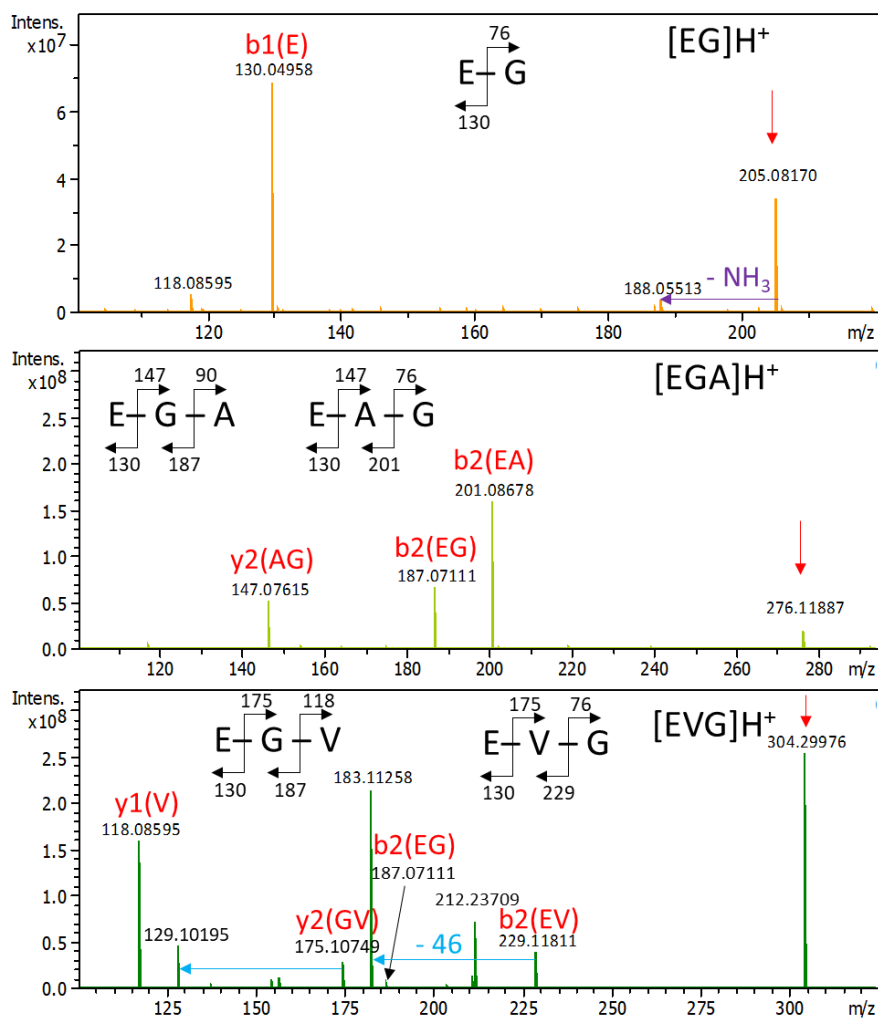


Figure A22. Spectres MS/MS des espèces  $[EG]H^+$ ,  $[EVG]H^+$  et  $[EGA]H^+$  enregistrés respectivement à 6 eV, 4 eV et 14 eV (ESI positif, FTICR).

## Annexe 21. Analyse IM-MS/MS des tripeptides $[D_2A]H^+$ et $[D_2G]H^+$ du système GAVD

La Figure A23 reporte les études IM-MS/MS des tripeptides  $[D_2A]H^+$  et  $[D_2G]H^+$ . Dans les deux cas, les séquences correspondant aux voies de fragmentation observées sont de type DDY, avec Y l'acide aminé aliphatique, démontrant une polymérisation par blocs du côté N-terminal pour ces espèces.

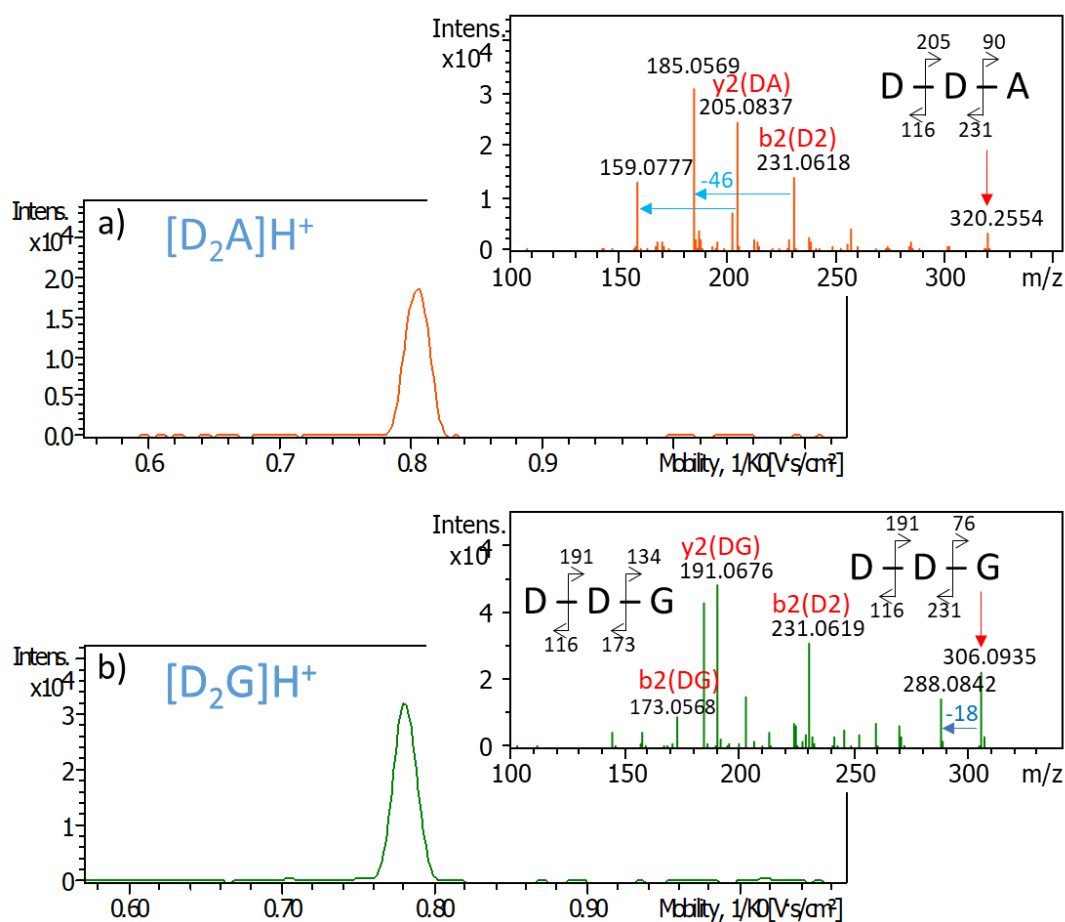


Figure A23. Mobilogrammes extraits et spectres MS/MS des ions  $m/z$  320 correspondant au tripeptide  $[D_2A]H^+$  (a) et  $m/z$  306 correspondant à  $[D_2G]H^+$  (b), obtenus pour l'analyse IM-MS/MS de l'échantillon GAVD/SiO<sub>2</sub> activé à 130 °C (ESI positif, Elab = 20 eV, instrument TIMS-TOF).



## Annexe 22. Mécanismes de fragmentation de $[\beta\text{EL}_2]\text{H}^+$

La Figure A24 montre les voies de fragmentations envisagées pour le tripeptide de séquence L- $\beta$ E-L. La Figure A25 montrent celles pour la seconde forme proposée.

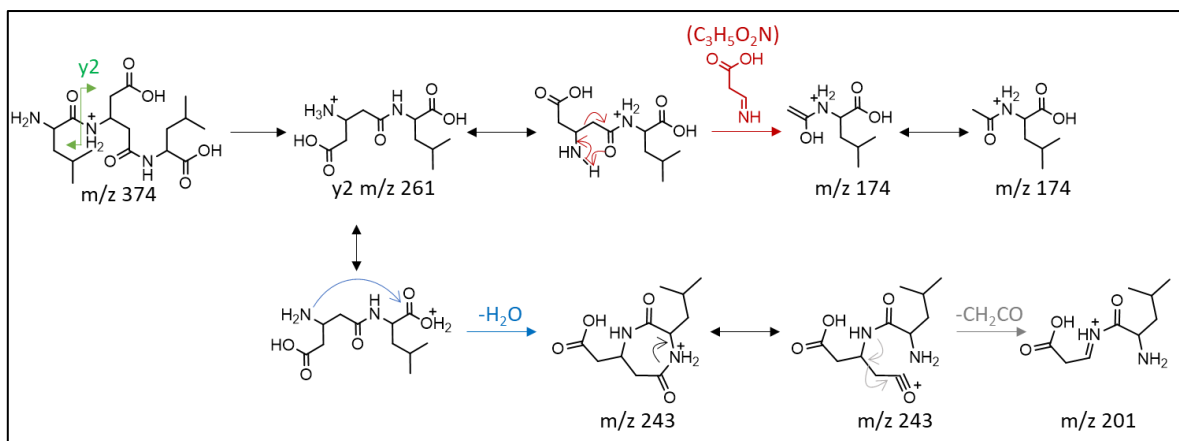


Figure A24. Voies de fragmentations envisagées expliquant les ions fragments observées à partir d'une séquence L- $\beta$ E-L

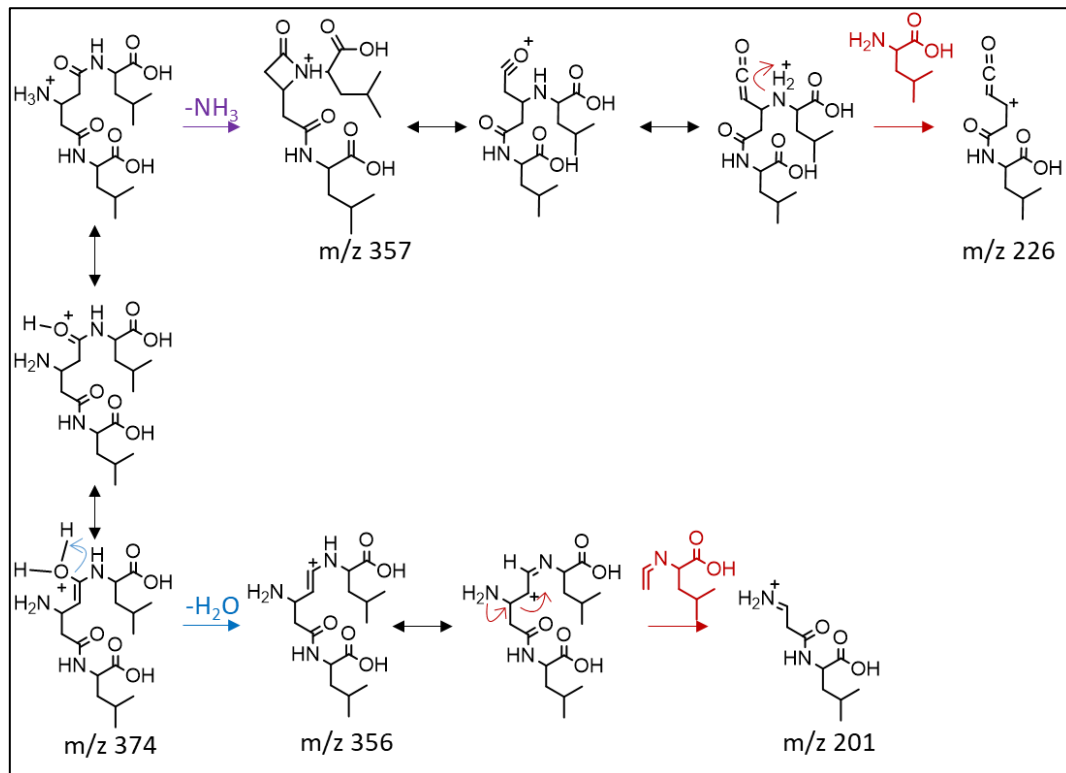


Figure A25. Voies de fragmentations envisagées pour la structure ramifiée correspondant au second pic chromatographique de la LC/MS de l'échantillon (L+ $\beta$ E)/ $\text{SiO}_2$ .

### Annexe 23. Analyse IM-MS/MS de $[\beta E_2 L]H^+$

La Figure A26 montre l'analyse IM-MS/MS du tripeptide  $[\beta E_2 L]H^+$ . D'après le spectre MS/MS obtenu, l'ion présent résulte d'un mélange des deux séquences  $\beta E-\beta E-L$  et  $L-\beta E-\beta E$ .

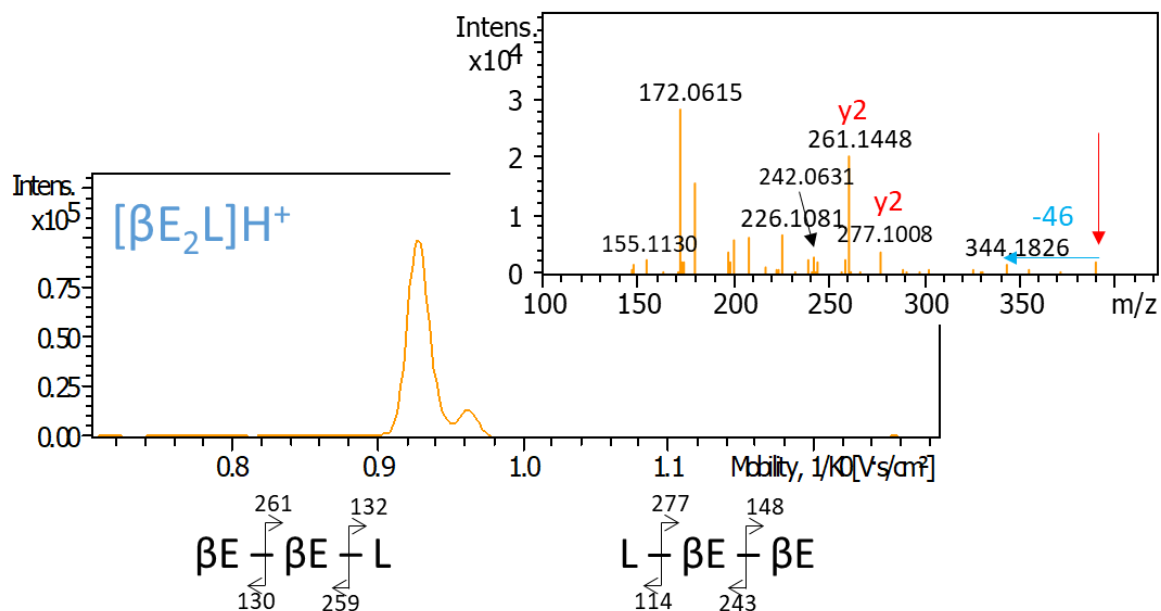


Figure A26. Mobilogramme extrait et spectre MS/MS de l'ion  $m/z$  390 correspondant au tripeptide  $[\beta E_2 L]H^+$ , obtenus pour l'analyse IM-MS/MS de l'échantillon  $(L + \beta E)/SiO_2$  activé à 160 °C (ESI positif, Elab = 35 eV, instrument TIMS-TOF).

## Annexe 24. Spectres MS/MS des dipeptides [DG]H<sup>+</sup>, [DA]H<sup>+</sup>, [DV]H<sup>+</sup> et [DL]H<sup>+</sup>

Les spectres MS/MS acquis sur l'instrument FTICR des dipeptides [DG]H<sup>+</sup>, [DA]H<sup>+</sup>, [DV]H<sup>+</sup> et [DL]H<sup>+</sup> obtenus dans les échantillons GAVD/SiO<sub>2</sub> et (L+D)/SiO<sub>2</sub> sont reportés sur la Figure A27. Les trois premiers montrent simultanément une perte de 46 u et la succession de pertes de 18 et 17 u), suggérant la présence des deux formes  $\alpha$  et  $\beta$ . Concernant le dipeptide [DL]H<sup>+</sup>, les pertes de 46 u et 18 u sont majoritaires (pas de perte consécutive de 17 u).

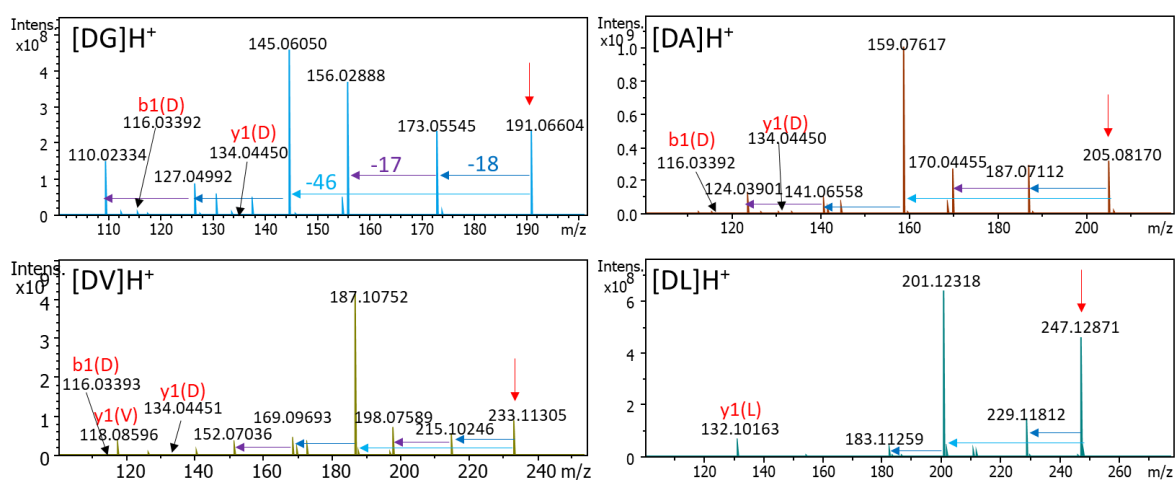


Figure A27. Spectres MS/MS des espèces [DG]H<sup>+</sup>, [DA]H<sup>+</sup>, [DV]H<sup>+</sup> et [DL]H<sup>+</sup> (ESI positif, Elab = 8 eV, FTICR).

## Annexe 25. Analyse IM-MS/MS du tripeptide $[\text{EL}_2]\text{H}^+$

La Figure A28 montre la comparaison des EIMs et spectres MS/MS de l'espèce  $[\text{EL}_2]\text{H}^+$  dans les échantillons (L+Q)/ $\text{SiO}_2$  et (L+E)/ $\text{SiO}_2$  activés à 160 °C. Les voies de fragmentations obtenues pour l'espèce dans les deux systèmes sont légèrement différentes, suggérant la présence probable de régioisomères entre les deux systèmes. Toutefois les fragments détectés sont cohérents avec des séquences ELL et/ou LEL.

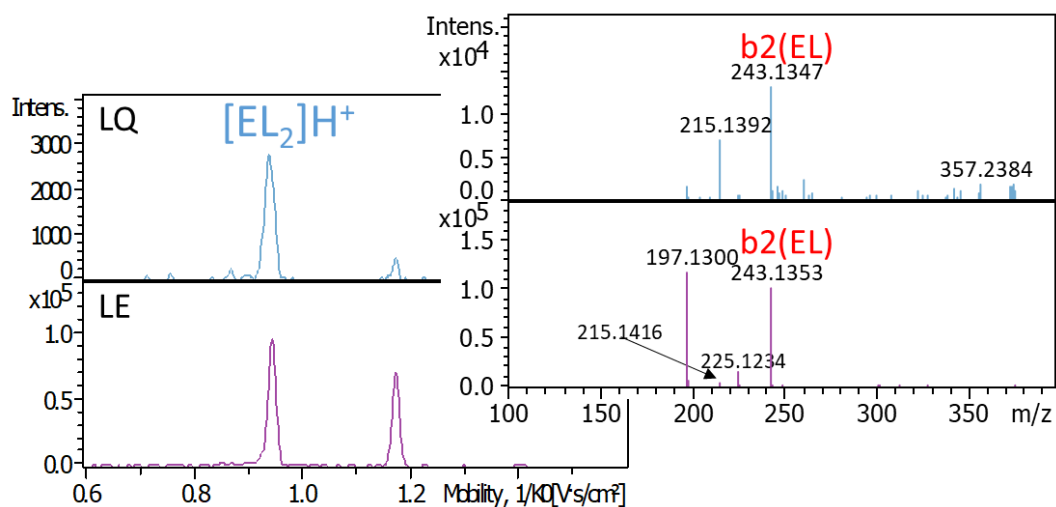


Figure A28. Mobilogramme extrait et spectres MS/MS de l'ion  $m/z$  356 correspondant au tripeptide  $[\text{EL}_2]\text{H}^+$ , obtenus pour l'analyse IM-MS/MS de (L+Q)/ $\text{SiO}_2$  (en haut) et (L+E)/ $\text{SiO}_2$  (en bas) activés à 160 °C (ESI positif, 20 eV, instrument TIMS-TOF). Pour chaque échantillon, seul le premier pic a pu être fragmenté.

## Annexe 26. Spectres MS/MS de $[DVA-H_2O]H^+$ et $[EVA-H_2O]H^+$

L'analyse LC/MS/MS des tripeptides  $[DVA-H_2O]H^+$  et  $[EVA-H_2O]H^+$  montrent que plusieurs formes déshydratées sont présentes pour chaque stœchiométrie. Les spectres MS/MS obtenus en infusion directe sur l'instrument FTICR sont reportés sur la Figure A29. Ils correspondent à un mélange des formes séparées en LC/MS/MS (voir Chapitre VI).

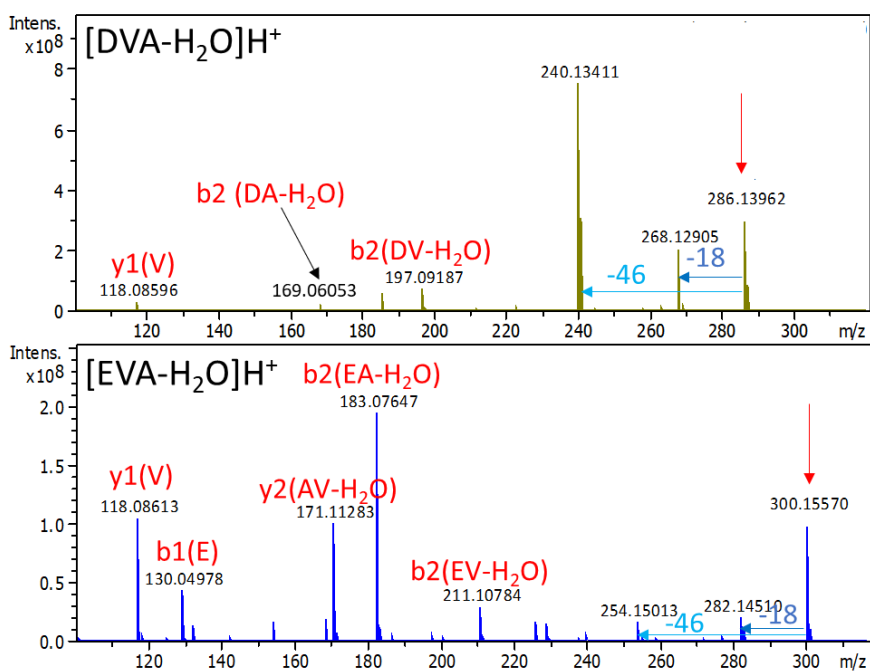


Figure A29. Spectres MS/MS des espèces  $[DVA-H_2O]H^+$  et  $[EVA-H_2O]H^+$  enregistrés respectivement à 6 eV et 10 eV (ESI positif, FTICR).

## Annexe 27. Analyse IM-MS/MS du tripeptide $[DL_2-H_2O]H^+$

La correspondance entre les formes intactes et déshydratées a également été vérifiée pour le système Leu + Asp. Trois pics distincts, montrant des voies de fragmentation différentes, sont observés sur l'EIM de l'ion  $m/z$  342 correspondant au tripeptide  $[DL_2-H_2O]H^+$  (voir Figure A30 où les structures envisagées sont reportées). La première correspond à une cyclisation de type DKP comme suggéré par la perte de 45 u à partir de l'ion moléculaire. Les deux autres correspondent à des cyclisations de type succinimide. Ces trois formes correspondent à des formes déshydratées des trois séquences de l'espèce intacte précédemment déterminées au Chapitre VI.

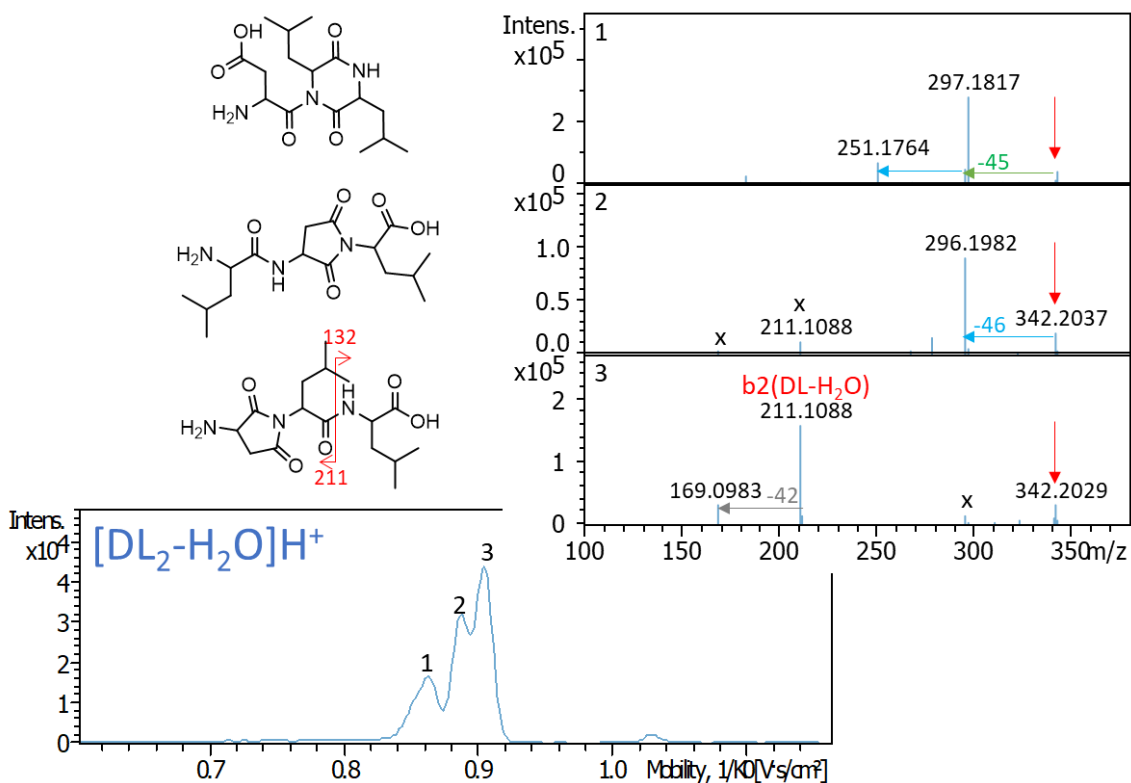


Figure A30. Mobilogramme extrait et spectres MS/MS de l'ion  $m/z$  342 correspondant au tripeptide  $[DL_2-H_2O]H^+$ , obtenus pour l'analyse IM-MS/MS de (L+D)/SiO<sub>2</sub> activé à 160 °C (ESI positif, 20 eV, instrument TIMS-TOF). Il a été vérifié que les pics notés d'une croix ne correspondent pas au spectre MS/MS considéré.